Multilateral methods and product specification

Claude Lamboray*

Paper prepared for the ISI 63th World Statistics Congress

Ottawa, Canada

July 16, 2023 - July 20, 2023

Abstract

Before using multilateral methods with scanner data, one needs to specify the product which is the unit for which a price is observed. Product specification has been recognized as a critical step that can have a significant impact on the results obtained with multilateral methods. Tightly specified products may cause a bias because there may not be sufficient product matches over time. Broadly specified products may cause a bias because there may be quality differences between the underlying transactions that make up the product. In order to assess the biases that result from using either a tight or a broad product specification, we consider the imputation Fisher index as the target index, and compare this index to matched and hybrid Fisher indices. We also define an alternative target index that we call a quality adjusted hybrid Fisher index. We then extend the analysis defined in the bilateral context to the multilateral context of GEKS-type indices. The approach is illustrated on scanner data sets for clothing and hygienic products. This type of analysis should help the compiler deciding which index formula and which product specification to use in practice.

Keywords — Scanner Data, Multilateral Methods, Product Specification

1 Introduction

The use of scanner data in a Consumer Price Index (CPI) gives rise to an aggregation problem that can be subdivided into three stages.

1. In the first stage, individual transactions are combined into an individual product for which an average price and a total quantity sold can be calculated.

*STATEC, Responsable Méthodologie, claude.lamboray@statec.etat.lu. This is an update of a previous version of this paper that was prepared while the author was working at Eurostat as a Seconded National Expert and that was presented at the 17^{th} meeting of the Ottawa Group, Rome, Italy, June 2022. I also gratefully acknowledge the helpful comments provided by Lucien May.

2. In the second stage, the previously specified individual products are further aggregated using, for example, a multilateral method, in order to obtain an elementary price index.
3. In the third stage, the previously calculated elementary price indices are further combined with other price indices in order to obtain price indices for the higher-level aggregates in the CPI.

The second stage has received a lot of attention from a theoretical and practical point of view. There are many studies on the properties, results and implementation issues of different bilateral and multilateral index formulas. Very often these studies rely on the assumption that the individual products together with their prices and quantities are given. However, any index compilation method applied in the second stage is conditional on the product specification applied in the first stage. This will be the focus of this paper.

Product specification has been recognized as a critical step that could jeopardize any gains in bias reduction that we would typically expect from using scanner data. While scanner data helps reducing lower-level substitution bias, other biases can appear because products are specified too tightly or too broadly (see section 3.1.2 in European Central Bank, 2021 [5]). Under certain pricing strategies, products may enter or exit with unusually high or low prices. It is known that this creates biases in traditional matched model methods (see Eurostat, 2021 [10]), and these biases propagate to scanner data and multilateral methods (see Melser and Webster, 2021 [6]). Konny et al., 2019 [14] stresses that multilateral index methods do solve the problem of chain drift but they are not fully satisfactory to cope with life-cycle pricing. For example, the multilateral methods do not solve the downward drift in the price index for new vehicles caused by the downward price trends of a given model year (see Williams et al., 2019 [4]).

Technically, tightly specified products may cause a bias as new and disappearing products in the two comparison periods are not taken into account in a matched price index. Broadly specified products may cause a bias as the underlying transactions that make up the product may not be of the same quality. This trade-off has been referred to as assignment bias versus assortment bias (Von Auer, 2017 [19]). This trade-off between homogeneity and stability over time has also been highlighted by Chessa, 2019 [3] who developed the MARS method as an operational tool for finding a compromise between these two objectives.

However, MARS is model agnostic in the sense that it does not take into account the index compilation (e.g. a multilateral method) that is applied to the data. In this paper, we examine the trade-off between tight and broad product specifications in the framework of a specific index formula, namely the Fisher price index. In order to assess the biases that result from using either a tight or a broad product specification, we define a target index that takes into account quality adjustment. Finally, we extend the analysis defined in the bilateral context to the multilateral context using the GEKS framework.

This paper is organized as follows. In section 2, we first discuss the dimensions of product specification. In section 3, we formalize the problem from an index number point of view. In section 4, we compare three types of indices in order to assess the matched-model bias and the unit value bias. In section 5, we introduce another target index that is based on quality adjusted average prices. Section 6 looks at the imputation methods that can be used. In section 7, the framework is extended from the bilateral to the multilateral case. The analysis is illustrated on two data sets in section 8. Finally, we draw some conclusions in section 9.

2 Product specification

In order to calculate a price index with scanner data, it is necessary to specify beforehand the individual product (see chapter 3 in Eurostat, 2022 [11]).An earlier discussion on this topic can be found in Dalèn, 2017 [16]. Conceptually, a single transaction specified by the product characteristics, the timing and place of purchase and the terms of supply is the most granular unit for which a price can be observed. In practice, we do not work with single transactions, but with individual products. The individual product is the statistical unit which is tracked over time and which corresponds to the input of, for example, a multilateral method.

When specifying individual products, one needs to consider the time, outlet and product dimensions. An average price (unit value) is calculated over days or weeks of the reference period, over outlets and possibly over item codes. 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. Alternatively, it can be defined in broader terms, for example comprising several item codes sold in several outlets for a given time period. The specification of the individual product is a critical step which can have a significant impact on the final index.

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. At the same time, there is a limit to this strategy. In principle, transacted products can only be combined as long as there are no significant quality differences between them. Quality differences must be evaluated with respect to the already mentioned time, outlet and product dimensions.

The treatment of the **time dimension** is the least controversial. In general it is appropriate to calculate a unit value when an item is sold at different prices at different times within the same month[1].

[1]For some products, such as accommodation or transport services, the timing is an important quality dimension. Traveling on a Friday evening may be considered as a different product (i.e. a different quality) from traveling on a Wednesday afternoon. The price may also depend on the moment of purchase. In such a context, one could argue to treat differences in the time of supply of the service, and possibly differences in the time of booking of the service, as differences in quality.

Ideally the average price should cover as much as possible of the reference month. Diewert, Fox and de Haan, 2016 [12] showed that aggregation over only one week of the month can be upward biased compared to aggregation over the full month. Production calendar constraints explain that in practice, very often only the first two or three weeks of the reference month are used in a CPI.

The treatment of the **outlet dimension** depends on circumstances. The individual product could be specified at the most detailed outlet level available in the data. Quality differences between outlets can be associated with different opening hours, different assortments, etc. Aggregation across outlets can be envisaged if data is only supplied at a more aggregated level or if for example price levels are similar in the outlets (for example of the same type and chain). The impact of the outlet dimension can be empirically assessed (see Ivancic and Fox, 2013 [17]). The outlet dimension is also examined in Azaircabe, 2022 [1]. This paper looks at unit value aggregation over different providers that offer on the one hand ride sharing services, and on the other hand taxi ride services.

The treatment of the product dimension is the most controversial one. The scanner data usually includes an item code such as a Global Trade Identification Number (GTIN) or a slightly broader Stock-Keeping Unit (SKU) code. In general, there is some product churn, meaning that the set of item codes is not stable over time. There are different strategies that can be used to cope with a dynamic product universe.

1. **Matching**. In many cases, it could be satisfactory to define the individual product at the GTIN or SKU level. With such a strategy, item codes are taken into account if they are available in two comparison periods. However, this approach is not satisfactory to capture relaunches. The approach is prone to downward biases if the pricing strategy depends on the life-cycle of a product. For example, it may happen that the last available price of the an item is a reduced price. This situation can be encountered at the end of a sales period and is especially common in clothing and footwear. Reduced prices can also be observed in situations of inventory clearing or closure of an outlet.
2. **Grouping** The item codes are combined into broader products, thereby reducing the lack of matching across time. However, this may create other problems as item codes may be grouped together which are not of the same quality.
3. **Imputation** In order to take into account the item codes that are not available in the two comparison periods, a price is explicitly imputed for these products in the periods in which they are not available. This allows then to estimate a price change for these unmatched item codes.

In principle, the matching approach is the easiest approach to apply. It only requires an appropriate product identifier whereas some kind of product characteristics are usually needed in order to group items together or to estimate a price of a missing item.

The grouping approach is relatively easy to explain. Sometimes the supplied data is already grouped and a more disaggregated approach is not possible. However, often, the supplied data can be grouped in various ways. This will make it possible to compare the index obtained from the item codes with the index obtained after grouping some of the item codes. It leads to the practical question of which one of the two approaches is most suitable. Grouping has been proposed by Chessa as a basis for processing scanner data in the Dutch CPI ([2]). Sometimes the grouping is implemented by first transforming the items of the same group into a common unit. This is the case in the French CPI where items are combined into equivalence classes by explictly using the volume of each item [21].

The imputation approach can be considered as a valid approach from an index number perspective. Practical implementations of such an approach in a multilateral context have been proposed by de Haan and Krsinich (2014) [9] and de Haan and Daalmans (2019) [8]. The practical challenge with this approach is that we need to estimate the missing prices.

Another approach referred to *linking* is sometimes adopted. Under this approach, a disappearing item is *linked* with an new item, thereby capturing the price change between these two items . Such one-to-one replacements are especially meaningful under a fixed-basket framework. If an item of the basket is missing, a replacement item is selected and a price comparison is made, possibly with a quality adjustment. However, one-to-one replacements are less natural in a dynamic product universe and ad-hoc procedures must be defined to implement such linkings[2].

In this paper, we will use the approach that require explicit imputations as benchmark that can be compared to matching and grouping strategies.

3 Fisher-type price indices

If both prices and quantities are available, a price index formula should be used that relies on the weights in the two comparison periods. In this study we will focus on a Fisher index. A Fisher index has good axiomatic properties and is consistent with a basket approach. In fact, it is defined as an average of two basket indices that rely on either base or current period quantities. Finally, from an analytical point of view, the Fisher index can be more easily related to and combined with unit values, which will be a key element in the analysis. The Fisher index is also the basis for some of the multilateral methods, such as GEKS.

We suppose that the 'item' is the most granular product identification available in the data (for example a GTIN code). We denote by p_i^t and q_i^t the price, and the quantity of the item i in period t. Let N_t be set of items available in period t. The set of items in the two comparison periods 0 and 1 is

[2]See for example Daalmans 2022 [15]) for a practical example on how such linking could be defined.

denoted by $M_{01} = N_0 \cap N_1$. Moreover, let $N_{01} = N_1 \backslash N_0$ be the set of items available in period 1 but not in period 0, and let $D_{01} = N_0 \backslash N_1$ be the set of items available in period 0 but not in period 1.

In addition, we suppose that broadly comparable items can be grouped together into 'broader product groups' (BPG). let H_k be set of items that belong to the BPG k. The average price and total quantity of the BPG k in period $t(t = 0, 1)$ can be derived from the initial data of items as follows:

$$\bar{p}_k^t = \frac{\sum_{i \in H_k} p_i^t q_i^t}{\sum_{i \in H_k} q_i^t} \tag{1}$$

$$Q_k^t = \sum_{i \in H_k} q_i^t \tag{2}$$

As a starting point, we calculate a Fisher index on the matched item codes. This means that the aggregate price change is only derived from the set of items that are available in the two comparison periods. The matched Laspeyres, Paasche and Fisher indices between periods 0 and 1 are defined as follows:

$$P_{ML}^{01} = \frac{\sum_{i \in M_{01}} p_i^1 q_i^0}{\sum_{i \in M_{01}} p_i^0 q_i^0} \tag{3}$$

$$P_{MP}^{01} = \frac{\sum_{i \in M_{01}} p_i^1 q_i^1}{\sum_{i \in M_{01}} p_i^0 q_i^1} \tag{4}$$

$$P_{MF}^{01} = \sqrt{P_{ML}^{01} P_{MP}^{01}} \tag{5}$$

One issue with the matched Fisher index is that items that are available in only one of two comparison periods are ignored. In order to overcome this limitation, we could estimate a price for an item in the period in which it is not available. We denote by $\hat{p}_i^t$ an estimated (i.e. not observed) price of an item in period $t(t = 0, 1)$. The imputation Laspeyres, Paasche and Fisher indices between periods 0 and 1 are defined as follows (see de Haan,2001 [7]):

$$P_{IL}^{01} = \frac{\sum_{i \in M_{01}} p_i^1 q_i^0 + \sum_{i \in D_{01}} \hat{p}_i^1 q_i^0}{\sum_{i \in M_{01}} p_i^0 q_i^0 + \sum_{i \in D_{01}} p_i^0 q_i^0} \tag{6}$$

$$P_{IP}^{01} = \frac{\sum_{i \in M_{01}} p_i^1 q_i^1 + \sum_{i \in N_{01}} p_i^1 q_i^1}{\sum_{i \in M_{01}} p_i^0 q_i^1 + \sum_{i \in N_{01}} \hat{p}_i^0 q_i^1} \tag{7}$$

$$P_{IF}^{01} = \sqrt{P_{IL}^{01} P_{IP}^{01}} \tag{8}$$

The imputation indices solve the lack of matching from which matched indices may suffer but it requires an estimation of the prices. An alternative strategy to increase the matching would be to first combine the initial items and create BPG.The index formula is then applied to these BPGs, instead of applying it to the initial items. Following the terminology used in Diewert, 2010 [13], we will refer to this as hybrid indices. The hybrid Laspeyres, Paasche and Fisher indices are defined as follows:

$$P_{HL}^{01} = \frac{\sum_k \bar{p}_k^1 Q_k^0}{\sum_k \bar{p}_k^0 Q_k^0} \tag{9}$$

$$P_{HP}^{01} = \frac{\sum_k \bar{p}_k^1 Q_k^1}{\sum_k \bar{p}_k^0 Q_k^1} \tag{10}$$

$$P_{HF}^{01} = \sqrt{P_{HL}^{01} P_{HP}^{01}} \tag{11}$$

Technically, these hybrid indices are defined for 'matched' BPGs that are available in the two comparison periods.

4 Matched model bias and unit value bias

We consider two nested options for specifying the product. On the one hand we have a tight product specification that is based on the item. On the other hand, we have a broad product specification based on BPGs which are obtained by grouping together the initial items. Tightly specified products may cause a bias as new and disappearing items in the two comparison periods are not taken into account in a matched price index. Broadly specified products may cause a bias as the underlying items that are grouped together may not be of the same quality. Our objective is to evaluate the two product specifications and find out which one works best. To do so, we will estimate matched model bias and unit value bias.

We consider the imputation Fisher index as the target index. Subject to a given imputation model, we quantify the matched model bias by comparing the matched index with an imputation index.

$$b_{MM}^{01} = \ln\left(\frac{P_{MF}^{01}}{P_{IF}^{01}}\right) \approx \frac{P_{MF}^{01}}{P_{IF}^{01}} - 1 \tag{12}$$

Subject to a given imputation model, we quantify the unit value bias by comparing the hybrid index with an imputation index.

$$b_{UV}^{01} = \ln\left(\frac{P_{HF}^{01}}{P_{IF}^{01}}\right) \approx \frac{P_{HF}^{01}}{P_{IF}^{01}} - 1 \tag{13}$$

It follows that the difference between the matched and hybrid index can be explained by these two biases:

$$\ln\left(\frac{P_{HF}^{01}}{P_{MF}^{01}}\right) = \ln\left(\frac{\frac{P_{HF}^{01}}{P_{IF}^{01}}}{\frac{P_{MF}^{01}}{P_{IF}^{01}}}\right) = b_{UV}^{01} - b_{MM}^{01} \tag{14}$$

In order to compare the matched and imputation index, we will introduce some additional notations. Let us define the following factor based on the total expenditure in the two comparison periods of all, or only the matched items.

$$\Delta^{01} = \frac{\left(\sum_{i \in M_{01}} p_i^1 q_i^1\right) / \left(\sum_{i \in M_{01} \cup N_{01}} p_i^1 q_i^1\right)}{\left(\sum_{i \in M_{01}} p_i^0 q_i^0\right) / \left(\sum_{i \in M_{01} \cup D_{01}} p_i^0 q_i^0\right)} \tag{15}$$

Moreover, let us define the following parameters that measure the impact of the imputed prices for the new or disappearing items.

$$\Pi^{01} = \frac{\left(\sum_{i \in M_{01}} p_i^0 q_i^1 + \sum_{i \in N_{01}} \hat{p}_i^0 q_i^1\right) / \left(\sum_{i \in M_{01}} p_i^0 q_i^1\right)}{\left(\sum_{i \in M_{01}} p_i^1 q_i^0 + \sum_{i \in D_{01}} \hat{p}_i^1 q_i^0\right) / \left(\sum_{i \in M_{01}} p_i^1 q_i^0\right)} \tag{16}$$

It can be shown that the matched model bias[3] can be decomposed into a component that is based on the observed prices (see equation 15), and a component that is based on the imputed prices (see equation 16).

$$b_{MM}^{01} = 0.5 \ln \left(\Delta^{01}\right) + 0.5 \ln \left(\Pi^{01}\right) \tag{17}$$

If there are no new and disappearing items, the second term vanishes and the first term collapses to zero. More interestingly, suppose that the missing prices are imputed using a fixed basket index of the matched items:

$$\hat{p}_i^1 = p_i^0 \frac{\sum_{i \in M_{01}} p_i^1 q_i^0}{\sum_{i \in M_{01}} p_i^0 q_i^0} \qquad \forall i \in D_{01} \tag{18}$$

$$\hat{p}_i^0 = p_i^1 \frac{\sum_{i \in M_{01}} p_i^0 q_i^1}{\sum_{i \in M_{01}} p_i^1 q_i^1} \qquad \forall i \in N_{01} \tag{19}$$

With such an inflation adjusted carry forward/backward imputation mechanism, it can be shown that $\Delta^{01} = \frac{1}{\Pi^{01}}$. In other words, the matched model bias is zero if the imputed prices are defined as shown in equations 18 and 19. The inflation adjusted carry forward/backward imputation may not be the best approach in the case of life-cycle pricing, see section 6.

In order to examine the unit value bias, we introduce additional notations. We denote by s_i^t the quantity share in period t of item i within its BPG $\kappa(i)$.

$$s_i^t = \frac{q_i^t}{Q_{\kappa(i)}^t} \qquad \forall i \in N_t, t = 0, 1 \tag{20}$$

Moreover, we denote by σ_i^t the quantity share in period t of BPG $\kappa(i)$ to which the item i belongs. Note that σ_i^t is the same for all items that belong to the same BPG.

$$\sigma_i^t = \frac{Q_{\kappa(i)}^t}{\sum_k Q_k^t} \quad \forall i \in N_t, t = 0, 1 \tag{21}$$

Let us now define the following hybrid measure using all, or only the matched items.

$$\Theta^{01} = \frac{\left(\sum_{i \in M_{01}} p_i^0 s_i^1 \sigma_i^1\right) / \left(\sum_{i \in M_{01} \cup D_{01}} p_i^0 s_i^0 \sigma_i^1\right)}{\left(\sum_{i \in M_{01}} p_i^1 s_i^0 \sigma_i^0\right) / \left(\sum_{i \in M_{01} \cup N_{01}} p_i^1 s_i^1 \sigma_i^0\right)} \tag{22}$$

Again, it can be shown that the unit value bias can be decomposed into a component that is based on the observed prices (see equation 22), and a component that is based on the imputed prices (see equation 16).

$$b_{UV}^{01} = 0.5 \ln \left(\Theta^{01}\right) + 0.5 \ln \left(\Pi^{01}\right) \tag{23}$$

Suppose that the items that make up a BPG have the same price in each period, which is therefore identical to the average price for that BPG (i.e. $p_i^t = \bar{p}_{\kappa(i)}^t \forall i \in \kappa(i)$). Moreover, the imputed prices also correspond to that same price. Under such circumstances, it can be shown that $\Theta^{01} = \frac{1}{\Pi^{01}}$, and hence the unit value bias will be zero. This could be a theoretical justification for grouping together items

[3] Note that the measure Δ^{01} is very similar to the matching measure used in the MARS method (see Chessa, 2019 [3]). However, the MARS method compares the quantities, and not expenditures, of the matched items and all items. Moreover, the MARS method ignores the imputation term Π^{01}.

based on similar price levels.

We can now combine the decomposition of the unit value bias (equation 23) and of the matched model bias (equation 17) in order to decompose the difference between the hybrid and the matched index (equation 14).

$$\ln\left(\frac{P_{HF}^{01}}{P_{MF}^{01}}\right) = b_{UV}^{01} - b_{MM}^{01} = 0.5\ln\left(\Theta^{01}\right) - 0.5\ln\left(\Delta^{01}\right) \tag{24}$$

This decomposition does not depend on the imputed prices anymore, which makes sense as neither the matched index nor the hybrid index require imputations. However, one needs to define the imputed prices in order to be able to judge if either $|b_{UV}^{01}|$ or $|b_{MM}^{01}|$ is smaller.

In practice, we could prefer the tight product specification over the broad product specification if the matched model bias is close to 0 but the unit value bias is very different from 0. Conversely, we would prefer the the broad product specification over the tight product specification if the matched model bias is very different from 0 but the unit value bias is close to 0. If both matched model bias and unit value bias are different from 0, the imputation index may be the best option. All these conclusions depend however on the appropriateness of the imputation model.

5 A quality adjusted hybrid Fisher index

The framework described in the section 4 is based on the idea that the imputation Fisher index is the target index. It is a target in the sense that the true imputation is not known, although we can attempt to impute the missing prices in practice. In this section, we are going to propose a method that can be seen as an alternative target and clarify the relationship between this method and the imputation Fisher index.

We consider the following 2-stage aggregation method:

1. In the first stage, items are combined into BPGs for which a *quality adjusted* average price and a total quantity sold can be calculated. The calculation of the average price and total quantities requires that a quality adjustment factor v_i is associated with each item i that transforms the quantities into a common scale.

$$\hat{\tilde{p}}_k^t = \frac{\sum_{i\in H_k} p_i^t q_i^t}{\sum_{i\in H_k} v_i q_i^t} \tag{25}$$

$$\hat{Q}_k^t = \sum_{i\in H_k} v_i q_i^t \tag{26}$$

2. In the second stage, the quality adjusted prices and quantities of the BPGs are further aggregated using a Fisher price index (called hereafter a quality adjusted hybrid Fisher index).

$$P_{QA-HL}^{01} = \frac{\sum_k \hat{\tilde{p}}_k^1 \hat{Q}_k^0}{\sum_k \hat{\tilde{p}}_k^0 \hat{Q}_k^0} \tag{27}$$

$$P_{QA-HP}^{01} = \frac{\sum_k \hat{\bar{p}}_k^1 \hat{Q}_k^1}{\sum_k \hat{\bar{p}}_k^0 \hat{Q}_k^1} \tag{28}$$

$$P_{QA-HF}^{01} = \sqrt{P_{HL}^{01} P_{HP}^{01}} \tag{29}$$

Note that the quality adjustment factors are scale independent. By multiplying all the quality adjustment factors referring to items of the same BPG will not change the quality adjusted hybrid Fisher index. This is because the Fisher index satisfies the commensurability test (see test T10 in [23]). The proposed aggregation method also facilities the comparison with the hybrid Fisher index as defined in 11. If all the quality-adjustment factors are the same for the items of a BPG, then $P_{QA-HF}^{01} = P_{HF}^{01}$ and hence there is no unit value bias.

Compared to an imputation Fisher index, the proposed approach requires that quality adjustment factors are imputed instead of prices. We consider two strategies.

<u>Primary unit</u>

In some cases, an item can be characterized with a primary unit (for example the weight of a package expressed in grams, etc.). In order to account for the differences in this unit of items that belong to the same BPG, it can be practical to first transform all prices into this common unit. This is an example of a quality adjusted unit values that is based on the external information of the weight for example. Such a transformation can help to neutralize quality differences between the items that belong to the same BPG.

<u>Banerjee index</u>

Another example is to define the quality adjustment factors as the average of the observed or estimated price in the two comparison periods.

$$v_i = \begin{cases} 0.5 * \left(p_i^0 + p_i^1\right) & if \quad i \in M_{01} \\ 0.5 * \left(p_i^0 + \hat{p}_i^1\right) & if \quad i \in D_{01} \\ 0.5 * \left(\hat{p}_i^0 + p_i^1\right) & if \quad i \in N_{01} \end{cases} \tag{30}$$

It has been noted by Von Auer, 2014 [18] that the change in the quality adjusted average price with these factors is equivalent to the Banerjee index. It can also be shown that such an index approximates the imputation Fisher index. In our framework, we denote by $P_{IF|k}^{01}$ the imputation Fisher index calculated over the items that belong to the BPG k. We also denote by M_{01}^k,D_{01}^k and N_{01}^k the set of matched, disappearing and new items for a BPG k. We can show the following:

$$\begin{aligned} \frac{\left(\hat{\bar{p}}_k^1/\hat{\bar{p}}_k^0\right)}{P_{IF|k}^{01}} = & \left(\frac{\left(\sum_{i\in M_{01}^k} p_i^1 q_i^1 + \sum_{i\in N_{01}^k} p_i^1 q_i^1\right)^{0.5} \left(\sum_{i\in M_{01}^k} p_i^0 q_i^1 + \sum_{i\in N_{01}^k} \hat{p}_i^0 q_i^1\right)^{0.5}}{0.5\left(\sum_{i\in M_{01}^k} p_i^1 q_i^1 + \sum_{i\in N_{01}^k} p_i^1 q_i^1\right) + 0.5\left(\sum_{i\in M_{01}^k} p_i^0 q_i^1 + \sum_{i\in N_{01}^k} \hat{p}_i^0 q_i^1\right)} \right) \\ & / \left(\frac{\left(\sum_{i\in M_{01}^k} p_i^1 q_i^0 + \sum_{i\in D_{01}^k} \hat{p}_i^1 q_i^0\right)^{0.5} \left(\sum_{i\in M_{01}^k} p_i^0 q_i^0 + \sum_{i\in D_{01}^k} p_i^0 q_i^0\right)^{0.5}}{0.5\left(\sum_{i\in M_{01}^k} p_i^1 q_i^0 + \sum_{i\in D_{01}^k} \hat{p}_i^1 q_i^0\right) + 0.5\left(\sum_{i\in M_{01}^k} p_i^0 q_i^0 + \sum_{i\in D_{01}^k} p_i^0 q_i^0\right)} \right) \end{aligned} \tag{31}$$

The previous fraction may approximate unity: in each fraction, the nominator is a geometric average and the denominator is an arithmetic average of the same two terms.

Therefore, the method proposed in this section approximates a 2-stage aggregation method, where in the first stage (aggregating items into BPGs) an imputation Fisher index is applied and in the second stage (aggregating BPGs) a Fisher index is applied. This index is likely to approximate the (1-stage) imputation Fisher index. In other words, by choosing the quality adjustment factors as in 30, we obtain an index that will be close to the imputation Fisher index: $P^{01}_{QA-HF} \approx P^{01}_{IF}$. If there are no missing items, then we obtain an index that is close to the matched Fisher index $P^{01}_{QA-HF} \approx P^{01}_{MF}$.

6 Imputation method

In order to calculate an imputation index, a price must be estimated for the items that are only available in one of the two comparison periods. There are many options that can be considered for imputing the missing prices.

In the context of product churn, special attention should be given to situations where the pricing strategy depends on the life-cycle of a product. For example, it may happen that the last available price of the an item is a reduced price. This situation can be encountered at the end of a sales period and is especially common in clothing and footwear. Reduced prices can also be observed in situations of inventory clearing or closure of an outlet.

In case of life-cycle pricing, inflation adjusted carry forward or backward imputation methods may not be satisfactory. Suppose that the last price of an item observed in period $t-1$ is a reduced price. The prices of the other items remain stable between periods $t-1$ and t. By imputing a price for the disappeared item in period t based on the price change of the other items, the reduced price will be kept in the calculations, thereby having a downward impact on the index (see the examples in Eurostat, 2021 [10]). That is why we consider here alternative imputation methods.

Recall that the missing item i could be grouped together with other, similar, items. Let $\kappa(i)$ be the BPG to which item i belongs $\left(i \in H_{\kappa(i)}\right)$.The price of the missing item i is then set equal to the average price of $H_{\kappa(i)}$ in that period [4].

$$\hat{p}^t_i = \bar{p}^t_{\kappa(i)} \tag{32}$$

From a practical point of view, no additional information is needed apart from the assignment of the items into BPGs. The quantity for the missing item is zero as no purchase took place for item i in that

[4]This is a simple average price but the approach can be naturally extended if quality adjustment factors are available for each item. The missing price could then be estimated as the quality adjusted average price as defined in equation 25.

period.

This imputation method can also be formalized with a regression in which the dependent variable is the price and the independent variables are dummy variables for the groups. Formally, let G_{ki} be a dummy variable that is set to 1 if the item i belongs to the BPG k, and that is set to 0 otherwise. Consider the following model to be estimated in period t.

$$p_i^t = \alpha + \sum_{k \neq 1} \beta_k G_{ki} + \epsilon_{it} \quad \forall i \in N_t \tag{33}$$

If each item i in this regression is weighted by its quantity q_i^t, it can be shown that the estimated price $\hat{p}_i^t$ for an item obtained from model 33 corresponds to the average price defined in equation 32.

7 Multilateral extension

A bilateral Fisher index can be applied as a fixed base index or as a chained index. None of these two strategies is satisfactory in the context of scanner data. A fixed base Fisher index compares prices in a fixed base period with prices in the current period. The choice of the base period may have too much influence on the resulting index. Moreover, by moving 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 update the base period each month and chain link the resulting month-on-month Fisher indices. However, it has been found that a chained Fisher index can be subject to chain drift because the Fisher index is not transitive.

In order to overcome these limitations, transitive index formulas can be used. Transitivity is an index number property in which an index that compares periods a and b indirectly through period c is required to be identical to one that compares periods a and b directly. Several transitive index formulas have been proposed as a solution when using scanner data (see Chapter 10 in [22], and [11]). These index formulas are part of the family of multilateral methods. In a multilateral method, the aggregate price change between two comparison periods is obtained from prices and quantities observed in multiple periods, not only in the two comparison methods.

One specific example of a multilateral method is the Gini-Eltetö-Köves-Szulc (GEKS) method. This method is based on the bilateral Fisher indices calculated between any two periods of a given time window. These bilateral price comparisons are then averaged in order to obtain the GEKS price index. It can be shown that the GEKS index is the transitive index that is closest to its underlying bilateral indices.

In our context, we define the following GEKS indices based on the matched, imputation, hybrid and quality-adjusted hybrid Fisher indices. Let us consider a time window consisting of periods $0, 1, \dots, T$ over which the GEKS index is applied. The different GEKS indices are then defined as follows:

$$P^{0,t}_{GEKS-M} = \prod_{k=0..T} \left(P^{0k}_{MF} \cdot P^{kt}_{MF}\right)^{\frac{1}{T+1}} \quad \forall t \in 0, 1, \ldots, T \tag{34}$$

$$P^{0,t}_{GEKS-I} = \prod_{k=0..T} \left(P^{0k}_{IF} \cdot P^{kt}_{IF}\right)^{\frac{1}{T+1}} \quad \forall t \in 0, 1, \ldots, T \tag{35}$$

$$P^{0,t}_{GEKS-H} = \prod_{k=0..T} \left(P^{0k}_{HF} \cdot P^{kt}_{HF}\right)^{\frac{1}{T+1}} \quad \forall t \in 0, 1, \ldots, T \tag{36}$$

$$P^{0,t}_{GEKS-QA-H} = \prod_{k=0..T} \left(P^{0k}_{QA-HF} \cdot P^{kt}_{QA-HF}\right)^{\frac{1}{T+1}} \quad \forall t \in 0, 1, \ldots, T \tag{37}$$

Note that all four GEKS indices are transitive. This is because all Fisher-type indices actually satisfy the time reversal test. As a consequence, these indices do solve the problem of 'chain drift' caused by the bouncing of prices and quantities. This type of chain drift has been examined in Von Auer, 2019 [20]. However, the GEKS indices are not necessarily exempted from the matched model bias and unit value bias. As in the bilateral case, we can now distinguish, on the one hand, the matched-model bias from, on the other hand, the unit value bias. Note that these 'multilateral' biases can be defined as a GEKS-type average of the biases observed in the bilateral case.

$$b^{0,t}_{GEKS-MM} = \ln\left(\frac{P^{0t}_{GEKS-M}}{P^{0t}_{GEKS-I}}\right) = \frac{1}{T+1} \sum_{k=0..T} \left(b^{0,k}_{MM} + b^{k,t}_{MM}\right) \tag{38}$$

The matched-model bias can be decomposed into two components Δ and Π, as shown in equation 17. We now calculate multilateral versions of the Δ and Θ factors.

$$\Delta^{0,t}_{GEKS} = \frac{1}{T+1} \sum_{k=0..T} (\Delta^{0,k} * \Delta^{k,t}) \tag{39}$$

$$\Pi^{0,t}_{GEKS} = \frac{1}{T+1} \sum_{k=0..T} (\Pi^{0,k} * \Pi^{k,t}) \tag{40}$$

Note that Δ and Π satisfy the time reversal property, and therefore Δ_{GEKS} and Π_{GEKS} are transitive measures. Plugging these measures into 38, we obtain:

$$b^{0,t}_{GEKS-MM} = 0.5 * \ln(\Delta^{0,t}_{GEKS}) + 0.5 \ln(\Pi^{0,t}_{GEKS}) \tag{41}$$

Similarly, we have the following multilateral version of the unit value bias:

$$b^{0,t}_{GEKS-UV} = \ln\left(\frac{P^{0t}_{GEKS-H}}{P^{0t}_{GEKS-I}}\right) = \frac{1}{T+1} \sum_{k=0..T} \left(b^{0,k}_{UV} + b^{k,t}_{UV}\right) \tag{42}$$

The unit value bias can be decomposed into two components Θ and Π, as shown in equation 23. Similarly to equations 39 and 40, we now calculate a multilateral version of the Θ factor.

$$\Theta^{0,t}_{GEKS} = \frac{1}{T+1} \sum_{k=0..T} (\Theta^{0,k} * \Theta^{k,t}) \tag{43}$$

Note that Θ satisfies the time reversal property, and therefore Θ_{GEKS} satisfies the transitivity property. Plugging the measures defined in 40 and 43 into 42, we obtain:

$$b^{0,t}_{GEKS-UV} = 0.5 * \ln(\Theta^{0,t}_{GEKS}) + 0.5 \ln(\Pi^{0,t}_{GEKS}) \tag{44}$$

It follows that the difference between the matched and hybrid GEKS indices can be explained as follows:

$$\ln\left(\frac{P^{0t}_{GEKS-H}}{P^{0t}_{GEKS-M}}\right) = b^{0t}_{GEKS-UV} - b^{0t}_{GEKS-MM} = 0.5 * \ln(\Theta^{0,t}_{GEKS}) - 0.5\ln(\Delta^{0,t}_{GEKS}) \quad (45)$$

8 Examples

8.1 Example T-shirt

We illustrate the issues on a data set for t-shirts[5]. This data set is characterized by an overall downward trend in prices and significant item churn.

We calculate the following indices:

- A matched GEKS index that is calculated by matching the item codes (GEKS-M).
- An hybrid GEKS index that is calculated by grouping the items into BPGs. The variables Fabric (cotton or organic), Sleeves (long or short) and Number of items (1, 2 or 3) are used to specify the BPGs (GEKS-H).
- An imputation GEKS index that is calculated by matching the item codes and imputing the missing prices. The missing price corresponds to the average price of the BPG to which the missing item belongs (GEKS-I).
- A quality adjusted hybrid GEKS index that is calculated by grouping the items into BPGs. The quality adjustment factors are defined as the average of the observed or possibly imputed prices in the two comparison periods (GEKS-QA-H2).

The results are plotted in Figure 1. In this example, there are significant differences between the matched, imputation and hybrid indices. In fact, the matched GEKS index sits below the GEKS imputation index, which means that there is some negative matched-model bias. The hybrid GEKS index sits above the imputation GEKS index, which means that there is some positive unit value bias. As a consequence, an imputation index could be the preferred solution. Finally, the quality-adjusted hybrid GEKS index is slightly above the imputation GEKS index.

The matched-model and unit value biases become larger starting with period 3. This is because there are several items that are not available in periods 1 and 2, but are available thereafter. These items lead to both matched-model bias because of their non-inclusion in the matched approach and to unit value bias once they are grouped with other items.

[5]The data set was used by A. Chessa in a training session on the MARS method that was conducted during the 2018 Eurostat Workshop on Scanner Data organized by Statistics Norway in Oslo. We are thankful for the permission to reproduce this example in this paper.

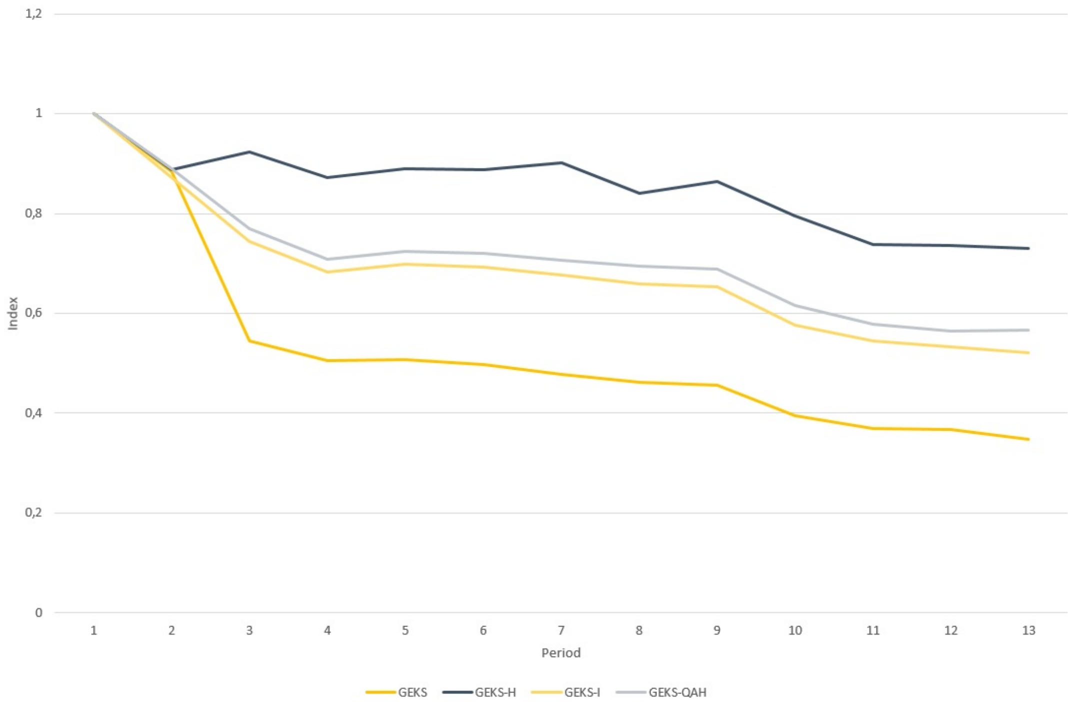

Figure 1: Price indices for the t-shirts scanner data.

We can also decompose the multilateral unit value bias, and the multilateral matched-model bias, as shown in equations 41 and 44. The results are included in Table 1. In this example, the matched model bias is larger (in absolute terms) than the unit value bias. Subtracting the term related to the imputation that is common to these two biases (the term $0.5\ln(\Pi^{0,t}_{GEKS})$ in the second column), we can still conclude that the matched model issue measured by the term $0.5 * \ln(\Delta^{0,t}_{GEKS})$ is larger (in absolute terms) than the unit value issue measured by the term $0.5 * \ln(\Theta^{0,t}_{GEKS})$.

8.2 Example diapers

We illustrate the issues on a real scanner data set that includes diapers sold across outlets in Luxembourg. This data set is characterized by a lot of product relaunches, with each relaunch containing a slightly different number of diapers included in the package. In total, there are 78 different GTIN codes and the data spans over 25 months (September 2020 to September 2022).

The initial data only contains a GTIN code and a text string. First, we extracted from the text string information in order to create the variables 'brand', 'type of diapers' and 'number of diaper included in the package'. We calculate the following indices:

- A matched GEKS index that is calculated by matching the GTIN codes (GEKS-M).
- An hybrid GEKS index that is calculated by grouping the items into BPGs. The variables 'brand' and 'type of diapers' are used to specify the BPGs (GEKS-H).

$0.5 * \ln(\Delta^{0,t}_{GEKS})$ (1)	$0.5 \ln(\Pi^{0,t}_{GEKS})$ (2)	$0.5 * \ln(\Theta^{0,t}_{GEKS})$ (3)	$b^{0,t}_{GEKS-MM}$ (1)+(2)	$b^{0,t}_{GEKS-UV}$ (3)+(2)
0,00	0,00	0,00	0,00	0,00
0,01	0,00	0,02	0,01	0,02
-1,48	1,17	-0,95	-0,31	0,22
-1,57	1,27	-1,03	-0,30	0,24
-1,61	1,29	-1,05	-0,32	0,24
-1,64	1,31	-1,06	-0,33	0,25
-1,61	1,26	-0,98	-0,35	0,29
-1,67	1,32	-1,07	-0,36	0,24
-1,70	1,34	-1,06	-0,36	0,28
-1,73	1,35	-1,03	-0,38	0,32
-1,76	1,37	-1,07	-0,39	0,30
-1,78	1,41	-1,09	-0,37	0,32
-1,78	1,37	-1,04	-0,41	0,34

Table 1: Decomposing the multilateral matched model and unit value biases.

- A quality adjusted hybrid GEKS index that is calculated by grouping the items into BPGs. The 'number of diapers included in the package' is used as the variable to calculate quality adjusted average prices (GEKS-QA-H1).
- An imputation GEKS index that is calculated by matching the GTIN codes and imputing the missing prices. The missing price corresponds to the quality adjusted average price of the BPG to which the missing item belongs (GEKS-I).
- A quality adjusted hybrid GEKS index that is calculated by grouping the items into BPGs. The quality adjustment factors are defined as the average of the observed or possibly imputed prices in the two comparison periods (GEKS-QA-H2).

The results are plotted in Figure 2. First of all, the hybrid GEKS sits above all the other price indices. This is because the average number of diapers in a package is increasing over time, and this has an upward impact on the average price captured by this index. The matched GEKS is still significantly upward biased compared to the imputation GEKS. The quality-adjusted hybrid index (GEKS-QA-H2) is the index that is closest to the imputation GEKS, as suggested by theory. The other quality-adjusted hybrid index (GEKS-QA-H1) is easier to calculate and approximates the imputation GEKS reasonably well, except in period 19.

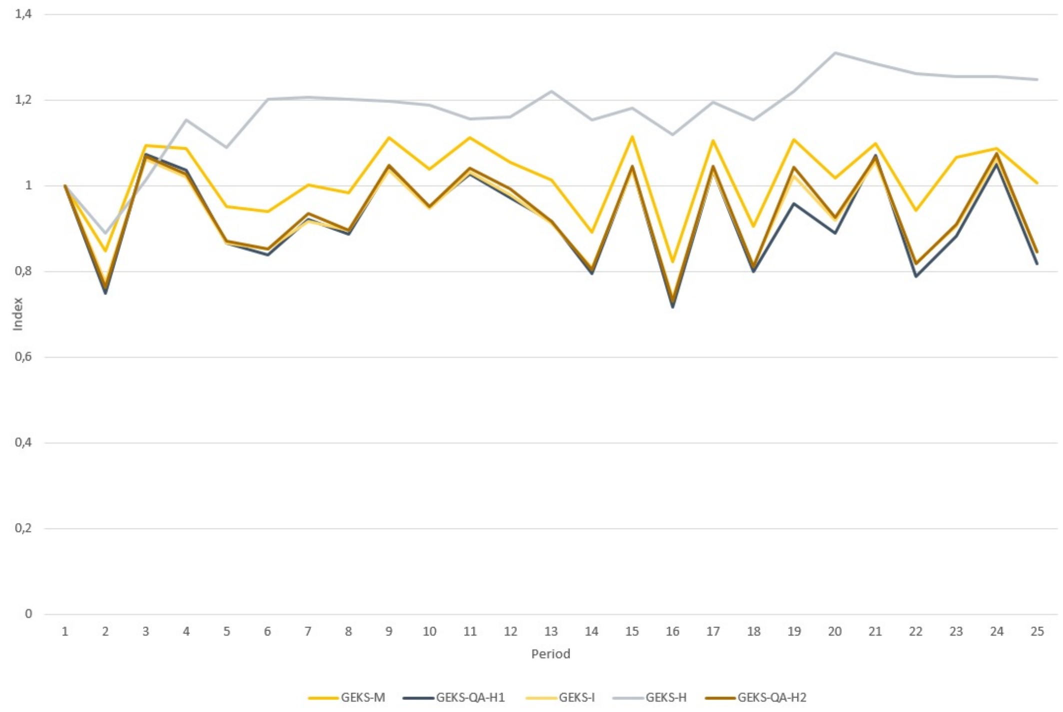

Figure 2: Price indices for the diapers scanner data.

9 Conclusion

The individual product that enters the price index can often be specified in different ways. The analysis in this paper is an attempt to formalize the problem of product specification In order to provide some guidance on which product specification to use, we compare a matched price index based on items to a hybrid index based on BPGs. We assess both unit value bias and matched-model bias by comparing the resulting matched and hybrid indices to an imputation index that acts as the target index. An alternative target index based on a quality adjusted hybrid Fisher index has also been proposed.

In principle, we prefer to use one of the target indices. A choice must be made between an imputation Fisher index and a quality adjusted hybrid Fisher index. We have found that these indices approximate each other fairly well if the quality adjustment factor of an item corresponds to the average price in the two comparison period of that item. More work is needed to disentangle the pros and cons of these two target indices.

The target indices require some kind of estimations. The imputation Fisher index requires values for the missing prices, and the quality adjusted hybrid Fisher index requires values for the quality adjustment factors. More work is needed on characterizing imputation methods in this context.

In practice, there may be a case for using either a matched Fisher index or an hybrid Fisher index if the matched-model bias or the unit value bias is found to be small. These measures can be seen as decision aid for selecting a product specification. Moreover, these measures are explicitly linked to the framework of Fisher price indices.

The framework is derived in the context of a GEKS index. It may also be applicable to other multilateral methods that are related to the Fisher index. It should be examined how the concepts of matched model bias and unit value bias can be extended to other multilateral methods such as the Weighted Time Product Dummy or the Geary-Khamis.

References

[1] Aizcorbe A. Outlet substitution bias estimates for ride sharing and taxi rides in new york city. *BEA Working Paper Series*, WP2022-1, 2021.

[2] Chessa A. A new methodology for processing scanner data in the dutch CPI. *EURONA*, 2016-1, 2016.

[3] Chessa A. MARS: A method for defning products and linking barcodes of item relaunches. *Paper presented at the 16th meeting of the Ottawa group, Rio de Janeiro, Brazil.*, 2019.

[4] Williams B. and Sager E. A new vehicles transaction price index: Offsetting the effects of price discrimination and product cycle bias with a year-over-year index. *BLS Working Paper*, 2021.

[5] European Central Bank. Inflation measurement and its assessment in the ECBs monetary policy strategy review. *ECB Occasional Paper Series*, 2021.

[6] Melser D. and Webster M. Multilateral methods, substitution bias and chain drift: some empirical comparisons. *Review of Income and Wealth*, 67(3), 2021.

[7] de Haan J. Generalized fisher price indexes and the use of scanner data in the CPI. *Paper presented at the 6th meeting of the Ottawa group, Canberra, Australia.*, 2001.

[8] de Haan J. and Daalmans. Scanner data in the CPI – the imputation CCDI index revisited'. *Paper presented at the 16th meeting of the Ottawa group, Rio de Janeiro, Brazil.*, 2019.

[9] de Haan J. and Krsinich F. Scanner data and the treatment of quality change in nonrevisable price indexes. *Journal of Business and Economic Statistic*, 32(3), 2014.

[10] Eurostat. *Recommendations on the bridged overlap method*, 2021. Available at `https://ec.europa.eu/eurostat/documents/272892/7048317/Recommendation-bridged-overlap-June-2021.pdf/`.

[11] Eurostat. *Guide on the use of multilateral methods in the HICP (Version 2022).*, 2022. Available at `https://ec.europa.eu/eurostat/product?code=KS-GQ-21-020`.

[12] Diewert E.W., Fox K., and de Haan J. A newly identifed source of potential CPI bias: Weekly versus monthly unit value price indexes. *Economics Letters*, 141(C), 2016.

[13] Diewert E.W. and von der Lippe P. Notes on unit value index bias. *Discussion Paper University of British Columbia*, 10-08, 2010.

[14] Konny C. G., Williams B.K., and Friedman D. M. Big data in the U.S. consumer price index: Experiences & plans.

[15] Daalmans J. Mmultilateral indices and the relaunch problem: product clustering and alternative solutions. *Paper presented at the 17th meeting of the Ottawa group, Rome, Italy.*, 2022.

[16] Dalèn J. Unit values in scanner data some operational issues. *Paper presented at the 15th meeting of the Ottawa group, Eltville, Germany*, 2017.

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

[18] Von Auer L. The generalized unit value index family,. *Review of Income and Wealth*, 60, 2014.

[19] Von Auer L. Processing scanner data by an augmented GUV index. *EURONA*, 2017-1, 2017.

[20] Von Auer L. The nature of chain drift. *Paper presented at the 16th meeting of the Ottawa group, Rio de Janeiro, Brazil.*, 2019.

[21] Leclair M., Léonard I., Rateau G., Sillard P., Varlet G., and Vernédal P. Scanner data: Advances in methodology and new challenges for computing consumer price indices. *INSEE économie et statistiques / economics and statsitics*, N. 509, 2019.

[22] ILO, IMF, OECD, EUROSTAT, UNECE, and World Bank. Chapter 10. Scanner data. In *Consumer Price Index Manual: Concepts and Methods*, 2020.

[23] ILO, IMF, OECD, EUROSTAT, UNECE, and World Bank. Chapter 3. The axiomatic or test approach to index number theory. In *Consumer Price Index Manual: Theory (draft version)*, 2021.