Using multilateral hedonic methods to capture product relaunches

Meeting of the Group of Experts on Consumer Price Indices

Geneva, Switzerland: 7-9 June 2023

Ken Van Loon, Adrien Mierop & Dorien Roels[1]

Abstract

Statistics Belgium currently uses the GEKS-Törnqvist/CCDI multilateral method for supermarket scanner data. For consumer electronics scanner data, we use the Imputation Törnqvist GEKS (or ITGEKS) to capture quality differences between new and discontinued products. In supermarket scanner data, there are also products that leave the market and others that enter the market. Although quality changes are less pronounced for traditional supermarket products, the relaunch of products in supermarkets may coincide with a price level change, especially in the context of shrinkflation. In shrinkflation, the price of a product tends to remain the same, but the quantity received by consumers decreases, effectively increasing the price faced by consumers. Failure to take account of these relaunches could potentially lead to a bias in the CPI. Traditional (multilateral) methods use unique product identifiers to match the same product over time. In the case of relaunches, a new product may be given a new product identifier and any potential price change is missed when using these matched model methods. We show that multilateral index methods that use hedonic quality adjustments (e.g. ITGEKS or Time Dummy Hedonic method) can be used with supermarket scanner data and that they solve or provide insight into the relaunch problem. Under hedonics, products with similar characteristics are considered as similar products and relaunches or shrinkflation are effectively captured.

[1] Statistics Belgium, email: Ken.Vanloon@economie.fgov.be, Adrien.Mierop@economie.fgov.be & Dorien.Roels@economie.fgov.be.
This research was funded by Eurostat grant 2021-BE-PRICE, for which the authors express their gratitude. The views expressed in this paper are those of the authors and do not necessarily reflect the views of Statistics Belgium.

Introduction

In 2023, Mars Incorporated reduced the package size of its cat food brand Whiskas by around 15% in most European countries, e.g. a pack of 100 grams was reduced to 85 grams while the price remained the same (Wood, 2023). This is an example of shrinkflation: the price stays the same, but the package size is reduced, increasing the price in grams or litres for the consumer.

Shrinkflation might not only be "*cat*astrophic" for cat owners, but it also causes problems for price statisticians who use new data sources such as scanner data in the compilation of the Consumer Price Index (CPI). Typically, matched-model methods are used when compiling price indices using scanner data. These methods use unique product identifiers such as barcodes or stock keeping units to compare the same product ("like with like") over time. However, in the context of shrinkflation these methods can fail and cause a (downward) bias in the CPI, as products usually get a new product identifier when the package size changes. As a result, the price change between the old and the new product is not captured when using these product identifiers. This is in contrast to classical price collection, where usually a quantity or package size adjustment is carried out using the information obtained from price collectors. This is typically difficult to do with scanner data due to the sheer number of observations.

This paper examines whether multilateral index methods that use hedonic quality adjustments, such as the Imputation Törnqvist GEKS or Time Dummy Hedonic method, can be applied to supermarket scanner data. It also analyses whether these methods can successfully capture product relaunches or shrinkflation. Under hedonics, products with similar characteristics could be considered as the same product (or correctly quality or quantity adjusted for), and the product relaunch would indeed be captured.

Section 1 gives an overview of the implementation of scanner data by Statistics Belgium and a summary of the different methods that have been used so far. Currently, we use supermarket scanner data obtained directly from supermarkets and scanner data for consumer electronics and household appliances which is obtained from a market research company. As will be made clear, the methods that we use are different for each of those two segments. Section 2 provides more insight into what product relaunches are, which product identifiers we use and how we try to capture these product relaunches semi-automatically using a combination of text mining and manual verification of price collectors. Section 3 describes the two multilateral hedonic methods that we will empirically evaluate. The two methods are the Time Dummy Hedonic method (TDH) and the Imputation Törnqvist GEKS (ITGEKS) with bilateral time dummy hedonic indices as inputs. These two methods are going to be compared to a GEKS-Törnqvist/CCDI method with unique product identifiers. In Section 4, an empirical evaluation of the three index methods is carried out. This section focuses on a comparative analysis, assessing the performance and robustness across for different product categories. We first focus on products that have experienced shrinkflation in recent months, as these provide interesting case studies for comparative analysis. We then extend the comparative analysis to other product groups, with a longer time window and compare a non-spliced to a spliced version of the indices to examine whether this has any effect. Section 5 concludes the paper, discussing the main findings and suggesting future work.

1. Background on scanner data methods used in the Belgian CPI/HICP

Statistics Belgium has been using scanner data from supermarkets to compile the CPI since 2015 and in the HICP since 2016. The method used initially was the so-called “dynamic method”, which used an unweighted monthly chained Jevons index for a sample of products (Van Loon and Roels, 2018). The sample of products was determined each month using the turnover figures at a product level in two adjacent months. With products being included in the sample if their turnover was above a dynamically determined threshold. The monthly chaining was necessary because scanner datasets tend to have a high attrition rate of products, making it difficult to use a fixed basket of products.

A disadvantage of this method was that the turnover available in scanner data was only implicitly used to select the products in the sample and not explicitly (e.g. every product getting an explicit weight based on its turnover). It is by now well known that incorporating the available turnover information into a chained monthly index calculation (e.g. superlative formulae such as Törnqvist) leads to chain drift. A solution for this is to use multilateral methods, these methods maximize the number of matches in the data and without running the risk of introducing chain drift. They do this by measuring the price change between two periods based on information observed in multiple periods (de Haan, Hendriks, Scholz, 2016).

After a few years of research, the dynamic method was replaced by a multilateral method in January 2020 (Van Loon, 2020). The method we ended up implementing was the GEKS-Törnqvist (also known as CCDI). This method (Ivancic, Diewert and Fox, 2011) uses all possible matching products and calculates the price index between two periods as an unweighted geometric average of all possible matching bilateral Törnqvist indices within a time window. The time window we use is a window length of 25 months. This window length performed very well for both non-seasonal and seasonal items (Radjabov and Van Loon, 2022). To compile non-revisable indices when using multilateral methods, a splicing or extension method needed to be used. We settled on implementing the half-splice on published indices (Chessa, 2021). This method combined with a rolling window of 25 months, causes the annual rate of change of the published index series to correspond to the annual rate of change of the last calculated multilateral index. This makes – from an index compilers point of view – the analysis, validation and decomposition much easier. An important benefit of this method is that it can take care of “base level effects” that can happen when changing methods or data sources using the traditional December overlap in the HICP.

In January 2022, we expanded the coverage of scanner data in the Belgian CPI/HICP. The indexes of consumer electronics and household appliances are compiled using scanner data from a market research company. Due to the specific characteristics of these market segments, the same methodology as for traditional supermarket scanner data could not be used. These segments have a much higher attrition rate (product churn) due to the relative short life cycle of its products in which products also tend to see their prices decrease drastically. New products that enter the market tend to also have "better” or “different” characteristics compared to the products leaving the market. Therefore, a traditional method of matching or chaining the same product throughout time using a (unique) product identifier is not a good strategy and leads to a downward bias when compiling price indices for consumer electronics and household appliances. After evaluating several methods (Van Loon, 2021), the multilateral Imputation Törnqvist GEKS (ITGEKS) method using bilateral time dummy hedonic indices was implemented since this method dealt the best with the aforementioned particular challenges. The hedonic part takes care of the quality adjustment part between new and discontinued products.

2. Capturing product relaunches "semi-automatically"

In supermarket scanner data, there are also products leaving and entering the market. While quality change is less pronounced for supermarket products, the relaunch of products in supermarkets might coincide with a price level change (Dalén, 2017), as the aforementioned example of Whiskas shows in the context of shrinkflation. With a relaunch, we mean a product gets a new packaging or package size and a new barcode (GTIN) or other product identifier (Stock Keeping Unit, SKU). Missing these relaunches could potentially cause a bias, i.e. when producers keep the price level unchanged but reduce the package size. If in the case of relaunches, "new" products get new product identifiers then any potential price change between the discontinued product and its replacement or relaunch is missed when using these matched model methods.

To identify products in scanner data segments, Statistics Belgium uses stock keeping units (SKUs) instead of official barcodes (GTIN). These codes are retailer specific codes which are normally used by retailers to track their inventory. SKUs also allow price indices to be compiled for seasonal products (i.e. fruit and vegetables) and fresh products such as meat. For instance, a certain type of minced meat is typically sold in different amounts of grams (e.g. 422 grams, 424 grams, …) with all of these different weights having other product codes. SKUs make it possible to aggregate the quantities and turnover with a standardized unit of measurement (e.g. 1 kg of a certain kind of minced meat). Dividing the turnover by the quantities sold then gives you the average price per kilogram.

Usually these codes are a level above the official barcode because a SKU can combine multiple GTINs. For example, if during Christmas a certain kind of chocolate is sold in a Christmas wrapping next to the traditional wrapping, these are sold using different GTINs. These are then combined by the retailer using the same SKU. Using SKUs might capture some purely cosmetic product relaunches; however it is mostly retailer specific and will not capture all of the product relaunches, especially when the relaunch coincides with a change in package size.

In our monthly CPI/HICP compilation, we try to link relaunches "semi-automatically" with manual verification. To help identify relaunches, fuzzy matching techniques are used to find products that have similar descriptions. For each possible combination of discontinued and new product a score is calculated based on the number of words in common in the 2 product descriptions. Lists of possible relaunches are created that include all combinations of new and discontinued products with a score higher than a predetermined threshold. This list is first reduced by considering only products with the same brand name (or brand names that are considered to be synonymous). Additionally, combinations in which the product description for the discontinued and new item are completely identical are automatically identified as relaunches. Given that the package size of the 2 products sold can be different, a coefficient is automatically calculated to perform a quantity adjustment (this coefficient can be adapted if necessary). The remaining cases are manually analysed by central price collectors. For the manual verification, the price collector looks up more metadata online. This extra information may help in determining whether it is a relaunch and if a quantity adjustment should be made. In the case the price collector determines there is a relaunch, the new and old product code are linked, and quantity adjustment is made if necessary.

Unlike in bilateral methods, which always compare the price of the current month with the price of a reference month, multilateral methods compare prices over a window of several months. Therefore, products may have several relaunches in the same window. In order to take into account all these potential relaunches, the unit value prices of products subject to relaunches are "combined" by adding the turnover and sales volumes of the different SKUs considered to be the same product.

To account for quantity adjustments, the sales volume of the discontinued product will be divided by a coefficient equal to the ratio of the content of the new product divided by the content of the discontinued product. In the next step, a new unit value is obtained by dividing the total turnover of the products considered to be the same by the quantity adjusted sales volume of these products.

The following table illustrates the case where product A is replaced by product B over a period of 4 months. In the second and third months, the 2 products are sold simultaneously. In the fourth period, only product B is sold:

Period	Product	Sales volume	Turnover	Price	content (ml)
1	A	10000	30000	3.00	500
2	A	15000	48000	3.20	500
2	B	5000	19800	3.96	600
3	A	9000	26100	2.90	500
3	B	10000	40000	4.00	600
4	B	16000	65600	4.10	600

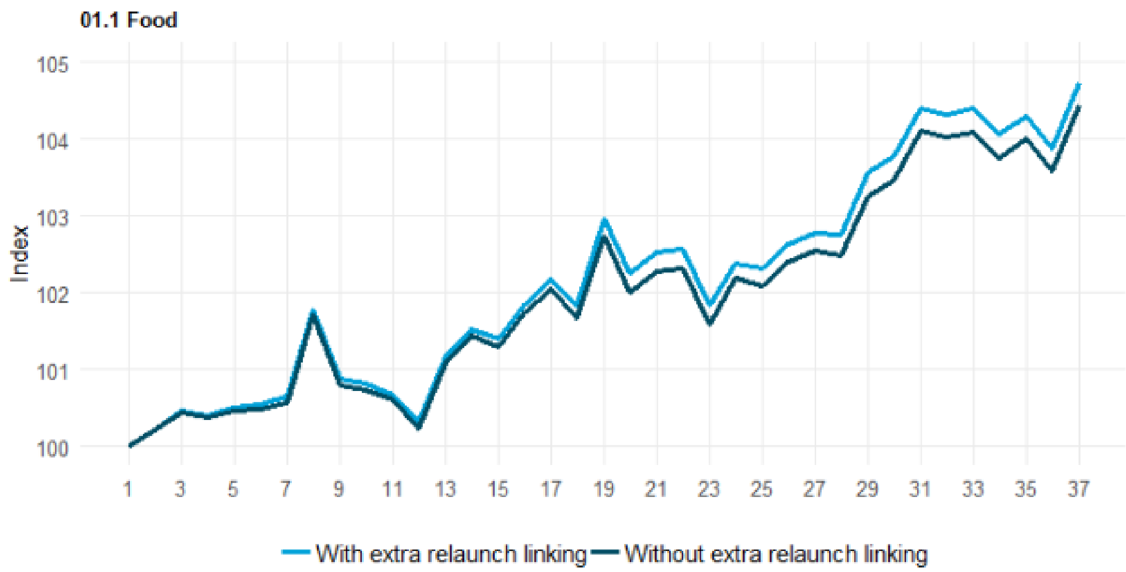
In this case, the quantity adjustment coefficient is 1.2 (600/500). In periods 1, 2 and 3, the sales volume of product A is therefore divided by this coefficient.

Period	Product	Sales volume (corrected)	Turnover	Price	content (ml)
1	A	8333.33	30000	3.00	500
2	A	12500.00	48000	3.20	500
2	B	5000.00	19800	3.96	600
3	A	7500.00	26100	2.90	500
3	B	10000.00	40000	4.00	600
4	B	16000.00	65600	4.10	600

The data that will be used to compile the index is shown in the table below:

Period	Product	Sales (corrected)	Turnover	Price	content (ml)
1	B	8333.33	30000	3.60	600
2	B	17500.00	67800	3.87	600
3	B	17500.00	66100	3.78	600
4	B	16000.00	65600	4.10	600

In periods 2 and 3, when the two products are sold simultaneously, the "corrected" sales volume and turnover are added together to calculate a new unit value price. Finally, the product identifier used in the index calculation is replaced with the SKU of the new product (B in this example). If the product B is replaced by a new product C in a subsequent month of the window, the same calculation is performed, and the SKUs of products A and B are replaced by the SKU of C. In this case, the corrected sales volume of product A is quantity adjusted a second time, if necessary. Previous research by us (Van Loon, 2019) has shown that performing this work has some effect. The effect of accounting for relaunches on the index is shown in the graph below. Carrying out additional product relaunch linking shifts the index levels upwards.



3. Index methods

The downside of linking products or relaunches with a manual intervention is that it is not only time consuming, but price collectors may miss some relaunches (e.g. due to human error or too much time between the new product and the discontinued product). Also not all product linkings carried out by our price collectors are relevant in the sense that they end up given a different price index. The only relaunches that are relevant from a price index point of view are those that coincide with a price level change.

It is therefore interesting to assess whether multilateral indices which use hedonic quality adjustment methods (e.g. ITGEKS or Time Dummy Hedonic method) can be used to compile price indexes for supermarket scanner data and also whether they solve or provide insight into the relaunch problem or shrinkflation. Under hedonics, products with similar characteristics can be considered to be similar (and correctly quantity or quality adjusted for) and thus product relaunches would in fact be captured. To apply hedonics, characteristics information needs to be extracted from the scanner datasets and in the context of shrinkflation obviously detailed information regarding package size or contents need to be available too.

Other methods have been proposed to solve the product relaunch puzzle, such as stratification or clustering (Chessa, 2019). By grouping similar products together using categorical variables, the prices of new and disappeared products are directly compared if they are attributed to the same strata or cluster. A potential drawback of this is that while these methods try to make a compromise between homogeneity and product match, they might introduce a unit value bias if the strata end up being too heterogeneous (Daalmans, 2022). These methods are out-of-scope of this paper and are not further examined.

3.1. Time Dummy Hedonic method

The first option we will look into is the Time Dummy Hedonic (TDH) method (de Haan, 2010). This method is quite straightforward. We use a log-linear specification and estimate it using weighted least squares (WLS), with expenditure shares in each period serving as weights.

$$\ln p_i^t = \alpha + \sum_{t=1}^{T} \delta^t D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t$$

Prices of all products (i) for several periods (t) are pooled in the same regression for every segment, on their characteristics (z_{ik}) and on dummy variables for the periods (D_i^t).

The main advantage of this method is its simplicity, since the index follows directly from the estimated time dummy parameters. By exponentiating the time dummy parameters, the index is obtained for all of the periods in the pooled regression. The base period is by definition 100.

In the equation above all of the data are pooled. This means that when new periods are added to the time window, the indices of the previous periods will change as they will be re-estimated. Continuously revising indices is problematic since it would imply that official indices are never definitive.

A solution to deal with the revision problem is to use a rolling window and apply splicing. Splicing here will, for comparability reasons, be limited to the option that we currently use in the monthly compilation of the CPI/HICP, namely the half splice on published indices (HASP) with a 25-month window. The index using a full window will also be shown.

3.2. ITGEKS (Imputation Törnqvist GEKS)

A disadvantage of the TDH method is that it forces parameter fixity for the whole window, which is quite restrictive, but may not be as much of a problem for supermarket products, since unlike consumer electronics, there is not (that) much technological progress. Another drawback is that, if there is no product attrition, there is no need to “quality adjust”; a matched index would be preferable in such cases. Which brings us to the ITGEKS method with hedonic bilateral indices as inputs for the CCDI/GEKS Törnqvist as proposed by de Haan and Krsinich (2014). This is the method that we now use for scanner data for consumer electronics and household appliances when compiling the CPI and HICP.

To illustrate this method we can start with the standard GEKS formula which uses all possible matched products and calculates the price index between months 0 and t as an unweighted geometric mean of $T + 1$ ratios of matched-model bilateral price indices P^{0l} and P^{lt}, with l running through $[0, T]$:

$$P_{GEKS}^{0,t} = \prod_{l=0}^{T} \left(P^{0l}/_{P^{tl}}\right)^{(1/_{T+1})} = \prod_{l=0}^{T} \left(P^{0l}P^{lt}\right)^{(1/_{T+1})}$$

The indices P^{0l} and P^{lt} are the bilateral Törnqvist indices between periods 0 and l and period l and t respectively. The Törnqvist index is defined as:

$$P_{T}^{0,t} = \prod_{i=1}^{n} \left(\frac{p_i^t}{p_i^0}\right)^{0.5\left(\frac{p_i^0 q_i^0}{\sum_{j=1}^{n} p_j^0 q_j^0} + \frac{p_i^t q_i^t}{\sum_{j=1}^{n} p_j^t q_j^t}\right)} = \prod_{i=1}^{n} \left(\frac{p_i^t}{p_i^0}\right)^{0.5\,(s_i^0 + s_i^t)}$$

With s_i^t (resp. s_i^0) the market share of product i in period t (resp. 0). The above GEKS formula is still a matched index, only the same products are matched over time using product identifiers. Recall that this is the method that we currently use in our CPI/HICP, except that we try to take into account product relaunches by semi-automatically linking them.

In the ITGEKS method proposed by de Haan & Krsinich (2014) the difference between new and disappearing products is taken into account by carrying out hedonic imputations for the unmatched items. This is done by using weighted bilateral time dummy hedonic indices as inputs for P^{0l} and P^{lt} instead of the Törnqvist-indices:

$$\ln p_i^t = \alpha + \delta^t D_i^t + \sum_{k=1}^{K} \beta_k z_{ik} + \varepsilon_i^t$$

Now the question remains, which weights do we need to use to estimate the bilateral time dummy hedonic indices when using WLS?

As de Haan (2004) has shown, using the mean expenditure shares for matched items and half expenditure shares for the new and disappeared items as weights, it is algebraically equivalent to:

$$P_{ITGEKS}^{0,t} = \prod_{i \in U_M^{0t}} \left(\frac{p_i^t}{p_i^0}\right)^{0.5\,(s_i^0 + s_i^t)} \prod_{i \in U_D^{0t}} \left(\frac{\hat{p}_i^t}{p_i^0}\right)^{0.5\,(s_i^0)} \prod_{i \in U_N^{0t}} \left(\frac{p_i^t}{\hat{p}_i^0}\right)^{0.5\,(s_i^t)}$$

With U_M^{0t} as the set of matched items, U_D^{0t} the set of disappeared items and U_N^{0t} the set of new items. The above index is a variant of the single imputation Törnqvist price, where imputations for the unmatched items are made with the estimated bilateral time dummy hedonic model. The expenditure shares used in the regression are identical to those used in the Törnqvist index. If there is no product attrition (i.e. no new or disappearing products), the above index is equal to a matched Törnqvist index and the ITGEKS is identical to the traditional GEKS Törnqvist index. This is a desirable property, as a matched index would be preferable in such cases.

A practical disadvantage of the ITGEKS compared to the TDH is its complexity and the fact that many bilateral time dummy regressions have to be estimated. For a window $[0, T]$ with a window length of $T + 1$ a total of $(T(T + 1))/2$ bilateral time dummy hedonic regressions have to be estimated. For a window period of 25 months, this requires the estimation of 300 bilateral time dummy hedonic regressions.

Using the bilateral time dummy hedonic method still forces fixity in the parameters, but now only between the two months compared, rather than over the whole window period as in the TDH. Other methods have been proposed to obtain the imputed prices $\hat{p}$ without parameter fixity, such as running a log-linear hedonic model for each period, once again using WLS (De Haan and Daalmans, 2019):

$$\ln p_i^t = \alpha^t + \sum_{k=1}^{K} \beta_k z_{ik}^t + \varepsilon_i^t$$

These imputed prices can then be used in the ITGEKS index formula mentioned above. In addition to a single imputation, a double imputation is also possible. In a double imputation the observed prices for the set of disappeared items U_D^{0t} and set of new items U_N^{0t} are replaced by their imputed values. The reason for this is that there may be omitted variable bias, which can be reduced by carrying out a double imputation. If there is not much omitted variable bias the difference between single and double imputation will be small. The advantage of running a log-linear hedonic model for each period is that the parameters are no longer fixed over time, unlike in the ITGEKS with bilateral time dummy hedonic indices as inputs. The downside of doing imputations for each period is that these cannot be carried out for new or disappeared characteristics. Put simply, new or disappeared characteristics have to be left out of the regression equation (Van Loon, 2021). For supermarket products, for example, it might for instance be a new segment in the classification of a supermarket chain or a new brand. For this reason, we do not investigate this method further and limit ourselves in the next section to the TDH and the ITGEKS with bilateral time dummy hedonic indices as inputs. As with the TDH, the ITGEKS index using the full window (without splicing) and the index using the half splice on published indices (HASP) with a 25-month window are both examined.

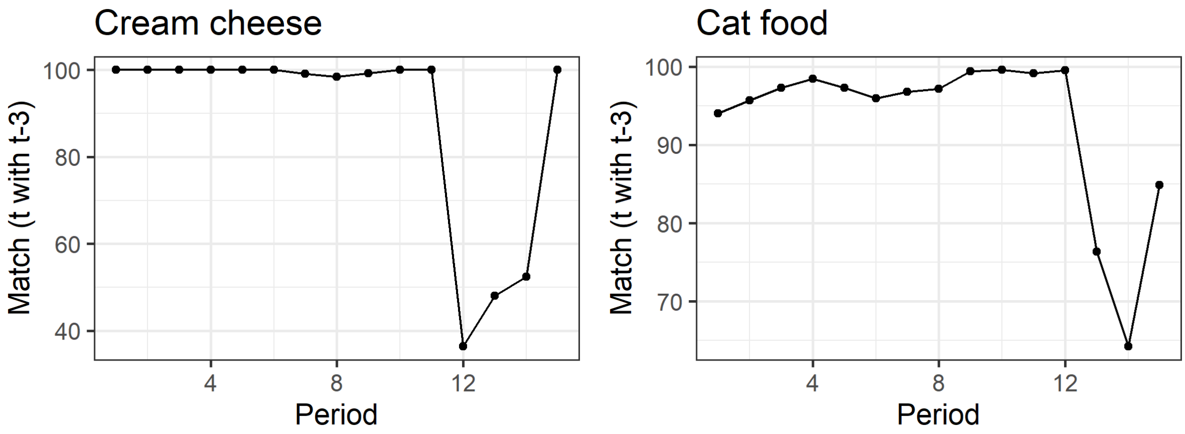
4. Empirical evaluation

In this section, we assess the feasibility of the TDH and ITGEKS and compare their results with the traditional GEKS using SKUs as product identifiers. To evaluate the methods, we do not look at the manual linkings. The data that is used are scanner datasets that we currently use to compile the CPI and HICP. These datasets are merged with the internal classifications of supermarket chains. The scanner datasets contain for each product the SKU, the number of sales, turnover, detailed product descriptions (i.e. separate variables for brand, type and other information), separate variables for package size and the unit of measure (kilograms, litre, …) and our classification to ECOICOP. The retailers' internal classification is structured hierarchically, creating thousands of possible combinations at the most detailed level. Obviously, some data cleaning was necessary; for instance the same brand was not always written the same way in the scanner data. With the final dataset, we can compile and compare (1) the GEKS index, (2) the TDH index and (3) the ITGEKS index and examine how well the latter two methods are able to capture product relaunches or shrinkflation.

To provide insight into our methodology and findings, we start with two specific product segments that experienced shrinkflation during our test period: cat food and cream cheese. These two examples immediately highlight the impact of product relaunches or shrinkflation on the 3 index methods.

The figure below shows the percentage of sales units from 3 months ago (t-3) that can be matched in a period t. As the figure shows, cream cheese and cat food experienced a significant "relaunch". For example for cream cheese in month 12, only ±40% of the number of sales from 3 months ago (=month

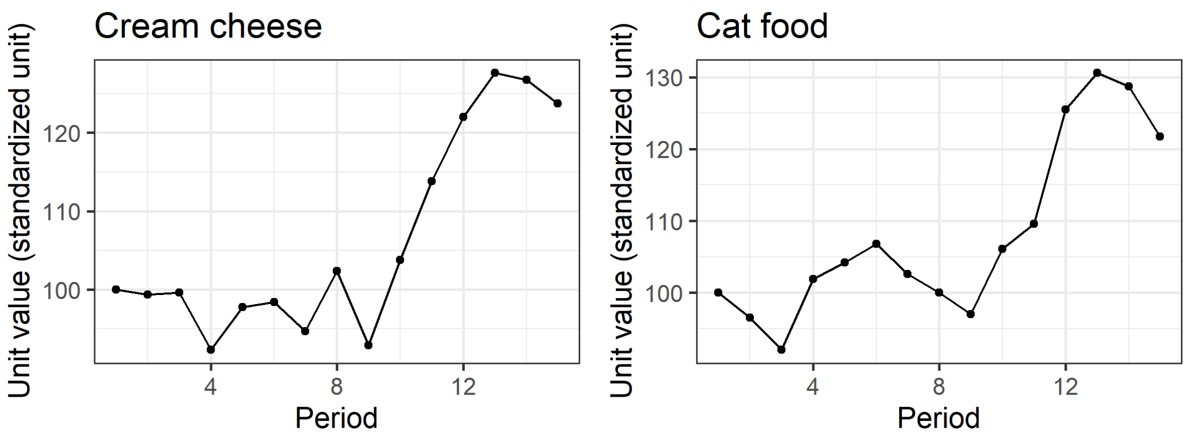
9) can be matched. While for cat food in month 14, only ±65% of the number of sales from month 11 can be matched.



We look at sales units rather than the number of products that can be matched, because sales (or expenditure) are a better indicator of a potentially problematic relaunch than the number of individual products that can be matched. A high attrition rate of products that hardly sell will not affect the price index that much. Apart from sales units, expenditure shares could also be used.

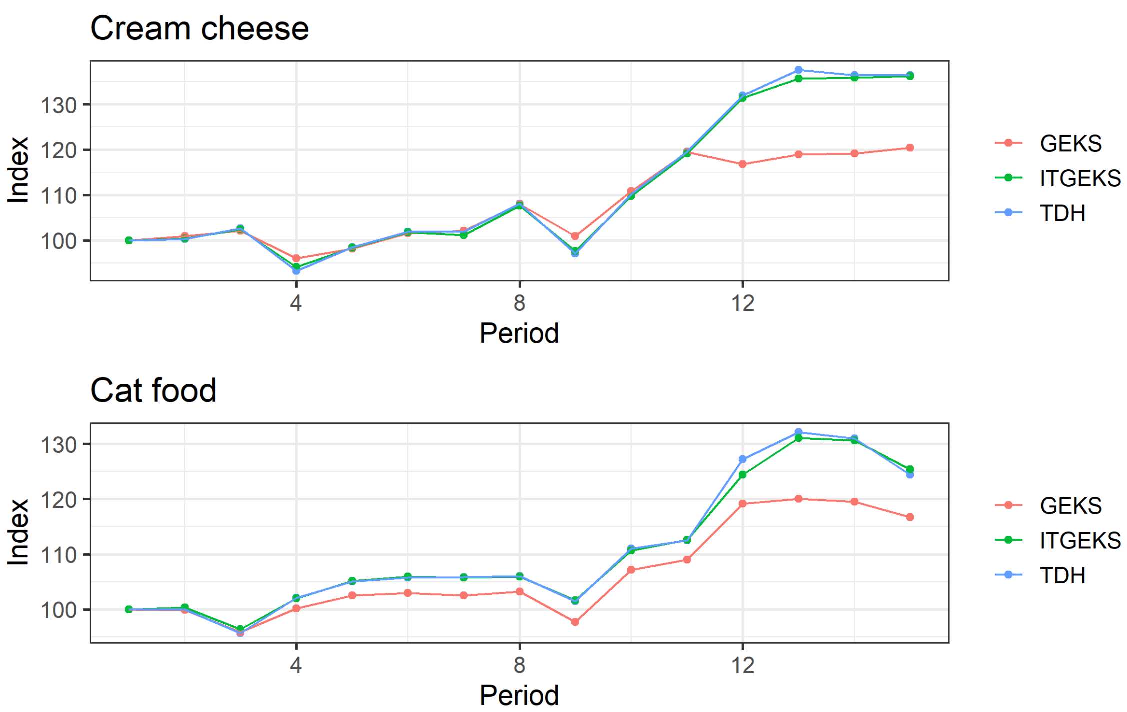
The reason why we look at the number of sales that can be matched with 3 months ago is because product relaunches can happen gradually, for instance the remaining stock of "disappearing" items usually needs to be sold first and this can take a while. If a comparison would be made between consecutive months, only gradual declines might show up in the graphs, which are analytically less interesting for detecting large shifts. Comparing the number of sales with to a fixed month is also analytically less interesting, as the graphs would end up showing gradually decreasing matching sales due to product churn, making it harder to detect significant shifts.

As emphasised earlier, a product relaunch is not necessarily problematic if it does not coincide with a price level change. For example, if the price change is not really different from the packaging size change, then there is no price change that would be missed when using matched-model methods. However, if we look at the unit value index for a standardized unit (e.g. price/kg), we see in the figure below (period 1 = 100) that both relaunches coincide with an increase in the standardised unit value price index, indicating that there might be a problematic product relaunch or shrinkflation. Of course, these "standardised" unit value indices cannot be used as a proxy for a proper price index since these indices show other compositional effects as well. However, they may indicate that something is going on that is worth examining in detail.



Further examination of the underlying data confirmed that these were both cases of product relaunches and, in this case, shrinkflation. In fact, the same product was relaunched with a smaller package size, but the price was not reduced (by the same amount).

We then compile the three price index methods: GEKS, ITGEKS and TDH. The GEKS will be compiled using a standard matched model with SKUs as product identifiers, no manual product relaunch linkings have been carried out. For the ITGEKS and the TDH, for each retailer we used as characteristics: variables related to the product description, package size and variables derived from the internal classification of the retailer. The R squared ranged between 0.903 and 0.963.

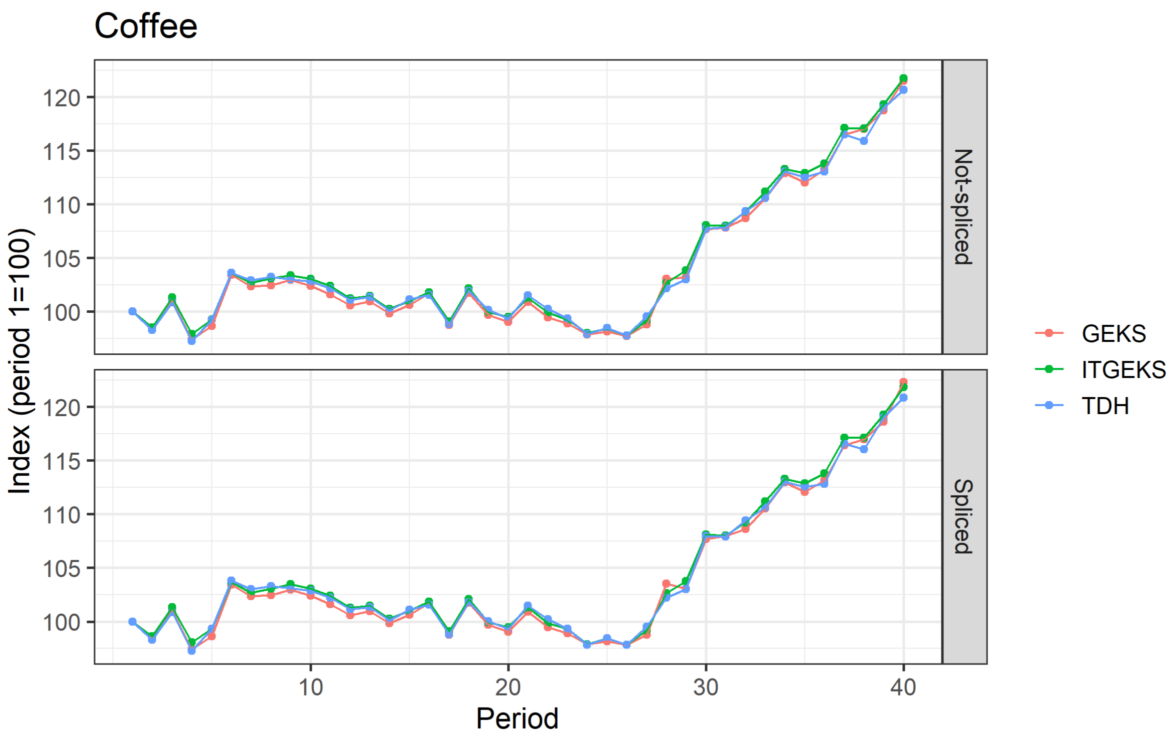


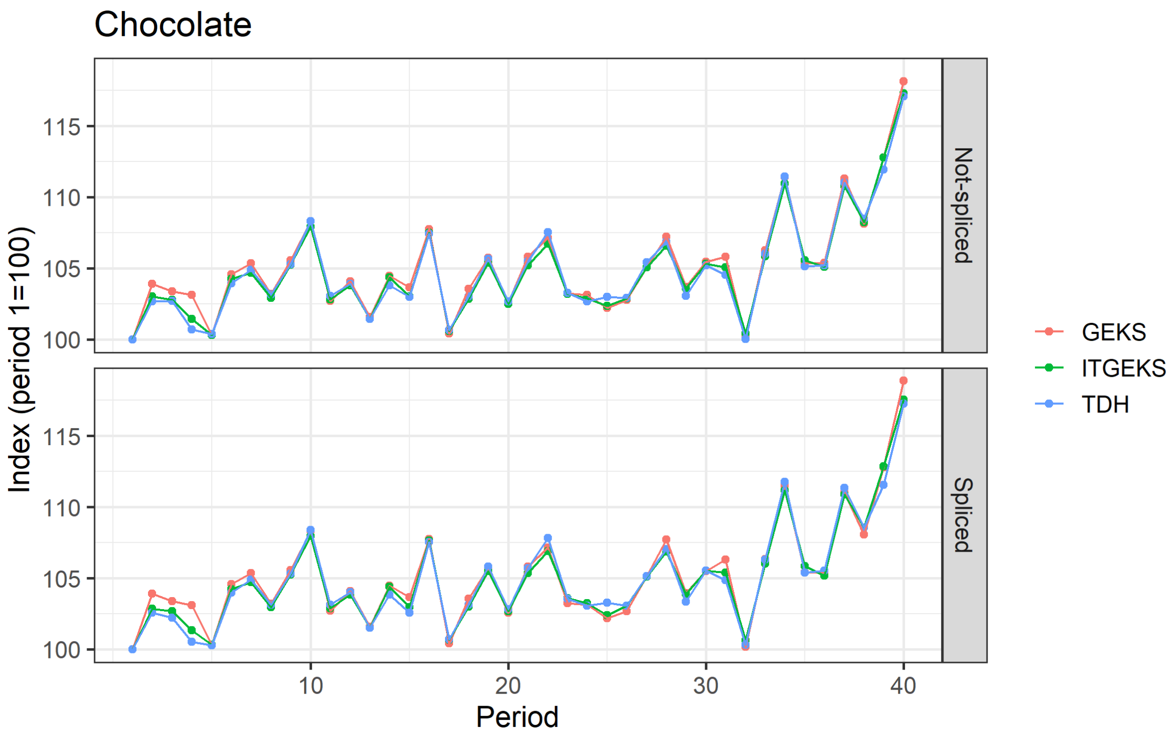
As the above figures show, three key observations can be made. First, as expected, the GEKS index with unique product identifiers does not capture the product relaunches and gives downward biased results. In both cases, the GEKS index shows no price increase at all. It underestimates the price level for cream cheese by around 15 index points and for cat food by 10 index points. In fact, the GEKS index for cream cheese even shows a small price decline when prices are increasing. Second, both the

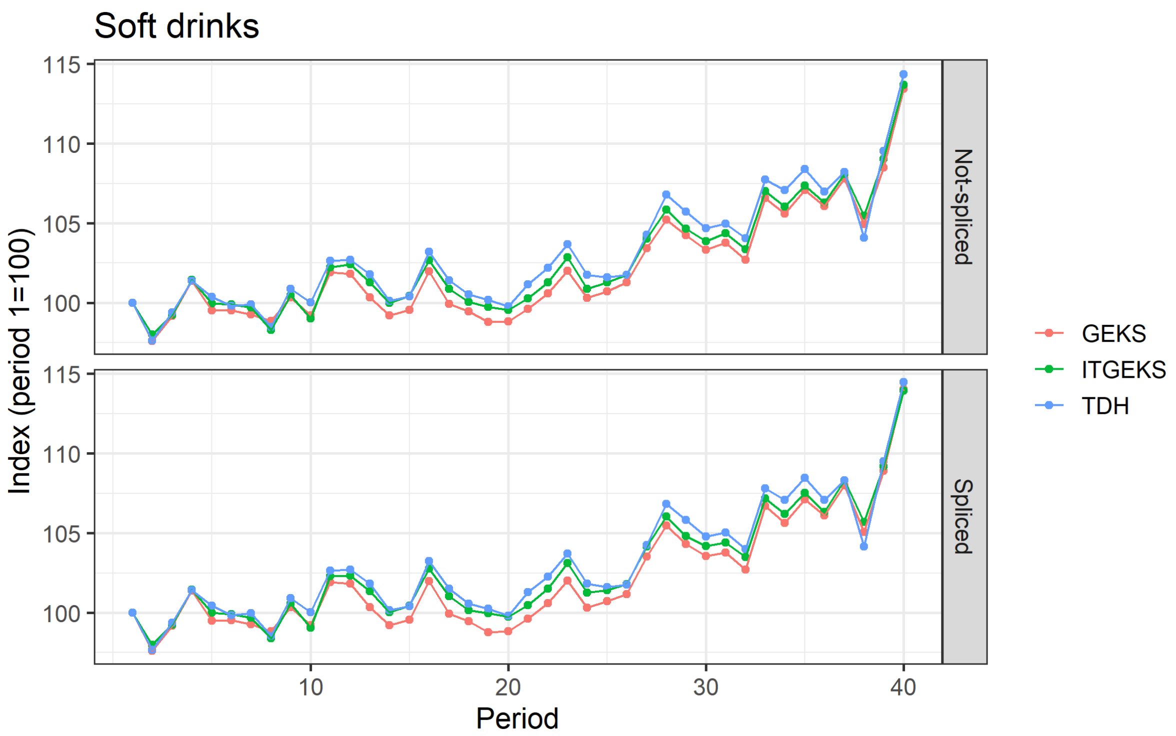
ITGEKS and the TDH do show an increase when the shrinkflation takes place in both product groups. This suggests that both methods can account for the change in product characteristics (including package size) to show that prices have increased in both segments. Third, the difference in results between the ITGEKS and the TDH is limited. This may be explained by the high R squared values and indicates that parameter fixity for the whole window in the TDH is not that problematic. As emphasised earlier, the TDH is much easier to compile than the ITGEKS. The ITGEKS more closely matches the GEKS index in both examples, although the difference with the TDH is very small in both examples.

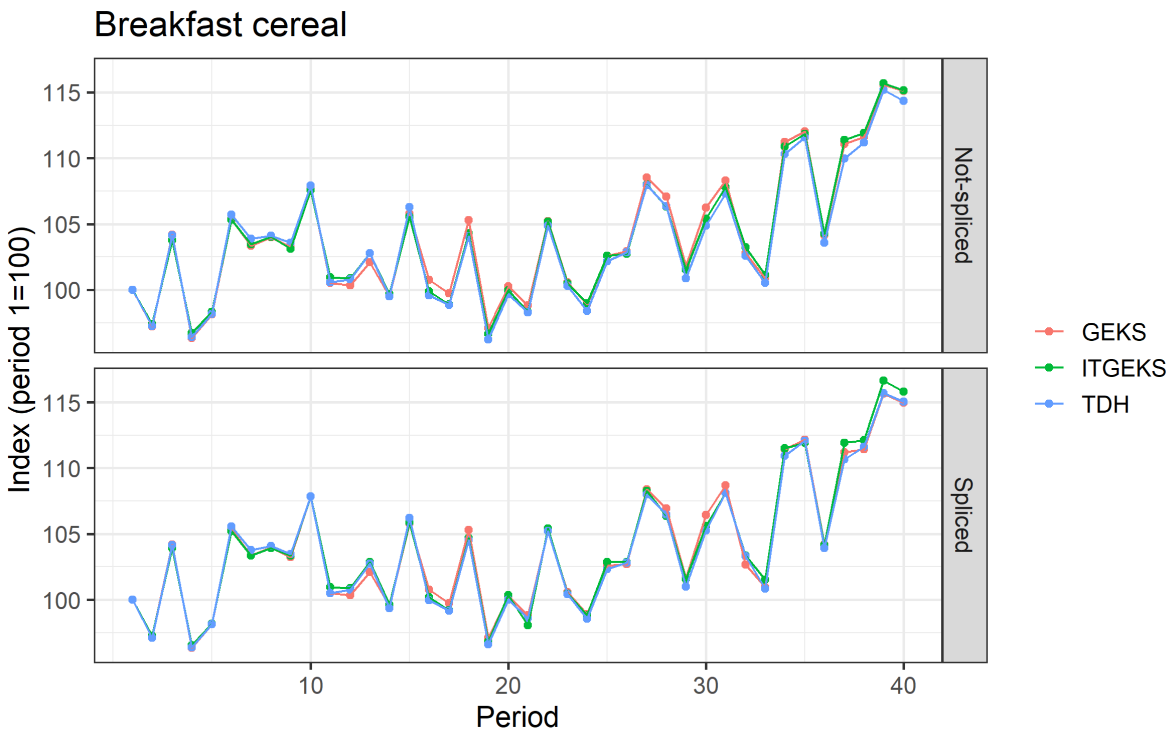
We now extend the analysis to 4 other product groups randomly selected from segments with product attrition, but where the linkings by price collectors showed no significant effect on the index: coffee, chocolate, soft drinks and breakfast cereals. For these product groups, the GEKS index could serve as a “benchmark” against which we can compare the TDH and ITGEKS. Of course, this is based on the assumptions that the use of our semi-automatic procedures for capturing product relaunches has not resulted in any missed problematic relaunches. A 40-month period is examined, requiring 780 bilateral time dummy hedonic regressions for the ITGEKS for each product group. We also compare an index using the entire window to an index with splicing (half-splice on published indices with a rolling window of 25 months) to see if splicing has any effect on the results. The results are shown in the figures below.

All four segments show similar results over the long-term, with minor deviations in the short-term. The results for the ITGEKS and the TDH also indicate that there are indeed no problematic product relaunches for those four product groups. The largest difference is noticeable for the TDH for soft drinks in certain periods. However, the R squared was also the lowest for this product group. The good performance of the much easier to calculate TDH, suggests that limited parameter fixity might not be that problematic. As with the cream cheese and cat food examples, the ITGEKS tends to be slightly closer to the GEKS. However, in most cases the differences between the TDH and the ITGEKS are marginal. Results between spliced and non-spliced indices tend to be almost identical and do not change the conclusions. This is good news from a production point of view, as only splicing can be used for the monthly production of the CPI/HICP.


Coffee
Index (period 1=100)
Not-spliced
Spliced
Period
GEKS
ITGEKS
TDH


Chocolate
Index (period 1=100)
Not-spliced
Spliced
Period
GEKS
ITGEKS
TDH


Soft drinks
Index (period 1=100)
Not-spliced
Spliced
Period
GEKS
ITGEKS
TDH


Breakfast cereal
Index (period 1=100)
Not-spliced
Spliced
Period
GEKS
ITGEKS
TDH

5. Conclusion

Using scanner data from the Belgian CPI/HICP, this paper has shown that multilateral index methods which use hedonic quality adjustments, such as the Imputation Törnqvist GEKS or the Time Dummy Hedonic method, can be used to compile accurate price indices for supermarket scanner data. Both the ITGEKS and the TDH gave comparable results to the GEKS using unique product identifiers for product groups where there was product churn but no problematic relaunches or shrinkflation. We have also shown that these methods can successfully capture product relaunches or shrinkflation using the examples of cat food and cream cheese. Under hedonics, products with similar characteristics are implicitly considered to be the same product (or correctly quality or quantity adjusted for), and the product relaunch is then captured. The GEKS, which uses unique product identifiers, did not capture any of the shrinkflation that occurred and resulted in a significant a downward bias. In all the examples the differences between the ITGEKS and the TDH were limited. This suggests that parameter fixity in the TDH is not that problematic. A major advantage of the TDH is that it is much easier to compile and is easier to implement.

Detailed product information is required to produce an accurate ITGEKS or TDH. Whether these methods save resources compared to the semi-automatic procedures we currently use depends on the quality of the data. For the analysed product groups, some (limited) data cleaning was necessary.

Even if countries are unable to extract useful product characteristics from their scanner data (or enrich them by scraping information from retailers' websites) in order to run reliable hedonic regressions, it may still be worthwhile to compile the other two indicators mentioned in this paper. The first indicator is the one that tracks, per product group, what percentage of sales or expenditure from a few months ago can be matched with the current month. Significant movements in this indicator may indicate that there is a problematic product relaunch and that it is worth examining the product group in more detail. If information on packaging size is available, a unit value index using standardized unit of measurement could also be compiled as a second indicator. A combination of a sharp drop in matched sales and an increase in the price per kg or price per litre might indicate that a problematic relaunch is happening and that the compiled index might be biased.

While our results suggest that using the characteristics available in our scanner data one can capture a certain problematic relaunch (e.g. in the case of shrinkflation), it might still not be able to capture all product relaunches. For example, the use of lower quality ingredients may also coincide with a product relaunch and should theoretically be quality adjusted for. Our data does not allow us to capture these effects, although we would argue that such product relaunches are also difficult to capture in classical price collection and are therefore not a problem inherent to scanner data.

In the coming months, we will extend our analysis to other COICOP groups to see if our results can be generalised to the full range of supermarket products. We will also compare our results with methods that use stratification or product clustering.

References

Chessa A. (2019), MARS: A Method for Defining Products and Linking Barcodes of Item Relaunches - *Paper presented at the 16th meeting of the Ottawa Group, 08-10 May 2019, Rio de Janiero, Brazil.*

Chessa A. (2021), Extension of multilateral index series over time: Analysis and comparison of methods - *Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 2-10 June 2021, Geneva, Switzerland.*

Daalmans J. (2022), Multilateral indices and the relaunch problem: product clustering and alternative solutions - *Paper presented at the 17th Ottawa Group meeting, 7-10 June 2022, Rome, Italy.*

Dalén J. (2017), Unit Values and Aggregation in Scanner Data – Towards a Best Practice - *Paper presented at the 15th meeting of the Ottawa Group, 10-12 May 2017, Elville am Rhein, Germany.*

de Haan, J. (2004), The Time Dummy Index as a Special Case of the Imputation Törnqvist Index - *Paper presented at the 8th Meeting of the Ottawa Group, 23-25 August 2004, Helsinki, Finland.*

de Haan J. (2010), Hedonic Price Indexes: A Comparison of Imputation, Time Dummy and 'Re-Pricing' Methods - *Journal of Economics and Statistics 230, 772-791*.

de Haan J. and Krsinich, F. (2014), Scanner Data and the Treatment of Quality Change in Non-Revisable Price Indexes - *Journal of Business and Economic Statistics 32(3), 341-358*.

de Haan J., Hendriks, R. and Scholz, M. (2016), A Comparison of Weighted Time Product Dummy and Time Dummy Hedonic Indexes - *Graz Economics Papers 2016-13, University of Graz, Department of Economics*.

de Haan J. and Daalmans J. (2019), Scanner Data in the CPI: The Imputation CCDI Index Revisited - *Paper presented at the 16th meeting of the Ottawa Group, 08-10 May 2019, Rio de Janiero, Brazil.*

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

Radjabov B. and Van Loon K. (2022), Seasonal Products and Multilateral Methods - *Paper presented at the 17th Ottawa Group meeting, 7-10 June 2022, Rome, Italy.*

Van Loon K. and Roels D. (2018), Integrating Big Data in the Belgian CPI - *Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 7-9 May 2018, Geneva, Switzerland.*

Van Loon K. (2019), Redefining what products are in the context of scanner data and web scraping, experiences from Belgium - *Paper presented at the 16th meeting of the Ottawa Group, 08-10 May 2019, Rio de Janiero, Brazil.*

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

Van Loon K. (2021), Implementing transaction data for consumer electronics and household appliances - *Report written for Eurostat grant project.*

Wood Z. (2023), Cat-astrophe? Whiskas pet food criticised over shrinking servings – *The Guardian, https://www.theguardian.com/business/2023/mar/24/cat-astrophe-whiskas-pet-food-creates-a-racket-over-shrinking-servings*