The Choice between Bilateral and Multilateral Index for Scanner Data: Case Study on Austrian Grocery Scanner Data 2022-2023

Adam Tardos
Statistics Austria
Department of Price and Parity

Summary
Multilateral index methods are often considered more appropriate than bilateral index methods when using scanner data for price index compilation purposes for various reasons. Nevertheless, multilateral methods require advanced index compilation procedures, longer data series, come along with additional complexity, and are difficult to explain to users. Since the introduction of scanner data in 2020 using a bilateral method, Statistics Austria has changed the index compilation method to the multilateral GEKS method. This paper analyses the advantages and pitfalls of the two main approaches for the compilation of CPIs and compares the performance of these methods, particularly with regard to the problem of chain drift. Furthermore, the paper discusses the issue of communicating the choice between bilateral and multilateral index methods to users and the pragmatic approaches taken in when introducing a multilateral method for the Austrian CPI/HICP.
Statistics Austria has been receiving regular data deliveries from the grocery and drugstore retail trade since the beginning of 2020. The original plan was to introduce scanner data into the Austrian CPI and HICP in January 2022 after a two-year transition period. However, the outbreak of the Covid-19 pandemic led to significant changes in the implementation plan. To protect the health of price collectors, scanner data were put into production almost immediately after the start of data deliveries. Due to the short time period and the available data series, a bilateral index calculation was applied. During this period, conditions were established to test multilateral methods. In addition, pragmatic decisions were taken on a number of issues, ranging from how to establish a good relationship with data providers, the method of data access, the classification of products and the choice of the appropriate index calculation and aggregation method. Finally, in order to avoid chain-drift effects as well as for practical and communication reasons, Statistic Austria have opted for the multilateral GEKS index method.
As another two years have passed since the introduction of the scanner data there is now a sufficiently long time window to test and analyse the chain-drift effect on Austrian supermarket data over a two-year period in 2022-2023. We have compared the multilateral index used in production with different bilateral methods, for example with the one used at the time of the Covid restrictions. The impact of chain drift was measured both at the level of the overall index and at the more detailed level of the ECOICOP 5-digit, with a particular focus on certain product groups that are more exposed to seasonality. This paper presents the results of this analysis.

Keywords:
CPI, HICP, Scanner data, Multilateral method, Bilateral method, GEKS, Windows length, Seasonality

Table of Contents

Background

The current and recent retail environment is characterised by continuing technological advances and a remarkable diversification, with significantly expanded product ranges and more pronounced segmentation of product categories. At the same time, shifts in pricing dynamics add further complexity and present new challenges to the process of consumer price index (CPI) measurement. In response to these challenges, the integration of scanner data into consumer price statistics represents a significant step forward in terms of quality. The use of sales volume, turnover values and comprehensive coverage of the reporting periods, together with a wide range of products, will further ensure the quality of the CPI in the future.

Since 2010, Statistics Austria had been working on obtaining scanner data and calculating price indices from them. Initial negotiations with potential data providers to provide data on a voluntary basis failed and therefore a legal obligation for mandatory scanner data deliveries had to be introduced. In December 2019, the Austrian national CPI-Regulation introduced the scanner data requirements and ensures scanner data deliveries by the major retailers, initially by the grocery and drugstore retail trade. After a two-year test period, scanner data were introduced into the Austrian CPI and HICP in January 2022, mainly for food and drugstore products.

Throughout the implementation of scanner data, and particularly during the testing phase, many decisions have to be made, often involving conflicting methodological and practical considerations. These decisions include selecting data providers, establishing data storage protocols, classifying products, filtering data by product or over time and, crucially, determining the most appropriate index calculation methodology.

It is known that one of the advantages of scanner data is that the time coverage of the data is much more comprehensive than the spot data from the conventional price surveys in the outlets. Ideally, scanner data are available for every week of the month. Obviously, from a theoretical point of view, the more weeks of data we build our index on, the better the representation of the given month. However, from a practical point of view, given the tight publication deadlines, it is questionable whether there is enough time to calculate the indices and implement thoroughly all quality control mechanisms, if one waits until the data of the last calendar week of a given month arrives.

When selecting an index method, a decision has to be made whether to choose between one of the well-established bilateral methods or a multilateral method that is more suitable for scanner data and more resistant to chain-drift effects.

Bilateral methods are effective for analysing static datasets where most prices have corresponding matches between the base and measurement months. This effectiveness is due to the fact that bilateral approaches typically involve a small number of products, and even when products leave the market, matches can often be ensured by product replacement. However, alternative data sources offer superior coverage of real markets, which are characterised by their dynamic nature. Dynamic markets are characterised by the continuous entry and exit of products, leading to inconsistencies in observed prices. Consequently, bilateral methods may be less effective when applied to dynamic data, where unobserved prices in either the base month or the measurement month may render a product unusable in the calculations. On the contrary, multilateral methods perform excellently when dealing with dynamic data. Given the objective of making full use of scanner data, multilateral indices offer a solution to accommodate the inherent dynamism of these large datasets.[1]

However, in some practical scenarios the use of a multilateral method with an appropriate window length may not be feasible, leading National Statistical Institutes (NSIs) to opt for bilateral index calculations. This decision may be based on a conservative approach to maintain methodological consistency despite changes in data collection methods. Alternatively, it could occur when the NSI

[1] Office for National Statistics (ONS), released 28 November 2022, ONS website, methodology article, GEKS Törnqvist: introducing multilateral index methods into consumer price statistics

is forced to adopt scanner data-based indices despite the lack of a sufficient dataset of scanner data (e.g. due to short time series). This was the case in Austria, where the planned two-year test phase was interrupted by the outbreak of the Covid-19 pandemic. The regular delivery of scanner data had only started three months before the first Covid-19 lock-down in March 2020. The use of scanner data allowed the suspension of traditional data collection in the grocery sector to protect the health of price collectors, as it became possible to produce the HICP without traditional data. Indeed, the opportunity to consistently monitor both traditional and scanner-based index series over a two-year period was lost. In addition, given the short timeframe between the start of the closure measures and the monthly HICP release date, Statistic Austria had to act quickly. Consequently, it was decided that bilateral indices derived from scanner data would be used due to the limited timeframe and lack of historical data. During the 2020-2021 trial period, a further four lock-down months resulted in the interruption of conventionally collected price data. Although it would have been possible to introduce a multilateral methodology in the second year of the test-period, Statistic Austria opted to maintain methodological consistency in the context of an already turbulent period. As a result, the final implementation of the multilateral method has been postponed until January 2022, in line with the original plan.

When implementing scanner data, even when opting for a multilateral index, there are various methods available, each with its own advantages and disadvantages. However, the process does not end with the selection of the most suitable method. Each method can be applied with different parameters. It is necessary to decide on the splicing method to be used each month to link the multilateral index chains and, most importantly, the number of consecutive months on which the multilateral index will be based.

The appropriate window lengths have been tested by several experts. Chessa[2] found that the use of 13-month windows can be sensitive to downward drift, especially in case of seasonal items. Kevin J. Fox, Peter Levell and Martin O'Connell[3] concluded that chain drift bias falls significantly as the window size increases.

From a methodological point of view alone, it seems advantageous to use a window of maximum duration, ideally at least 25 months. However, it must be recognised that, despite a two-year testing period prior to the implementation of a new methodology, the availability of sufficient data may remain uncertain. The accessibility of historical data depends on the co-operation of data providers, their technological capacity and the regulatory framework in place. In cases where historical data is not available, a period of 25 months of scanner data collection is required before experimentation with 25-month window lengths can begin. In practice, the length of the experimental period prior to the introduction of scanner data is limited in order to limit concurrent data collection efforts and to reduce respondent burden.

Another consideration is the seasonal composition of the target sector for scanner data coverage. As outlined in the academic discourse, indices that are predominantly seasonal in nature will benefit from longer window lengths. It is worth noting, however, that longer window lengths impose greater demands on computational resources, with a fourfold difference between 25 and 13 month window lengths.

For these practical reasons, Statistics Austria has introduced the scanner data into the CPI with a window length of 13 months. Later in 2023, Statistic Austria examined the extent of the difference between indices based on 25-month and 13-month window lengths over longer time series.[4] The results showed that the annual inflation rates calculated from the two respective windows did not differ significantly. While small discrepancies were observed, these discrepancies largely balanced each other out, especially at higher levels of aggregation. Nevertheless, it was observed that, overall,

---

[2] Chessa, A.G. (2021) Extension of multilateral index series over time: Analysis and comparison of methods, Paper written for the 2021 Meeting of the Group of Experts on Consumer Price Indices

[3] Fox, K. J., Levell, P., O'Connell, M. (2022) Multilateral index number methods for Consumer Price Statistics

[4] Tardos, A. (2023) Introduction of Scanner Data into Austrian CPI and HICP – practical implementation experience, with a focus on window length options

annual inflation rates for lower COICOP categories tended to be slightly higher when applying a 25-month window compared to a 13-month window. In addition, the analysis showed that in cases where an elementary aggregate showed seasonality, the difference between the two methods increased, with the longer window approach typically resulting in higher inflation measures.
In this study we examine the impact of using bilateral indices on dynamic scanner data over longer periods. In particular, we examine whether the chain-drift phenomenon leads to significantly lower inflation rates when using the bilateral method and whether certain product categories show more pronounced divergences in this respect.

Description of the Data

The Austrian CPI Regulation defines the periodicity of scanner data delivery, including turnover shares and the survey period. In contrast to the traditional survey, which usually only captures the current prices on a certain day (reference date), the scanner data provision covers a defined period, for which the turnover achieved and the quantities sold per item are determined and a so-called unit value (average value) is calculated. In order to ensure data consistency, weekly data provision is required, as well as timely transmission for processing purposes. The scope and characteristics of scanner data require a change in CPI/HICP calculation procedures and methods. Consequently, a gradual introduction of scanner data into the CPI/HICP production process has been foreseen, starting with the scanner data of the enterprises classified in (Ö)NACE classes 47.11 (Retail sale in non-specialised stores with food, beverages or tobacco predominating) and 47.75 (Retail sale of cosmetic and toiletry articles in specialised stores), selected by cut-off sampling in accordance with the Regulation (small and medium-sized enterprises are excluded). The data of the enterprises in these (Ö)NACE classes were particularly suitable for the introduction of scanner data, due to the dominance of the five largest retailers in the food and drugstore sectors, which together account for more than 85% of the market share. In addition, the product groups they mainly trade have a high weight of around 16% in the CPI shopping basket (including food, beverages, daily consumer goods, drugstore goods).
Table 1 describes the properties and characteristics of scanner data as provided by the obliged retailers for each item sold per postcode and calendar week.

Table 1 - Scanner data variables and values

Variables	Example(s)
Article number and EAN/GTIN (if available)	130404 (Art-nr.); 9100000742175 (GTIN)
Article name or description	Red Bull 250 ml DS
Content quantity and unit	250 ml
Classification code and name of the article-related product group, in as much detail as available.	Drinks/alcohol-free drinks/energy drinks
Sales volume	235
Sales value	315 EUR
Date (from - to, or calendar week)	07.11.22-13.11.22; (2022_45)
Postcode to which the local shop relates	1060

In 2023, the retail sectors covering the sale of clothing (NACE 47.71) and the sale of footwear and leather goods (NACE 47.72) have been selected from a number of potential sectors for the expansion of the consumer price index (CPI) price survey using scanner data. These markets are more fragmented and therefore traditional data collection in smaller shops and automated online price collection (web scraping) will continue to be important alongside scanner data. This initiative is currently in a trial phase until 2025 and requires a partly different approach to that applied to scanner data for food, which is not the focus of this paper.

Data Preparation and Verification

The process of managing data from suppliers involves a number of automated steps to ensure accuracy and reliability. First, files received from data providers are automatically transferred, imported and subjected to a series of validation checks. These checks are designed to verify the integrity of the incoming data and cover various aspects such as weekly sales per supplier, the number of postcodes from which data was sourced in the current week, and the number of product groups and new products sold.

In cases where discrepancies or irregularities are detected within the data patterns, proactive measures are taken. This includes contacting the data provider to resolve the discrepancies. Subsequent actions include confirming the plausibility of the data or requesting a repeat delivery of the data to ensure accuracy and consistency.

Following the overall validation procedure, the validated data is then loaded into a DB2 database for further analysis and use. This database serves as a central storage and management system for the validated data, allowing efficient access and data processing.

Product Classification

Product classification is one of the most complex tasks of the scanner-data-based method. The classification process was taken place in two stages: following the initial data deliveries in the primary classification phase, a significant number of products have been categorised to create a comprehensive classified product base, which subsequently served as a training set. Following this initial phase, thousands of new products are delivered each month which demands a regular classification. To reduce the workload associated with these recurring classifications, a high degree of automation is required in their processing.

During the test period, a blended classification system was developed, based partly on an automated matching procedure using GTINs and product names, partly on several machine-learning methods and partly on a manual procedure. At COICOP-5 level, 90-95% of products are classified fully automated based on three models: Support Vector Machine and Naive Bayes or more recently on Long Short-Term Memory Neural Network. Input variables are item labels and retail categories. In the context of classifying products into ECOICOP-5 classes, several machine learning (ML) approaches were evaluated to identify the most effective model for the task. However, a priori determination of the optimal model proved challenging due to the inherent strengths and weaknesses of each approach. Empirical evaluation on the dataset was therefore essential to determine the most appropriate algorithm. The results of the classification task showed that none of the ML models employed had a clear superiority over the others, as evidenced by comparable hit rates above 90% for all methods. This parity of performance underlined the high level of consensus between the models.

Attempts to improve classification accuracy by implementing a higher-level meta-ML model did not provide significant improvements over the individual models. Consequently, this direction of research was not followed. Nevertheless, the analysis showed that disagreements between models could be a valuable source of insight. Such differences often indicated the presence of particularly difficult products to classify, suggesting an increased likelihood of misclassification by the models. Conversely, consensus between models was an indicator of reliable classification results. Thus, three models are used for automatic classification, and agreement between all three models was required, a criterion met in over 90% of cases. In cases where there was no consensus between the models are considered as problematic products.

The COICOP-5 classification of such problematic products, as well as the classification into finer categories than COICOP-5, is done manually. Manual classification is carried out in an in-house developed application.

Index Calculation

There are different approaches - bilateral and multilateral methods - to calculate a price index with scanner data at the elementary aggregate level.
Bilateral concepts are based on the comparison of two periods (base and comparison period). Such approaches are based on the standard theory of bilateral price indices. This approach is well understood, transparent and can be easily explained to users.
However, as mentioned above, bilateral indices based on scanner data have limitations and drawbacks: limited product coverage due to decreasing product matches over time due to product discontinuations, lack of consideration of item sales in the sample and also the risk of chain drift when updating the base period or monthly chaining or due to over-consideration of items with promotional prices.
These disadvantages of bilateral approaches can be avoided by multilateral methods. In fact, chain drift is a violation of the multi-period identity test that must be prevented. This test requires that if all prices and quantities in a period T return to their values observed in the base period 0, the index should show no price change. Multilateral indices satisfy this test[5].
During the transition period in 2020 and 2021, we compared a number of bilateral and multilateral index calculation methods, which allowed us to choose the most suitable solution for us according to theoretical and practical criteria.

Temporal Basis for the Indices

A key consideration is how much data should be used for the index calculation. Since data providers deliver data on a weekly basis, using data from one, two and three calendar weeks per month is optional. Four calendar weeks were out of the question, as not every month contains four full calendar weeks, and the aim was of course to cover the same length of time each month.
Initial test calculations indicated that the scanner data indices are somewhat more volatile than traditional CPI indices. However, the more calendar weeks the index is based on, the more moderate the fluctuations are. The intention is therefore to maximise the use of calendar weeks, aiming at three weeks per month.
It's important to note that there is a delay of a few days between data arrival and processing, which is a practical challenge, particularly in terms of meeting publication deadlines. To overcome this problem, the Austrian CPI/HCIP flash estimate published at the end of the month is compiled using scanner data from two calendar weeks of the current month, and the final index is completed with data from the third week.

Content Data Basis for the Indices

Scanner data provide comprehensive information on a wide range of products, suggesting the potential to extend index calculation beyond the narrowly defined CPI basket items (elementary aggregates) below the COICOP-5 level. Instead, consideration could be given to compiling the index at the COICOP-5 level, covering all products within the relevant COICOP category.
This shift would be attractive as it would allow indices to be based on a rich set of product data while potentially simplifying classification procedures. However, such a transition would have required the discontinuation of long-standing time series of elementary aggregate indices covering many years. Therefore, the transition to the 5-digit COICOP level was not carried out. Consequently, the index calculation continued to be based on products aligned to narrowly defined CPI basket positions (elementary aggregates).

---

[5] Guide on Multilateral Methods in the Harmonised Index of Consumer Prices, 2022 edition, Luxembourg: Publications Office of the European Union

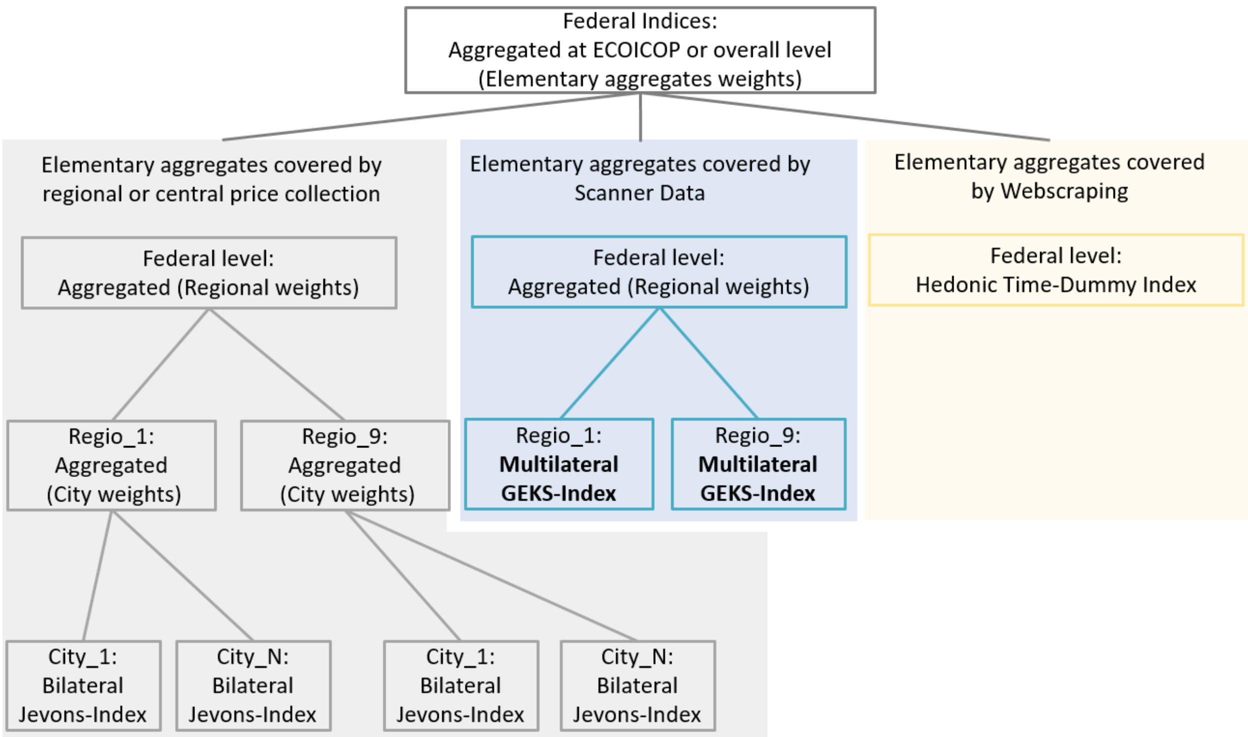
Junk Filtering, Outlier Filtering

Junk filtering removes products whose characteristics do not match the focus of the index. Classification, which is essentially a sorting process, allows the selection of products that fit into the appropriate COICOP category. Classification also eliminates products that lack sufficient information, as they cannot be categorised. Therefore, no further junk filtering is required after classification. In addition to the control mechanisms during data entry and the classification, an outlier search is carried out among the calculated unit values to exclude unrealistically high or low unit values before the index calculation. This outlier detection is used to identify products with extreme prices or price fluctuations that may indicate errors. We use broad relative fences to ensure that we do not flag highly volatile prices as outliers, but only unrealistic movements or errors.

Regionality and Aggregation Level

The CPI Regulation in Austria defines "survey regions for scanner data deliveries [...] by postcodes [...] ". The areas which are defined by the 346 postcodes listed in the annex to the CPI Regulation were selected to ensure representativeness at regional level. These postcodes are distributed across all 9 federal states, covering the 20 bigger cities where the traditional price surveys are carried out, as well as rural areas. This way, the elementary aggregate used to calculate the index is the unit value of products by retail chain and by federal state (region). At this level of aggregation, nine regional indices are compiled at the federal state level and then aggregated into a national index. By doing so, the procedure is harmonised with the index calculation methodology of the other survey types, the calculations of which are still based on a traditional, likewise hierarchical methodology: cities, regions (federal states) and country. For the regional weights, the same values are used for all items, regardless of whether it is the traditional or the new methodology.

Figure 1 – Aggregation levels of the CPI/HICP-Index



Index Calculation: Bilateral vs. Multilateral Method and Window Length

Multilateral methods are a special type of index compilation method that can be applied to scanner data. A price index usually measures the aggregate price change (at CPI basket position or COICOP 5-digit level) of the current period compared to a base period.
In multilateral methods, the aggregate price change between two comparison periods is determined from prices and quantities observed in several periods, not only in the two comparison periods. This is the great advantage of multilateral methods: they consider all products that are available in at least two periods of the observed time interval (time window). Multilateral methods have been used for many years for geographical price comparisons (e.g. between different countries or regions) of purchase price parities and have been adapted for temporal comparisons. Scanner data is typically dynamic. New products are constantly being added to the product range, while obsolete products that were previously available are removed. Bilateral price index methods compare the prices of products in the current period with prices in a past base period. However, as time passes, the overlap of products decreases, making it difficult to calculate price comparisons. One way to increase the overlap of products is to frequently update the base period and chain the resulting bilateral price indices. However, it has been shown that such an approach can be subject to chain drift, especially when products are explicitly weighted. Chained indices often lead to systematic distortions and therefore do not measure a plausible price change over longer periods.
Multilateral methods offer a solution to the problems of bilateral approaches. They take into account all products that are available in the different periods. They allow the explicit weighting of each product according to its importance in each period. Finally, they avoid the chain drift problems that arise with chained bilateral indices. Given these advantages, multilateral methods have been recommended as appropriate price index compilation methods for transaction data, despite their additional complexity compared to bilateral methods[6].
In order to use multilateral methods in the compilation of price indices, some data requirements must be met:

- Access to historical data: since multilateral approaches use the data of many months at the same time (time window), sufficiently long data series from the past are required to test and implement these methods (therefore the relatively long test period and implementation phase from December 2019 to December 2021).
- The raw data received must be pre-processed and classified (see check and classification steps above). As the multilateral methods are essentially based on all transactions, it is not necessary to select items by means of random sampling or to filter them out due to low turnover. Each product is included according to its importance. In practice, however, item records will still be excluded during processing and data control mechanism, if important information is missing (e.g. the turnover or commodity group code) or if they contain inconsistent values.

A multilateral index is constructed over a given time window length T consisting of a sequence of consecutive months. The index formula takes as input the prices (unit values) and quantities or turnover of the individual products available in the months of the given time window.
The first step in the calculation of all multilateral indices is to determine the length of the time window, which in practice means how many months of data a particular calculation should take into account. Given the seasonality of certain products, one of the most commonly used time window length is the number of months in the year plus 1, i.e. 13. This time window allows products that are only sold in one month of the year to be linked and thus have an impact on the index. Of course, it is possible to calculate with a longer time window (e.g. two years + 1 = 25), but this implies a longer data series

[6] Guide on Multilateral Methods in the Harmonised Index of Consumer Prices, 2022 edition, Luxembourg: Publications Office of the European Union
https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-21-020

and more calculation effort. Our calculations were tested with different time windows, but for the reasons given in the background chapter we considered 13 to be the optimal choice.
We tested the three theoretically well-founded methods recommended by Eurostat[6], the Gini, Eltetö and Köves, and Szulc (GEKS), the Weighted Time Product Dummy (WTPD), and the Geary-Khamis (GK) index, respectively.
As we found only minor differences between the indices for most items, we have opted for the GEKS index for practical reasons. Although all multilateral indices are based on a relatively complex methodological background, the logic of the GEKS index is most similar to that of the traditional bilateral indices and is therefore the easiest to communicate and to comprehend.
To calculate the GEKS index[7], a matrix of bilateral indices at a given time window must be constructed, and the corresponding bilateral index must be calculated for all possible pairs of months. This implies 13x13 = 169 index calculations for a time window of 13. If we consider the symmetry of the matrix and the fact that the diagonals of the matrix are all equally 1, this means in practice that 78 bilateral indices are calculated. At a window length of 25 months, the number of bilateral indices to be calculated increases by a factor of almost four (25x25-25)/2 = 300. The value of the GEKS index for a given time is the geometric mean of the corresponding bilateral indices.
The GEKS index between time periods 0 and t is calculated for a given time window W as follows:

$$\mathrm{I}_{W}^{0,t} GEKS = \prod_{k \in W} \left(I^{0,k} * I^{k,t}\right)^{\frac{1}{|w|}}$$

The bilateral Törnqvist approach can be interpreted as the GEKS method with a two-month window. When calculating bilateral indices from scanner data, the Törnqvist index is often preferred because it weights turnover data, which is one of the main advantages of scanner data. In addition, the Törnqvist index takes into account revenues in both months, making it a suitable choice for calculating bilateral indices from scanner data. In analysing the long-term differences between bilateral and multilateral indices at different levels of aggregation and for different products, we compare our introduced 13-window GEKS Törnqvist index with the bilateral Törnqvist index. We hold all parameters constant except the window length to ensure that the observed differences are solely due to the use of multilateral or bilateral indices.
Before proceeding, a brief digression is required. At the start of the testing phase, unforeseen Covid lock-down made it unexpectedly necessary to compile indices based on scanner data, as the price collectors were prevented from visiting the shops. With only three months of data available, the feasibility of using the multilateral method was not assured, prompting the calculation of a bilateral index. Consequently, we chose the bilateral index, but instead of the Törnqvist index, we opted for the unweighted Jevons index, for several compelling reasons:

- Alignment with traditional methods: Given the temporary nature of our reliance on scanner data for 1-2 months, after which we planned to revert to the conventional price survey index in the test period, we were looking for consistency with traditional approaches. The Jevons index used in traditional price collection methods fitted seamlessly into our transition strategy.
- Flexibility and convenience: The Jevons index offered practical advantages by allowing the calculation of product-level rates of change and then their geometric means. This methodology enabled us to avoid the need to calculate average prices, which posed challenges due to incomplete or inaccurate volume data, while still achieving comparable results.

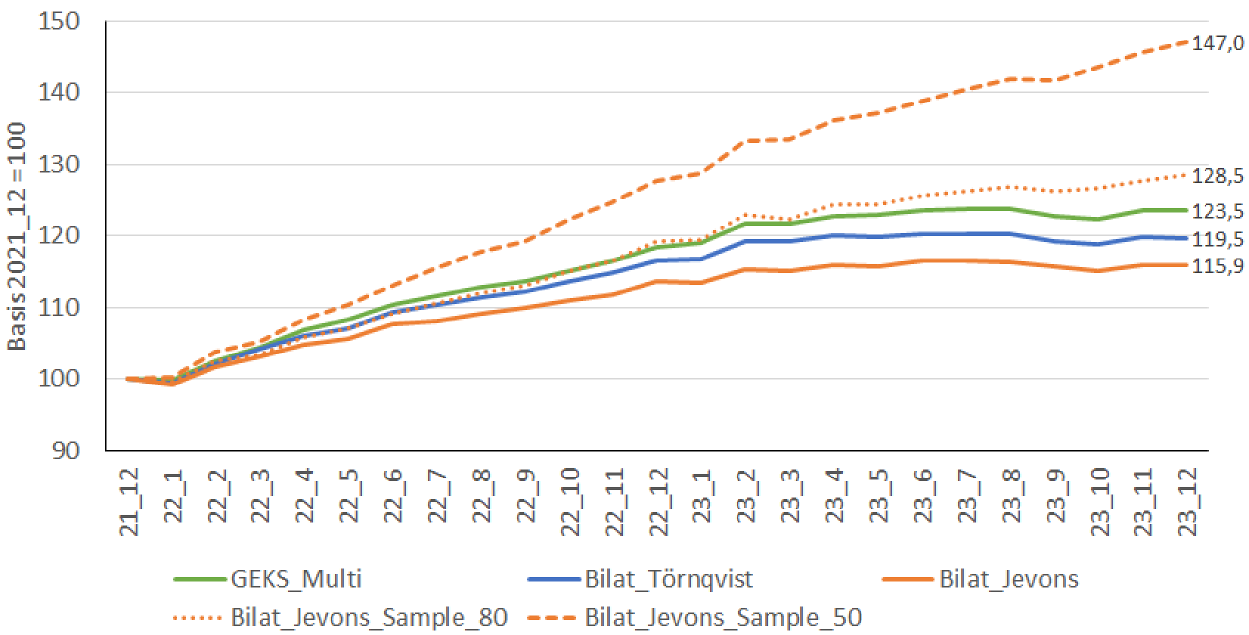
[7] Whenever a GEKS index is calculated, it is linked to a bilateral index method. This leads to many variants of the GEKS (e.g. GEKS-Fisher, GEKS-Törnqvist, GEKS-Jevons). The different variants are usually close to each other. It was decided to use the GEKS method with the Törnqvist index, accordingly by GEKS we actually mean GEKS-Törnqvist.

- Reducing potential chain drift: While we acknowledged the potential for chain drift associated with sales-weighted bilateral indices, our limited familiarity with scanner data led us to adopt a cautious approach. By opting for the unweighted Jevons index, we were able to avoid this concern.

In addition, we decided to use a product sample, a decision aimed at aligning our methodology even more closely with traditional indices and reflecting the limited subset of products typically monitored by price collectors. However, unlike the relatively stable product samples in price surveys, our approach was based on revenue data, resulting in monthly fluctuations. Our sampling method involved selecting the best-selling products within each chain, region and elementary aggregate until we reached a threshold representing 50% of total sales for each month. Hence, we excluded low-volume products, recognising that although their individual contribution to sales is minimal, their price movements have a similar impact on the unweighted Jevons index as the top selling products. Nonetheless, it is important to recognise that this sampling strategy on long-term introduces significant bias, which we discuss in more detail in the next chapter.

Comparing Bilateral Jevons, Törnqvist, and Multilateral GEKS Approaches for Food and Non-Alcoholic Beverages

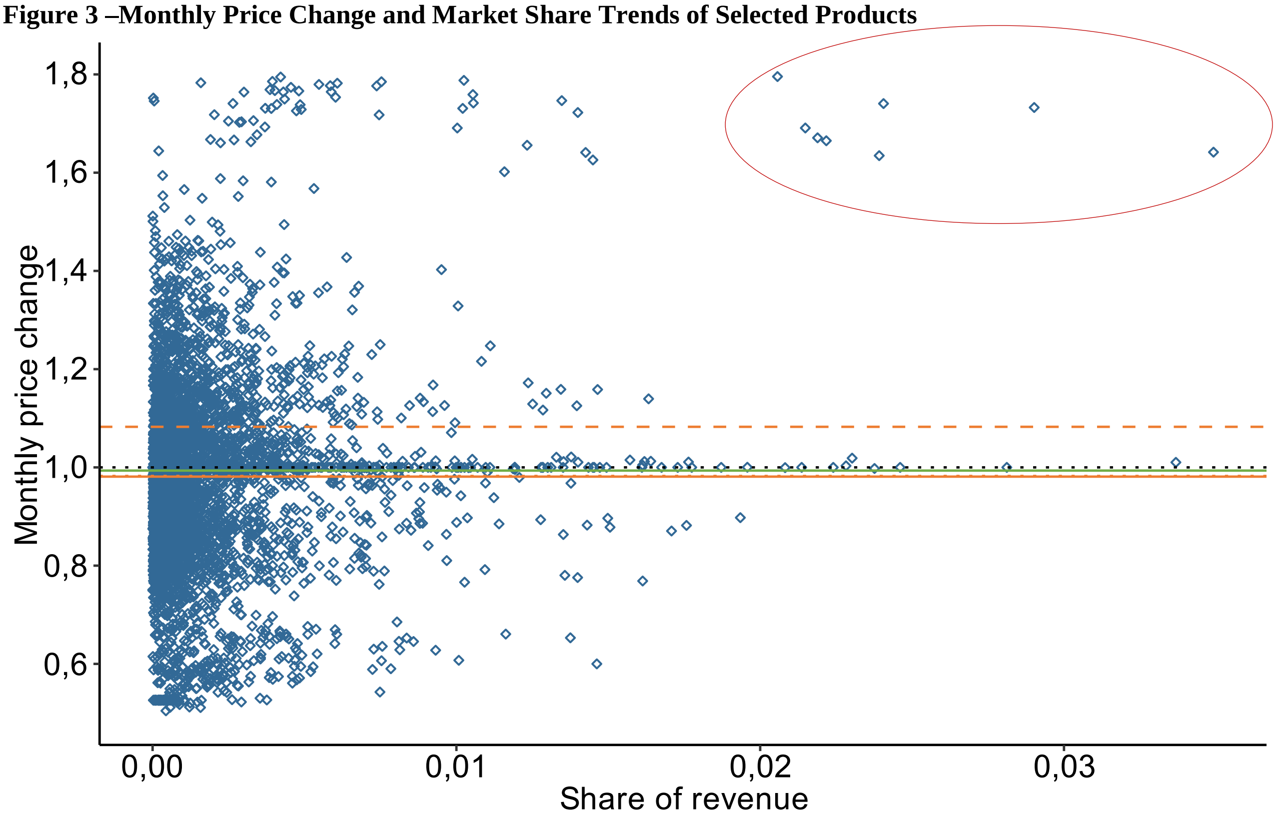
Figure 2 – Indices



The figure shows the index for the COICOP food and non-alcoholic beverages division (01), based on scanner data only, using both the multilateral GEKS and various bilateral methods. It can be seen that the unsampled bilateral indices, especially the unsampled Jevons index, show significantly lower values than the multilateral index over a two-year period. Calculated on the basis of December 2021, the multilateral index stands at 123,5 in December 2023, the bilateral Törnqvist at 119,5 and the bilateral Jevons at 115,9. The lower value of the bilateral Törnqvist is due to the well-known chain-drift effect. Similarly, chain drift affects the bilateral Jevons index, with additional distortions. The scanner data contains a large number of products that have a negligible revenue in a given month. While weighted indices mitigate these problems, unweighted Jevons indices give these products an undue influence, similar to that of the products with the largest market shares, and thus have a pronounced, albeit negative, distorting effect in our context.

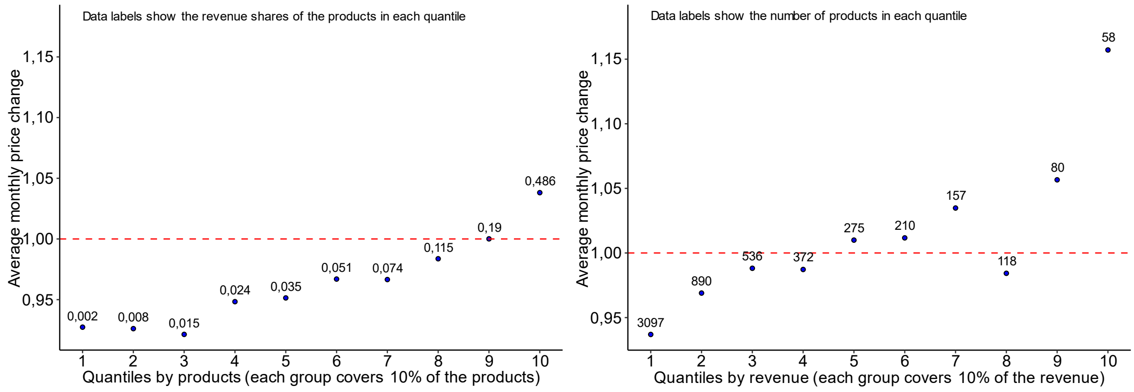
For the above reasons, one might consider excluding products that generate minimal revenue. However, as shown in the previous graph, the result is unpredictable. By taking a larger sample of products that account for 80% of revenue in a given month, the bias is already upward. If even fewer products, representing only 50% of the turnover, are included, the bias becomes significantly stronger. In our case, this methodological choice did not have a significant impact on the results because it was only implemented for a few months. The purpose of this transitional period was to maintain the use of the traditional method for production while testing the scanner data. Consequently, at the end of the lock-down periods, we resumed price collection, ensuring consistency by comparing current prices with the last valid price collected. This approach ensured that the bilateral Jevons method we used had no lasting effect on the index, except for the months in which it was applied, but it remains instructive to understand why the use of such a method may not be advisable. The accompanying graphs provide a deeper insight into how low-revenue products, or their exclusion, affect the Jevons Index.

Figure 3 –Monthly Price Change and Market Share Trends of Selected Products



The scatterplot shows the price change of more than 5.700 products (unit-values per chain and per region) of coffee beans in the database from April to May 2023 as a function of the market share of the product. On the y-axis, points greater than 1 represent products whose prices have increased and points less than 1 represent products whose prices have decreased. As a reference, we note that the multilateral GEKS measure for this month is -0,65 price decrease (green line), with a value of 0,9935 on this y-scale. By comparison, if we calculate using the unsampled Jevon index, the price decline is, as expected, higher at -1,9%, i.e. 0,9812 on the y-scale (orange line). On the other hand, if we sample the products, taking into account only the top products that generate 50 percent of the total turnover in the current months, we measure a price increase of 8,3 percent (1,083) with the Jevons index (orange dotted line). Of course, our choice of this particular month and product is based on the strong contrast it presents, making the distortion easy to see. In particular, products within the red circle have a significant market share of 2-3%, accompanied by a significant increase in price.

Figure 4 – Monthly Price Change of Selected Products by Quantiles



In Figure 4, we see that around 570 products, representing 1/10 of the total, contribute a massive 48,6% of total sales, while the most profitable 58 products alone account for 1/10 of sales. Conversely, the least profitable products, numbering 3.097, contribute only 10% of revenue.
The chart also shows a strong relationship between the price change of products and market share in this month, with the prices of the higher market share products tending to increase, so that the sample is strongly biased upwards and the non-sample downwards. This difference poses a challenge in matching the Jevons method with scanner data.
Adopting a fixed sample over a longer period - say, a year - similar to traditional price collection, could mitigate this bias, but introduces its own set of challenges such as product churn and substitution. It also risks compromising two fundamental strengths of scanner data in index production: comprehensive sales data and product diversity.
With this exploration in mind, we now look at the bilateral method, specifically the Törnqvist bilateral approach, which is better equipped to handle the issues raised by scanner data and thus deserves comparison with the multilateral method.

Comparing the Törnqvist Bilateral and Multilateral GEKS Approaches Across Different Levels of Aggregation

In this comparative study, the GEKS multilateral index and the Törnqvist bilateral index were calculated using the same data series. The only difference was the time period over which these indices were calculated: the bilateral approach was based on a two-month period, while the multilateral approach was based on a thirteen-month period. The analysis was carried out from the start of the scanner data integration into the Austrian Harmonised Index of Consumer Prices (HICP) until December 2023, covering a two-year period with December 2021 as the base month. In both methodologies, the index was calculated 24 times, similar to production cycles, allowing for monthly increments. The new index values for each month were then linked to the existing index series using a movement splice technique.
We calculated annual inflation rates from the two indices for each month in 2023 and compared these annual inflation rates and their averages. The differences were compared at different COICOP levels, starting from 1-digit level (total CPI) up to COICOP 5-digit level. The comparison has been restricted, as appropriate, to the COICOP groups involved in the introduction of the scanner data.

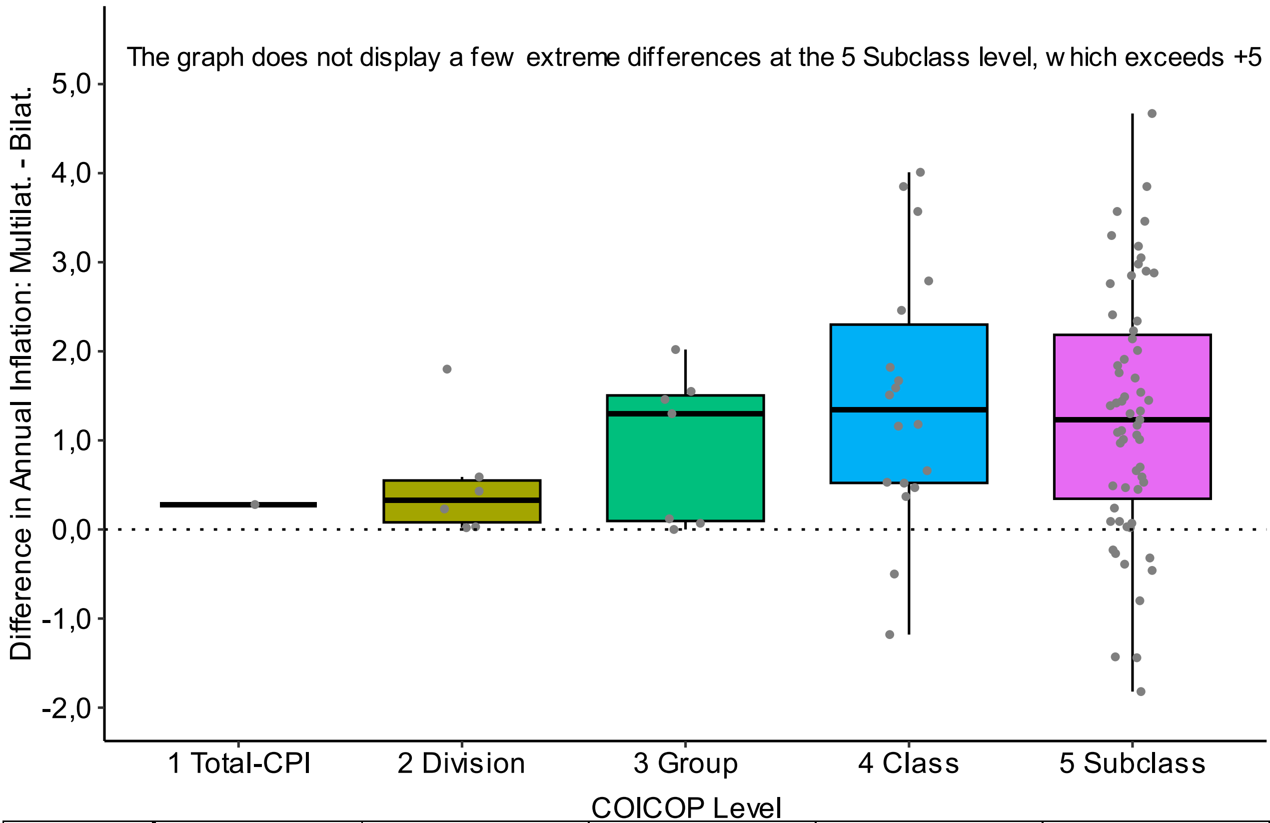
Impact of Törnqvist Bilateral on the Overall Index

Table 2 - Number of COICOP categories affected by introduction of scanner data at different COICOP levels

COICOP level	Number of categories	Average weight of the scanner data
1	1	16%
2	6	30%
3	7	65%
4	19	99%
5	62	100%

COICOP 1-digit level covers the entire consumer basket. The coverage of the scanner data on this level is 16%. At 2-digit level, the scanner data covers for instance division 01 (food and non-alcoholic beverages), and partly division 02 (alcoholic beverages, tobacco), or division 12 (miscellaneous goods and services)
The coverage for food is close to 100%, while for example the coverage for group 12 is 15%. The average for the 6 groups is 30% as shown in the table. Once again it is important to note that groups not covered at all by the scanner data (e.g. 07 Transport) are not included in the average.
The lower the COICOP level, the higher the coverage of the groups. At COICOP 5-igit level, the coverage of the groups concerned is 100%.
Of course, if the indices in a given group are calculated using not only scanner data, this reduces the impact of the different methodologies, as the sub-indices calculated using the traditional method are not affected by the method applied to the scanner data. Still, it is very important to see what impact the bilateral index would have had on the overall index.

Figure 5 – Difference in average annual inflation by COICOP level (2023)

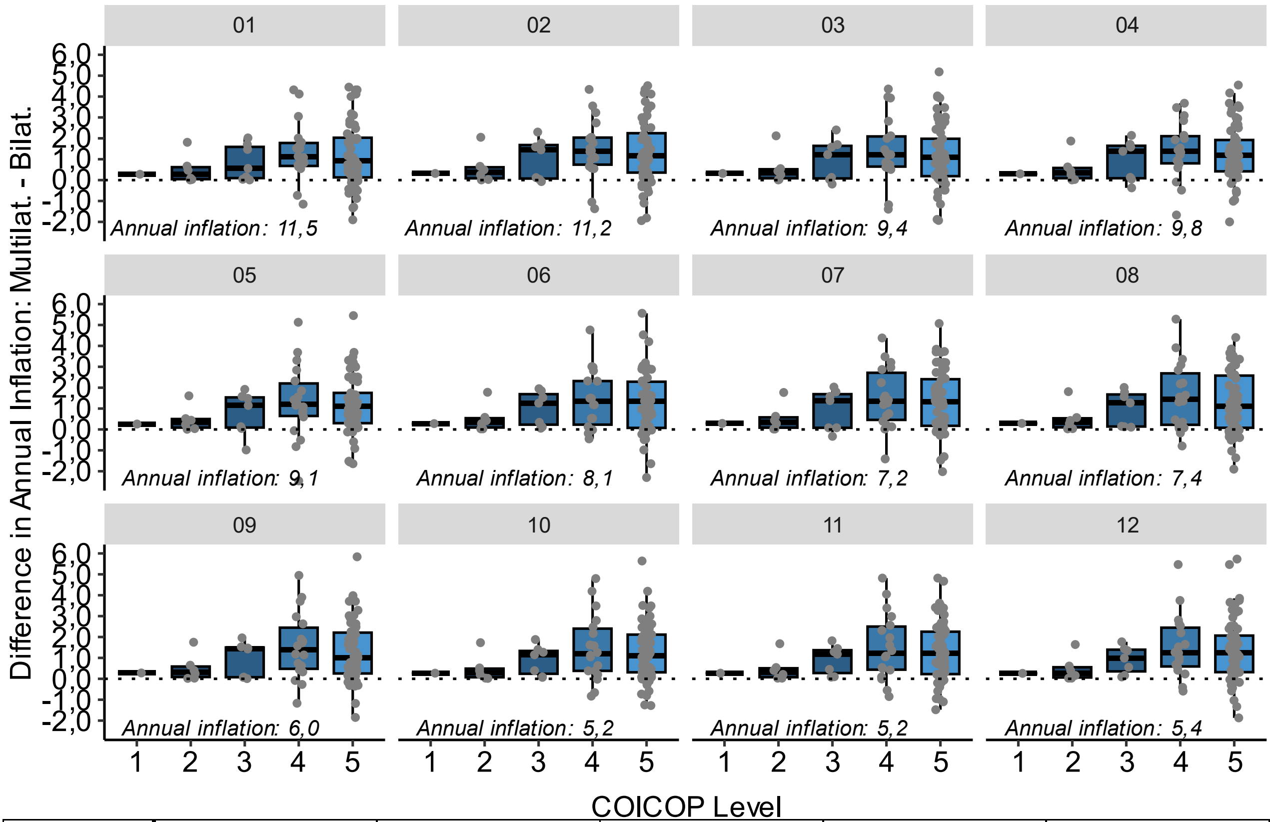


Avg. M	8,0	7,6	9,3	10,1	12,3
Avg. B	7,7	7,1	8,3	8,2	10,5
Δ(M-B)	+0,3	+0,5	+1,0	+1,9	+1,8

The box plot in Figure 5 shows the differences in average inflation in 2023 at different COICOP levels, depending on whether a multilateral GEKS or a bilateral Törnqvist is used. The grey dots show the differences within each COICOP category. A positive difference means that the inflation calculated using the multilateral method is higher, and a negative difference means the opposite. The horizontal jittering of the points along the symmetry axes of the box plots is for illustrative purposes only, so that the overlapping points can be seen. The lower the level of COICOP, the greater the spread of differences around 0. The points are distributed in both positive and negative directions around 0, but there are far more categories of COICOP with a positive spread. Out of 62 5-digit COICOP sub-classes, 53 have positive differences and only 9 have negative differences.

The table below the graph in Figure 5 shows that the average difference at the COICOP 5-digit level is +1,8 percentage points. Differences at this level range from -1,8 to +21,3 percentage points. At lower COICOP levels the difference is smaller but still significant: average annual inflation would have been 0,3 percentage points lower (7,7% instead of 8,0%) if the bilateral index had been implemented.

Figure 6 – Difference in annual inflation by COICOP level and by month (2023)



COICOP	1	2	3	4	5
			January 2023		
Avg. M	11,5	8,8	11,1	14,4	17,4
Avg. B	11,2	8,3	10,2	12,5	15,7
Δ(M-B)	+0,3	+0,5	+0,9	+1,9	+1,7
			December 2023		
Avg. M	5,4	5,4	5,9	4,9	5,6
Avg. B	5,2	4,9	5,0	2,9	3,9
Δ(M-B)	+0,2	+0,5	+0,9	+2,0	+1,7

Expressing the difference between the two methods in terms of the monthly value of annual inflation instead of the average annual inflation (see Figure 6), it can be seen that at the COICOP 1 level the difference does not vary significantly between January (+0,3) and December (+0,2), despite the moderation of high inflation over the 12-month period. The size of the average

difference also remains relatively stable at the lower COICOP levels (4 to 5), increasing from 1,9 to 2,0 percentage points and stagnated at 1,7 percentage points respectively. Annual inflation calculated using the multilateral method is consistently higher than that calculated using the bilateral method. The dispersion of the differences, shown in the monthly graphs, shows little change despite the steady decline in inflation. At the COICOP 5 level, differences range from -5,9 to 21,5 percentage points in January and from -1,9 to 21,5 percentage points in December.
It is important to note that annual inflation for COICOP 5 categories was 11,5% in January (using the multilateral method for COICOP categories covered by scanner data), while it fell to 5,4% by the end of the year. In other words, the difference between the two methods does not seem to be directly related to the rate of price increases.

Impact of Bilateral Index on CPI Food and on Food and Non-Alcoholic Beverages

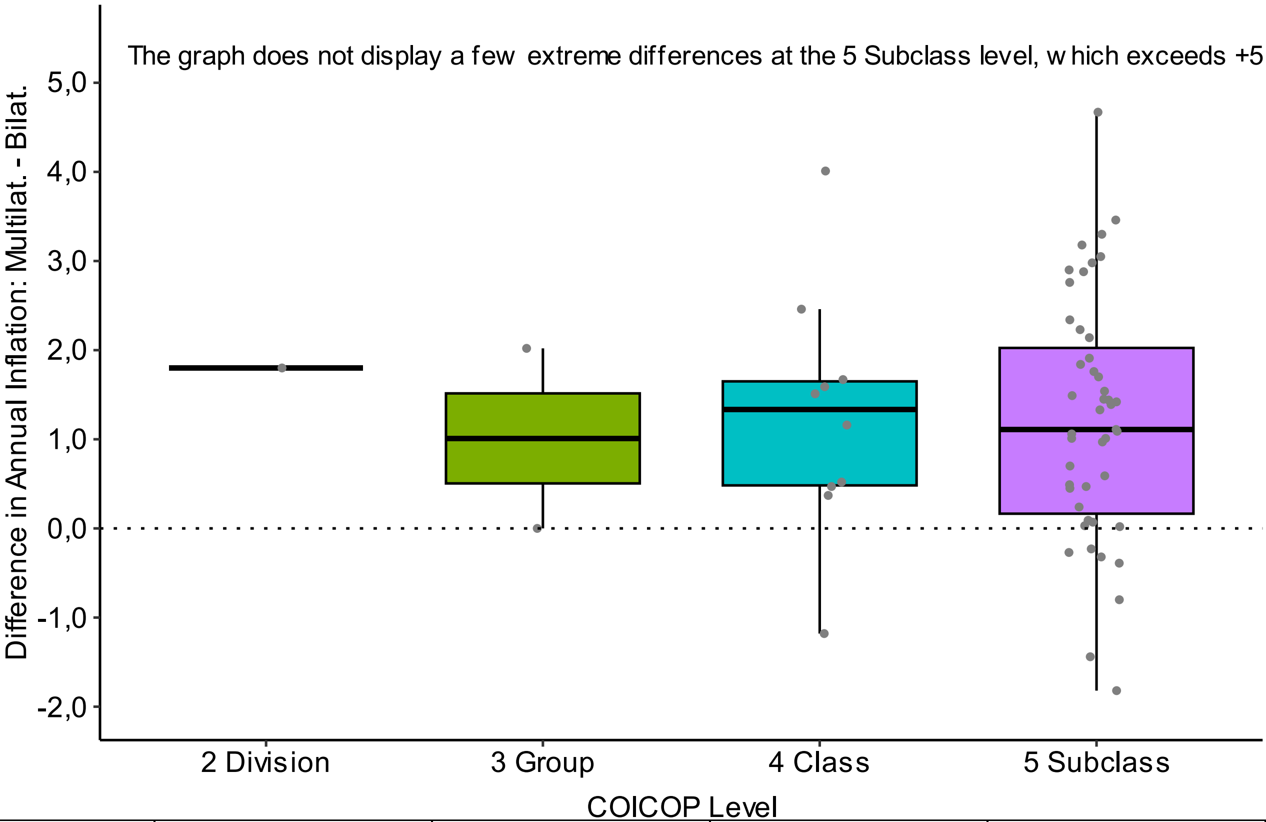
Although it is very important to see how bilateral method would have affected the overall index, it is nevertheless a logical step to limit our analysis to the COICOP categories that were fully covered by scanner data after the methodological change. Since the coverage of scanner data is complete in Division 01 (food and non-alcoholic beverages), we focus our analysis on this division.

Table 3 - Number of COICOP categories affected by scanner data at different COICOP levels

COICOP level	Number of categories	Average weight of the scanner data
2	1	100%
3	2	100%
4	11	100%
5	50	100%

In Table 3 we see that we have fewer categories in the analysis, but they are all fully covered with scanner data. In this case, it should be noted that the lowest level of examination is the division, so in the following figures and tables we will show four COICOP levels instead of the previous five.

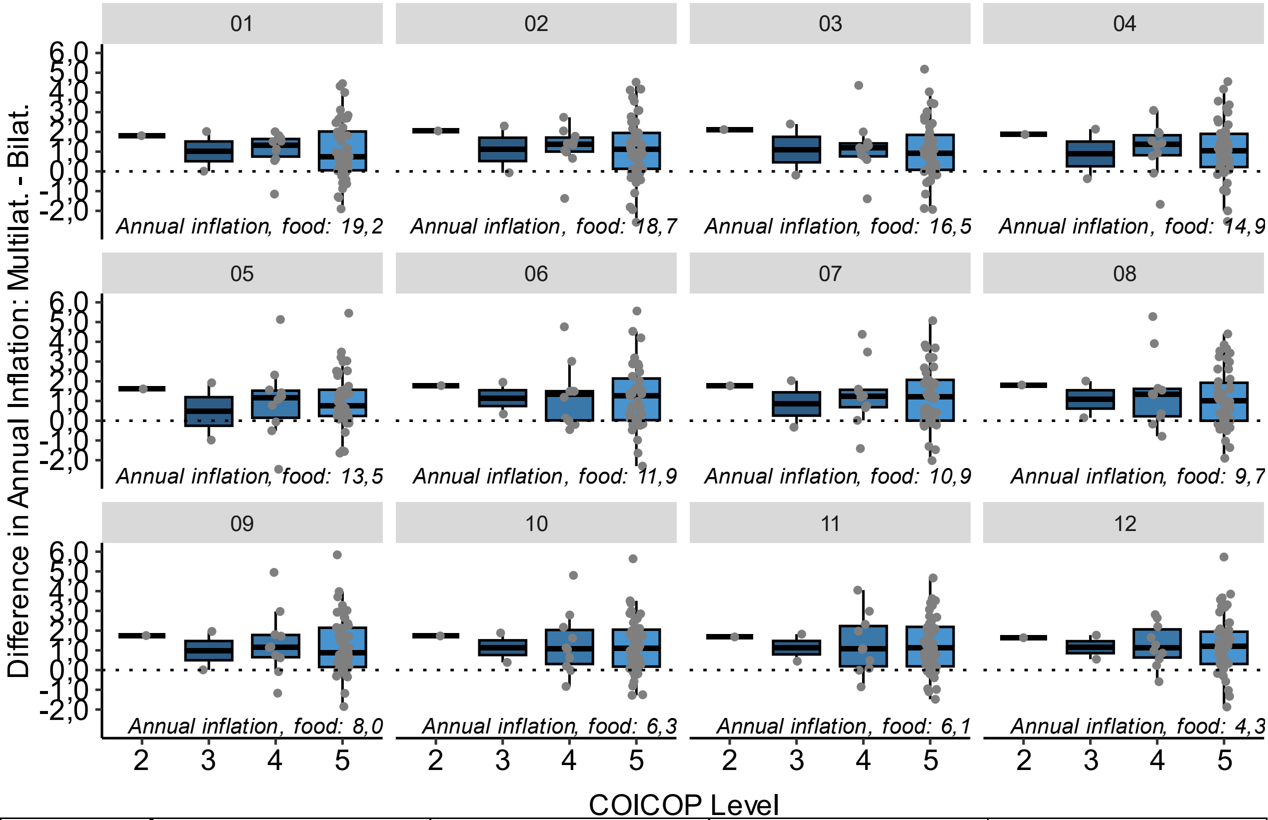
Figure 7 – Difference in average annual inflation by COICOP level: food only (2023)



Avg. M	11,7	12,2	11,4	13,1
Avg. B	9,9	11,2	9,3	11,2
Δ(M-B)	+1,8	+1,0	+1,9	+1,9

The average annual inflation in division 01 (food and non-alcoholic beverages) calculated with multilateral method is +1,8 percentage points higher than the inflation calculated with bilateral method. At COICOP 5-digit level, we again see relatively larger differences in the range -1,8 to +21,3 percentage points.

Figure 8 – Difference in annual inflation by COICOP level and by month – food only (2023)



COICOP	2	3	4	5
	January 2023			
Avg. M	19,2	17,7	18,7	19,4
Avg. B	17,4	16,7	16,7	17,6
Δ(M-B)	+1,8	+1,0	+2,0	+1,8
	December 2023			
Avg. M	4,3	6,2	4,5	5,6
Avg. B	2,7	5,1	2,5	3,9
Δ(M-B)	+1,6	+1,3	+20	+1,7

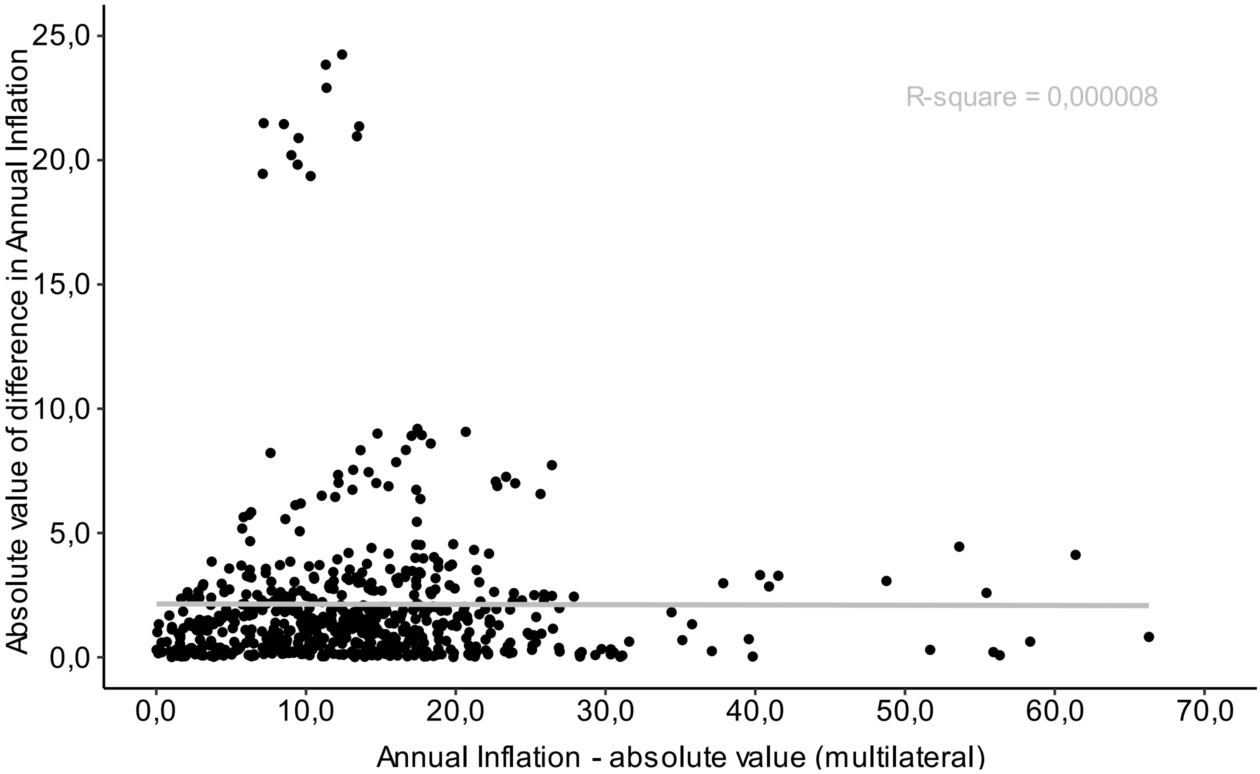
The monthly annual inflation values obtained by both methods reflect a consistent pattern observed in the average annual inflation: positive average differences persist across all COICOP levels and months. Despite a significant decrease in annual food inflation from January to December, the size of the difference remains relatively stable over the period considered. For example, in January the average difference between the two methods is 1,8 percentage points, while in December the difference decreases slightly to 1,6 percentage points at COICOP 2 level. An analysis of the data over two years suggests that there is no correlation between the inflation rate and the differences in the methodological results. Even when the annual inflation rate falls, the gap between the bilateral and multilateral method results remains relatively constant. This is particularly important because a difference of 1-2 percentage points for inflation above 10 percent may not carry the same weight as a similar difference for inflation below 10 percent.

The lack of correlation between the two indicators is discussed in more detail below

Analysing the Relationship Between Annual Inflation and Methodological Differences

In the chart below, each point represents a COICOP 5 category for 12 consecutive months (January to December 2023). The x-axis shows the extent of inflation for the respective month (calculated with multilateral method) for the respective COICOP 5 category, and the y-axis shows the differences between annual inflation measured with multilateral i.e bilateral approaches. We took the absolute value of both the x-axis and the y-axis, i.e., to remove the sign of both the price change and the difference, to make the potential relationship between the two variables somewhat more visible.

Figure 9 – The absolute difference according to the absolute value of annual inflation



The regression line and the R-squared value, which is close to 0, confirm that there is no relationship between the magnitude of price changes and the differences between the methods. This is further illustrated in the table below, which categorises price changes and assesses the differences accordingly. For example, for price changes between 0 and 5 percent, the average difference is 1,1 percentage points, rising to 2,7 percentage points for price changes between 5 and 10 percent. Beyond this range, however, there is no further increase in the difference: even when the price change exceeds 40 percent, the average difference remains at 2,1 percentage points.

Table 4 - The absolute difference by absolute value of annual inflation, split by categories

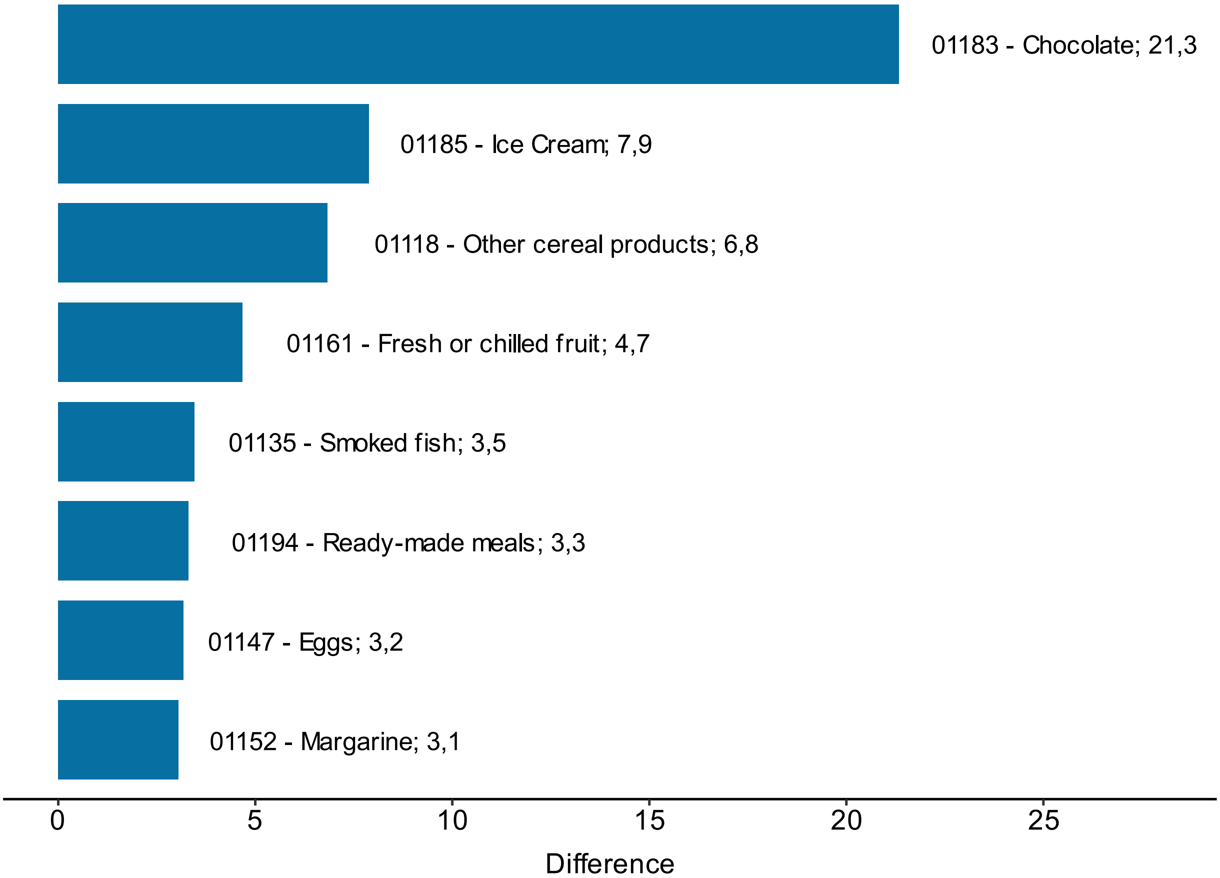
Annual inflation (absolute value of change)	Absolute value of difference
0-5	1,1
5-10	2,7
10-20	2,3
20-40	1,7
40+	2,1

Our data suggest that annual inflation measured by the bilateral method differs significantly from that measured by the multilateral method, regardless of the annual inflation rate. Given that the bilateral method consistently measures lower inflation than the multilateral method, this implies that even in a moderate inflation environment, the bilateral method can underestimate inflation by a relatively large margin.

Seasonality and Methodological Differences

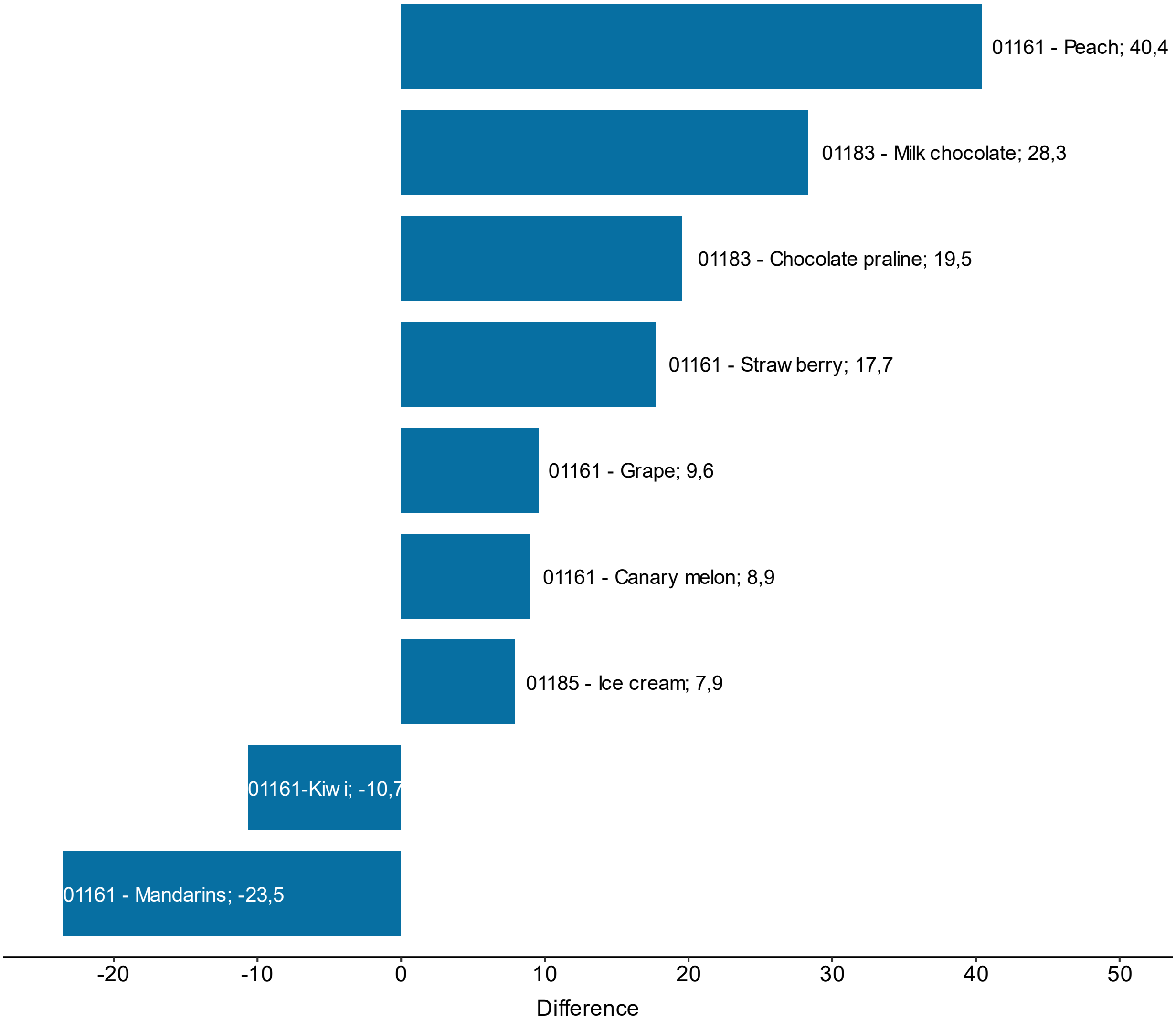
The largest positive difference in average annual inflation is observed for chocolate at COICOP 5 digit level, where the inflation rate calculated according to the multilateral method is 21,3 percentage points higher than that calculated according to the bilateral method. This is followed by the second largest difference for ice cream and the third largest for other cereals. The contrast is particularly pronounced in the case of chocolate, where the bilateral method indicates an average annual inflation decrease of 11% in 2023, while the multilateral method indicates a significant price increase of 10,3% (difference is 21,3%). This discrepancy is significant due to the seasonal nature of chocolate supply, which is often linked to holidays such as Christmas or Easter, resulting in products being available for sale for only 1-2 months of the year. As a result, the bilateral index does not accurately capture price fluctuations for these seasonal products.

Figure 10 – Top COICOP Subclasses: 5-Digit Level, 2023 Annual Average Inflation Difference (Multilateral vs. Bilateral)
Displayed are COICOP Subclasses with Differences Greater than 3 or Smaller than -3



The COICOP 5-digit level is not the elementary aggregate at which the index is calculated, so it is worth looking at the top differences at this lowest elementary level to see the negative and positive differences. The 50 COICOP 5 categories (see in Table 3) contain a total of 130 elementary aggregates.

Figure 11 – Top elementary aggregates, 2023 Annual Inflation Difference (Multilateral vs. Bilateral)
Displayed are* elementary aggregates *with Differences Greater than 7 or Smaller than -7



At the elementary aggregate level, six out of the nine top categories showing large differences between the methods presented relate to different types of fruit. Interestingly, in some rare cases the bilateral method shows the ability to measure higher inflation rates for seasonal products compared to the multilateral method. The latter scenario can occur when there is a consistent trend of increasing product prices for two consecutive months, followed by either stability or a decrease compared to the previous months. Using the multilateral method can reduce price increases observed in consecutive months, especially for highly seasonal products, resulting in more accurate price assessments. For example, in November 2023, we observed 354 mandarin unit values. Almost all products (353 out of 354) could be included in the multilateral index calculation, as only one product was not sold in at least one other month within the previous 13 months (the window length of the multilateral method). In contrast, only 185 products are included in the bilateral method, reflecting those available in the previous month. Consequently, on a month-on-month basis, the bilateral index shows a price increase of 0,6%, while the multilateral index shows a price decrease of -9,2%. It's worth noting that the bilateral index can be seen as a form of sampling based on product availability in the previous month. As seen earlier with the Jevons bilateral index, such sampling methods tend to bias the index upwards.

In addition to fruit, there are other products such as chocolate and ice cream that are seasonal and for which the multilateral method measures significantly higher inflation than the bilateral method. This observation is in line with the existing literature, which emphasises the increased importance of multilateral index calculation, especially for seasonal products. To explore this phenomenon

further, quantifiable indicators of seasonality have been developed to provide a clearer understanding of the correlation between seasonality and the deviation in annual inflation calculated by both bilateral and multilateral methods.
If we can express seasonality in terms of some quantifiable indicator, we can get a more accurate picture of the strength of the relationship between seasonality and the deviation of annual inflation calculated using bilateral or multilateral methods. Two indicators have been defined to express seasonality. One is based on the volatility of revenue data per elementary aggregate over the 13-month window. This was defined using the standard deviation of turnover. As each elementary aggregate generates different amounts of revenue, one of the indicators finally chosen was the coefficient of variation (CV), also known as the relative standard deviation (RSD), defined as the ratio of the standard deviation to the mean. The other indicator, which measures seasonality, measures the average number of months per elementary aggregate that products are available over the period defined by the 13-month window. For seasonal products, this value is lower because the products are not available out of season or are substituted by alternative products (e.g. imported products for fruit).
The strength of the relationship between these variables was measured using Pearson's correlation. In addition to seasonality, we also included in our analysis the magnitude of annual inflation, which we have already seen is not related to the magnitude of the difference between the methods. Our aim is to understand the strength of the relationship between seasonality and the difference between methods. We express both the difference and annual inflation in absolute terms, as before at Figure 9.

Table 5 - Pearson's correlation matrix at elementary aggregate level

	Difference (abs)	Revenue (RSD)	Number of months on sale	Annual Inflation (abs)
Difference (abs)	1,00	**0,78** <,0001	**-0,59** <,0001	-0,18 0,05
Revenue relative standard deviation (RSD)	**0,78** <,0001	1,00	**-0,58** <,0001	-0,24 0,006
Number of months on sale	**-0,59** <,0001	**-0,58** <,0001	1,00	0,31 0,0003
Annual Inflation(abs)	-0,18 0,05	-0,24 0,006	0,31 0,0003	1,00

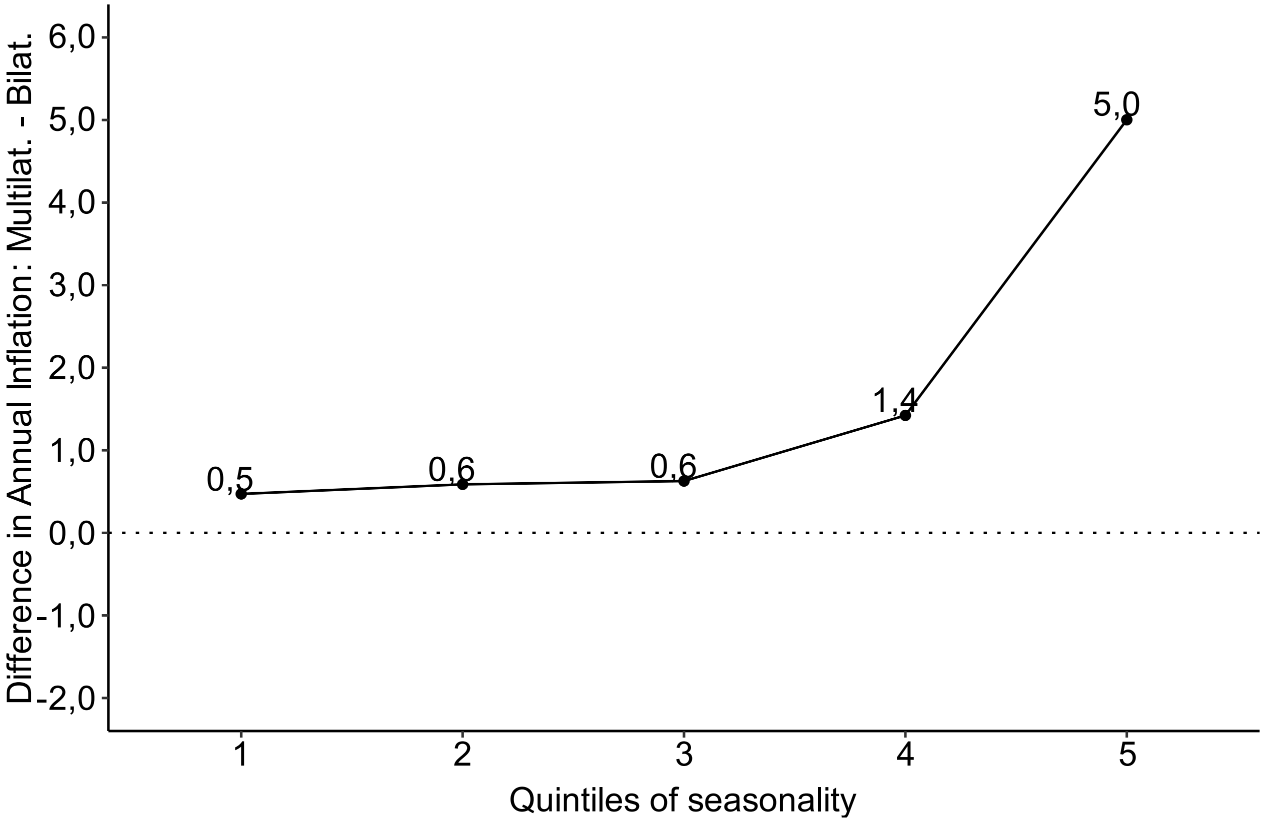
The correlation matrix shows the pairwise correlations between each variable.
There is a strong positive linear relationship between the difference in average annual inflation rates calculated with a multilateral and bilateral method and the relative standard deviation of the revenues of each elementary aggregate. This means that the higher the monthly volatility of revenues, the larger the difference between the two methods. There is also a significant linear relationship between our other indicator of seasonality and the absolute difference, but the direction is negative and the relationship is less strong. The negative direction is consistent with our expectations, since the fewer months on average a product is on sale, the more we can assume the seasonal character of the elementary aggregate, which is associated with a larger absolute difference. Consistent with the above, our two seasonal indicators are also strongly negatively correlated.
There is no significant relationship between the magnitude of the annual price change, currently defined as the absolute value of annual inflation measured by the multilateral method, and the magnitude of the difference between the two methods. This is consistent with Figure 8, which showed that in the first half of the year, when annual inflation was typically higher, we saw approximately same differences between the methods than in the second half of the year when inflation was lower.

To determine whether the difference is positive or negative, we used an additional seasonality indicator formed from our two previous seasonal variables. This indicator takes into account both the relative standard deviation of revenues and the number of months in which products are on sale.

$$\text{Saisonality} = \frac{\sigma(\text{revenue})}{\mu(\text{revenue})} \text{ X } \left(1 - \frac{\mu(\text{number of months on sale})}{13}\right)$$

We divided the 130 elementary aggregates into 5 quintiles along this new seasonality variable and evaluated the differences between the methods. To identify the signs, this time we used the original differences rather than the absolute values.

Figure 12 – Seasonality and difference in average annual inflation on elementary aggregate level for food, 2023



The top 20 percent of elementary aggregates (quintile 5), identified by our seasonality indicator as the most likely to show seasonal patterns, appear to have a significantly larger positive difference compared to other less seasonal elementary aggregates. Quintile 4 also shows above-average seasonality, with a similarly above-average difference, while the least seasonal 60 percent of elementary aggregates show a smaller but consistently positive deviation in favour of the multilateral method. In particular, quintile 5 includes in addition to the items seen on the Figure 11, a wide range of items, including fruit, chocolate, ice cream, vegetables and fish. Conversely, the least seasonal group (quintile 1) consists of items such as bread, milk, cheese, canned goods, chips, soft drinks and bananas.

Overall, the multilateral method, although dependent on elementary aggregates, tends to register higher inflation rates, with the difference being particularly pronounced for seasonal products. Specifically, the average deviation for the multilateral method is +5,0 percentage points for quintile 5 and +1,4 percentage points for quintile 4. In contrast, the deviation for quintiles 1-3 is much smaller, ranging between 0,5 and 0,6 percentage points.

Conclusion

- The introduction of scanner data into consumer price statistics represents a significant qualitative improvement. After extensive preparations, Statistics Austria introduced scanner data into the Austrian CPI and HICP in January 2022. This paper specifically examines the decision-making process surrounding the choice of index calculation method, focusing on the choice between multilateral and bilateral methods. For practical reasons, Statistics Austria chose a window length of 13 months using the GEKS index methodology for the introduction of scanner data into the CPI. The primary objective of this study was to assess, two years after implementation, how the index path would have been different if a bilateral index had been used, thus providing valuable insights for other National Statistical Institutes (NSIs) in the early stages of scanner data implementation.
- We have carried out a thorough comparison of price indices using bilateral and multilateral methods to assess their impact on indices and annual inflation rates. Our main focus was on contrasting the 13-window GEKS Törnqvist index, which we introduced, with the bilateral Törnqvist index, as it is better suited to the challenges posed by scanner data than other bilateral methods. In addition, we briefly examined the bilateral Jevons method used during the lock-down months prior to the introduction of scanner data. Our analysis showed that the bilateral Jevons index is vulnerable to chain drift and additional biases. Scanner data often include products with minimal turnover, which disproportionately affect the unweighted Jevons indices. Attempts to mitigate this bias through revenue-based sampling have lacked precision, resulting in a transition to positive bias when sampled. This confirms that scanner data and the unweighted Jevons index are difficult to reconcile.
- In a comprehensive analysis, we calculated annual inflation rates for each month in 2023 using both the bilateral Törnqvist and GEKS methods and compared them at different COICOP levels, from the 1-digit level (total CPI) to the 5-digit level. The scanner data covered only 16% of the consumer basket at the COICOP 1-digit level, but achieved full coverage (100%) at the 5-digit level within the food division (COICOP division 01). In particular, significant differences between the two methods were also observed at the level of the total index. The implementation of the multilateral method resulted in an average annual inflation difference of 0,3 percentage points higher in 2023 compared to the implementation of the bilateral method. At the COICOP 5-digit level, the differences ranged from -1,8 to +21,3 percentage points, with an average difference of +1,8 percentage points. These results underline the significant impact of methodological choices on the calculation of inflation rates at different levels of aggregation.
- In 2023, as inflation showed a gradual decline from month to month, we examined whether this declining trend affected the differences between methods. Our analysis revealed no significant correlation between the magnitude of price changes and the differences between methods at the COICOP 5-digit level. These results suggest that annual inflation rates calculated using the bilateral method differ significantly from those derived using the multilateral method, irrespective of the overall inflation rate. In particular, our data show that the bilateral method consistently leads to lower inflation measures than the multilateral method. This underlines that even in a moderate inflationary environment, the bilateral method can underestimate inflation by a significant margin.
- We used two indicators to measure seasonality in relation to the deviation of average annual inflation calculated using the bilateral or multilateral method. These indicators were the relative standard deviation of sales and the average duration of product availability per elementary aggregate. Pearson's correlation coefficient was used to quantify the strength of the relationship between these variables. Our analysis of the differences between methods across all food elementary aggregates revealed a robust positive linear correlation between the absolute value of the differences in average annual inflation rates and the relative standard

deviation of sales. In addition, we observed a significant linear correlation between the other seasonality indicator and the absolute difference in average annual inflation rates, albeit with a negative direction and weaker strength of the relationship.

- We found a correlation between seasonality and the deviations of the indices obtained by two different methods: the multilateral GEKS and the bilateral Törnqvist, which gives indication of the direction of these differences. Elementary aggregates were divided into quintiles according to their seasonality. For the most seasonal products (quintile 5), the multilateral method consistently produced higher annual inflation rates. Overall, the multilateral method tended to give higher inflation rates, especially for seasonal products. For example, the average deviation for the multilateral method was +5,0 percentage points for quintile 5 and +1,4 percentage points for quintile 4. Conversely, the deviations for quintiles 1-3 were much smaller, ranging between 0,5 and 0,6 percentage points.
- In summary, the significant differences we have found between the two methods strongly argue against the use of the bilateral method with scanner data, as it consistently shows a tendency to underestimate inflation, especially for seasonal products.

References

Chessa, A.G. (2021) Extension of multilateral index series over time: Analysis and comparison of methods, Paper written for the 2021 Meeting of the Group of Experts on Consumer Price Indices

Fox, K. J., Levell, P., O'Connell, M. (2022) Multilateral index number methods for Consumer Price Statistics

Tardos, A. (2023) Introduction of Scanner Data into Austrian CPI and HICP – practical implementation experience, with a focus on window length options

Publications Office of the European Union (2022) Guide on Multilateral Methods in the Harmonised Index of Consumer Prices, 2022 edition, Luxembourg

Office for National Statistics (ONS), released 28 November 2022, ONS website, methodology article, GEKS Törnqvist: introducing multilateral index methods into consumer price statistics