17th Ottawa Group Meeting
Rome, 7 – 10 June 2022

The use of weighted GEKS for the calculation of consumer price indices: an experimental application to Italian scanner data

Alessandro Brunetti (Istat), Stefania Fatello (Istat), Tiziana Laureti (Università della Tuscia), Federico Polidoro (Istat)

Abstract

The paper focuses on a generalization of the standard GEKS method, based on Törnqvist indices, used for the compilation of transitive consumer price indices. Unlike the usual application of this method, in its weighted version different weights for different months in the reference period enter into the calculation of indices. In this study, different alternative systems of weights are considered for introducing information on the reliability of the underlying binaries for the weighting matrix. The use of different system of weights, based on the similarity of each couple of months being compared in terms of (a) the share of matching items and (b) the share of the corresponding turnover, proved to have moderate effect on the dynamic of GEKS. However, the evidence suggests that the weighted version of rolling windows GEKS, under different splicing options, tend to be slightly closer to the full window counterpart, as compared to standard GEKS. In other words, the use of weights seems to reduce the impact of the constraint of non-revising the indices. As a further line of research for exploring the different performance of weighted and unweighted GEKS, in the last part of the paper, we present the preliminary results of an analysis, aimed to compare standard and weighted GEKS, which is based on the calculation of a target "true" cost of living index under the hypothesis of Constant Elasticity of Substitution (CES) purchaser preferences. In this analytical setting, our findings based on real scanner data seem to confirm the weighted GEKS tend to perform better than its unweighted version.

1. Introduction

Since the introduction of scanner data as new source of data to estimate inflation, the methodological debate developed over the years about their use in the context of CPI compilation. The attention was mainly focused on the potentialities that derive from one of the main value added of the scanner data that is the availability of the information about turnover (in most cases at level of Global Trade Item Number, GTIN and of each outlet). Indeed, the availability of this information has opened new perspectives to the use of weighted indices at level of elementary aggregates instead of the traditional approach to CPI compilation, mainly based on unweighted formula, allowing the use of the results obtained by Balk in the 80s about the translation of multilateral methods to the time series context.

This debate meant the evolution from the static and dynamic approach to sampling (mainly based on the use of unweighted formula to calculate indices at elementary level) to the studies on the use of multilateral indices for scanner data, moving the methodological choices that were mainly implemented in the field of Purchasing Power Parities (PPPs) to the field of estimation of the temporal evolution of Consumer Prices.

The evolution of this debate was represented also in the evolution of the EU regulations regarding the Harmonised Indices of Consumer Prices. Indeed, the issues related to lack of information about turnover at the most disaggregated level (elementary aggregate level) of index compilation were embedded in the Regulation 1749/96 (on initial implementing measures for Council Regulation (EC) No 2494/95) that substantially established that the price indices for elementary aggregates had to be compiled by using Jevons or Dutot indices. The last implementing Regulation 1148/2020 (laying down the methodological and technical specifications in accordance with Regulation (EU) 2016/792 of the European Parliament and of the Council as regards harmonised indices of consumer prices and the house price index), established that "the prices of individual products shall be aggregated to obtain elementary price indices using either of the following options:

(a) an index formula that ensures transitivity. The price index of prior periods shall not be revised when using transitive index formulae; or

(b) an index formula that ensures time reversibility and compares the prices of individual products in the current period with the prices of those products in the base period. The base period shall not be changed frequently if such change leads to significant violation of the transitivity principle."

This change implies that for the compilation of HICP, still Dutot and Carli formula can be used to compile indices of elementary aggregates but also multilateral indices that ensure transitivity and this represents an important innovation that was pushed by the increasing use of scanner data for CPI compilation.

One of the crucial issue that emerge from the adoption of the multilateral method is given by the constraint of not revising the indices referred to a prior period and the related issue of the window length to impose transitivity.

This paper represents a contribution on how to solve this issue taking advantage from the availability in the most used scanner database of the information about turnover and quantities (that means about weights). Indeed, what emerges in the experimentations analysed in this paper is that the use of weights in the compilation of specific types of multilateral indices, seems to reduce the impact of the constraint of non-revising the indices themselves.

2. Multilateral index methods based on bilateral comparisons

Scanner data offer new opportunities for price index calculation compared to traditional price collection, especially regarding the use of weighted index formulae based on product weights within elementary aggregates.

However, several empirical studies showed that high-frequency chaining of weighted price indices, including superlative price indices, can lead to strong chain drift[1] (Silver and Heravi, 2005; De Haan, and Van der Grient, 2011; De Haan and Krsinich, 2014).

In order to deal with the chain drift problem Ivancic, Diewert and Fox (2009; 2011) proposed the use of a multilateral indices in the scanner data context following the work of Balk (1981) and Kokoski, Moulton and Zieschang (1999) who noted that multilateral index methods, originally developed for price comparisons across countries, can be easily adapted to price comparisons across time. However, the basic idea of adapting a multilateral method to the time series context is due to Balk (1981) who set up a framework that is very similar to the one suggested by Ivancic, Diewert and Fox (2011). Balk (1981) used an index number formula due to Vartia (1976), in place of maximum overlap bilateral Fisher indices, as his basic building blocks.

[1] As stated in the CPI Manual (2020) "A chain index is said to drift if it does not return to unity when prices in the current period return to their levels in the base period."

Multilateral methods produce transitive price indices. For price comparisons across time, this means that the indices are independent of the choice of base period and can be written in chained form. Therefore, they are free from chain drift. Various multilateral methods have been suggested in literature following different approaches with the common characteristic that price indices are constructed simultaneously for the entire sample period (Diewert and Fox, 2022).

2.1 The Gini–Eltetö–Köves–Szulc (GEKS) method

The GEKS method, proposed by Gini (1931), Elteto and Koves (1964) and Szulc (1964), is designed to construct transitive multilateral comparisons from a matrix of binary/pairwise comparisons derived using a formula which does not satisfy the transitivity property. Using the GEKS approach it is possible to obtain a transitive index that deviates the least from a given matrix of binary comparisons between M entities. If non-transitive binary indices are indicated by I_{jk} $(j, k = 1,2, \ldots, M)$ then the GEKS transitive indices are obtained by minimizing:

$$\sum_{j=1}^{T} \sum_{k=1}^{T} \left[ln I_{jk}^{GEKS} - ln I_{jk}\right]^2 \qquad [1]$$

subject to $I_{jk}^{GEKS} = I_{jl}^{GEKS} \cdot I_{lk}^{GEKS}\ for\ all\ j, k$.

The computational form for the GEKS index is given by

$$I_{jk}^{GEKS} = \prod_{j=1}^{M} \left[I_{jl} \cdot I_{lk}\right]^{1/M} \qquad [2]$$

Equation [2] defines the GEKS index as an unweighted geometric average of the linked (or chained) comparisons between entities j and k using each of the entities in the comparisons as a link.

Ivancic, Diewert and Fox (2011) suggested to use the GEKS method to price indices across time by treating each time period as an entity. Let us consider to have data on prices and quantities at our disposal for periods 0, 1. . . T. Choosing 0 as the index reference period and denoting the comparison periods by t (t = 1, . . . , T), we can write the GEKS price index between 0 and t as an unweighted (or equally weighted) geometric average of all possible price comparisons where each link period l across the sample period serves as the base:

$$I_{0t}^{GEKS} = \prod_{l=1}^{T} \left[I_{0l} / I_{lt}\right]^{1/T+1} = \prod_{l=1}^{T} \left[I_{0l} \cdot I_{lt}\right]^{1/T+1} \qquad [3]$$

provided that the bilateral indices satisfy the time reversal test. In that case the GEKS index also satisfies this test, i.e. $I_{t0}^{GEKS} = 1/I_{0t}^{GEKS}$. The transitivity property implies that the GEKS index can be written as a period-to-period chained index

$$I_{0t}^{GEKS} = \prod_{s=1}^{t} \left[I_{t-1,t}^{GEKS}\right]^{1/T+1}$$

which should be free of chain drift (Ivancic, Diewert and Fox, 2011).

The bilateral indices are all matched-item indices, that is only price relatives of items that are purchased in the two periods compared enter the indices. The GEKS approach thus makes maximum use of all possible matches in the data between any two periods, which can be seen as its most important property. In addition, it is worth noting that in the GEKS indices all possible base months contribute to the overall index values.

The GEKS method in its original form uses the binary Fisher indices as the starting point. Let p_i^0 and s_i^0 denote the price and expenditure share of good (or item) *i* in the base period 0; p_{it} and s_{it} denote the corresponding values in the comparison period t (t>0). For a matched set of goods U the Fisher I_{0t}^F is defined as follows:

$$I_{0t}^F = \left[\frac{\sum_{i\in U} s_{i0}\left(p_{it}/p_{i0}\right)}{\sum_{i\in U} s_{it}\left(p_{it}/p_{i0}\right)^{-1}}\right]^{1/2} \qquad [4]$$

Feenstra et al. (2009) and de Haan and van der Grient (2011) suggested that the Törnqvist price index formula I_{0t}^T could be used instead of the Fisher price index in the Gini methodology:

$$I_{0t}^T = \prod_{i\in U}\left(p_{it}/p_{it}\right)^{\left(s_{i0}+s_{it}\right)/2} \qquad [5]$$

Both indices in formula [4] and [5] are superlative, as defined by Diewert (1976)[2].

Caves et al. (1982) used the GEKS idea with the Törnqvist index as a base in the context of making quantity comparisons across production units (the CCD method). Consequently, in the article by Diewert and Fox (2022), the multilateral price comparison method involving the GEKS method based on the Törnqvist price index is called the CCDI method.

2.2 The Weighted Gini–Eltetö–Köves–Szulc (WGEKS) method

When transferring multilateral methods from the spatial domain, where the set of countries or regions is fixed, to the time domain, countries are replaced by periods (e.g. months) and the set of periods considered is not fixed but changes when data of the next period become available. In order to avoid the revision of the already published indices, Ivancic, Diewert and Fox (2011) developed the rolling year GEKS method or RYGEKS. The RYGEKS method is an extension of the widely used GEKS approach for imposing transitivity on bilateral indices of a moving time interval (or window) of fixed length (13 months in their original work) which are then linked to each other in order to calculate the non-revisable index. However, the length of the windows is an issue that has been widely discussed in the literature. In fact there are trade-offs to consider when deciding the size of "time window" over which the multilateral index is applied. More specifically, shorter time windows could lead to unstable results and may not solve the chain-drift problem. However, the longer the time window, the more data from the past will impact the current-month compilations. Krsinich (2014) suggests a rolling 13-month window, which is used to calculate chained year-on-year indices each month. However, this method has given quite volatile and biased results (Chessa, 2015).

The choice may also depend on the product type; for seasonal products, the window should be sufficiently long to cover two successive in-season periods while minimizing the loss of "characteristicity". In practice, the time window should cover at least 13–14 months, if not longer such 25 months, as recently suggested by Eurostat (2022).

A loss of characteristicity means that price changes in the distant past disproportionately affect the results. Characteristicity, defined in the context of spatial comparisons (involving many countries and multilateral methods) following the seminal paper by Drechsler (1973), requires that any set of

[2] A superlative index number formula has the property that it is exactly equal to a Konüs (1924) true cost of living index (COLI) provided that the purchasing households have preferences that can be represented by certain functional forms, where these functional forms can approximate arbitrary preferences to the accuracy of a second order approximation.

multilateral comparisons satisfying the transitivity property should retain the essential features of the binary comparisons constructed without the transitivity requirement. In other words, the characteristicity property requires that distortions resulting from adherence to the transitivity property should be kept at a minimum. The GEKS method is especially defined with the characteristicity in mind since the GEKS technique requires to minimize such loss of characteristicity as specified in equation [1].

Similarly, in the context of temporal comparison using multilateral indices, as time passes, recent price movements will be increasingly affected by prices and price changes in the distant past. This will result in a loss of characteristicity. In other words, the choice of window length involves the tension, on the one hand, between using as much of the data as possible and, on the other, of using only bilateral comparisons which are reliable.

An approach for solving this tension between the different levels of reliability of the bilateral indices has been suggested by Rao (1997). It can be argued that in practice it is possible to show that some link comparisons are intrinsically more reliable than others. For example, in practice it is possible to find that some pairwise Fisher or Törnqvist indices are based on price data for many commodities while in other cases comparisons are based on prices for only a few items.

Melser (2016) argued that rather than finding a certain optimal window length a natural solution is to instead use weighted GEKS (WGEKS). This, through suitable choice of weights, allows for the fact that comparisons tend to become less reliable as the periods being compared become further apart.

Following Rao (2001) in order to generalize the GEKS method to incorporate weights to various linked comparisons involved in equation [3], it is necessary to look at the GEKS method from a different angle as illustrated in [1]. Even though this optimisation problem appears to be difficult to solve, it can be handled with considerable ease once the problem is reparametrised using the following commonly known simple result: A multilateral system of index numbers, I_{rt} (r,t=1,2,…,T), satisfy the transitivity property if and only there exist T numbers $\Pi_1, \Pi_2, ..., \Pi_T$ such that, for all r and t $lnI_{rt} = \Pi_r - \Pi_t$.

Using this result on transitive index numbers, the optimization problem can be restated as one finding $\Pi_1, \Pi_2, ..., \Pi_T$ which minimizes:

$$\sum_{r=1}^{T}\sum_{t=1}^{T}\left[\Pi_r - \Pi_t - lnI_{rt}\right]^2 \qquad [6]$$

Then the required index can be shown to be equal to the ratio $\exp(\widehat{\Pi_r})/\exp(\widehat{\Pi_t})$ in which $\widehat{\Pi_r}$ and $\widehat{\Pi_t}$ are solutions to the minimization problem (Rao, 2001; Rao and Timmer, 2003). After some algebraic manipulation, the GEKS can be expressed as follows:

$$GEKS_{rt} = \frac{\exp(\widehat{\Pi_r})}{\exp(\widehat{\Pi_t})} = exp\left(\widehat{\Pi_r} - \widehat{\Pi_t}\right) \qquad [7]$$

It can be demonstrated that $\widehat{\Pi}$ are the ordinary least squares estimators of Π in the following model specification (Rao, 2001):

$$lnI_{rt} = \Pi_r - \Pi_t + u_{rt} \qquad [8]$$

with $E(u_{rt}) = 0$ and $Var(u_{rt}) = \sigma^2$.

Given the model specification in [8], it is possible to specify the bilateral index formula I_{rt}, i.e. Fisher or Törnqvist, and to discriminate between different pairwise comparisons using some indicators of reliability as weights w_{rt}. The larger weights w_{rt} the larger the reliability of the indices I_{rt} .

The following model with Törnqvist indices can be obtained:

$$lnI_{rt}^{T} = \Pi_r - \Pi_t + u_{rt} \qquad [9]$$

with $E(u_{rt}) = 0$ and $Var(u_{rt}) = \frac{\sigma^2}{w_{rt}}$

The Törnqvist WGEKS can be obtained from the following model that is characterized by a general structure underlying the process of according weights to different linked comparisons:

$$\sqrt{w_{rt}}lnI_{rt}^{T} = \sqrt{w_{rt}}\Pi_r - \sqrt{w_{rt}}\Pi_t + e_{rt} \qquad [10]$$

with $E(e_{rt}) = 0$ and $Var(e_{rt}) = \sigma^2 \;\; \forall r, t = 1, \dots, T, r \neq t$.

To make the method operational it is necessary to specify the matrix weights. In section 3 various sets of weights will be considered.

2.3 Index extension methods

Following equation [10] a sequence of index series is generated for successive time windows, which must be linked for compiling the next index to be published. Methods that perform this linking are also known as "index extension methods"[3].

The most widely used index extension methods can be subdivided into two main classes: *Splicing methods*, which all use a rolling window and a moving linking month, and *Fixed base methods* using a fixed base month as linking month, which is usually December of the previous year. In this case, the published index in the base month is the linking index. A first version of this type of method are the fixed base expanding window proposed by Chessa (2016) where a direct index is calculated with respect to the base month. The index in December of the present year will be equal to the index of the transitive series when linking this series to the base month, for a 13-month window. This implies that the series of published indices will be free of chain drift by construction, at least, in a piecewise sense.

Concerning the splicing methods, a well-known option is the "*window splice"* proposed by Krsinich (2014) in which the first month of the adjusted window is used as linking month. Due to this property, the method is sometimes also called "*full window splice*". The most recent recalculated index is used as linking index, that is, of the most recently linked series. De Haan and van der Grient (2011) as part of the Rolling Year GEKS method have proposed another method, called "*movement splice*". In this case, the penultimate month of the adjusted window is taken as linking month, and the month on month index of the adjusted window is chained to the published index of the previous month. De Haan (2015) also suggested that the link period t should be chosen to be in the middle of the first window time span; i.e., choose t = T/2 if T is an even integer or t = (T+1)/2 if T is an odd integer.

[3] Several different index extension methods have been proposed in literature over the past years. An interesting classification is provided by Chessa (2019) who underlined that once the length of the time window is chosen, three choices are made that characterise index extension methods: i) The adjustment of the time window from month to month; ii) The linking month; iii) The index in the linking month. Different choices can be made for each of these three aspects, which produce different index extension methods.

The Australian Bureau of Statistics (2016) called this the "*half splice method"* for linking the results of the two windows. Diewert and Fox (2022) suggested the "*mean splice method*", defined as the geometric mean of the estimators, as the "best" estimator for the period T+1 price level.

We will study these alternative linking methods in the context of the GEKS multilateral indices since this facilitates comparisons.

3. Alternative weighting schemes for the WGEKS

The general idea in weighted GEKS is to use information on the reliability of the underlying binaries for the weighting matrix. Given the nature of the generalizations involved, it is possible to arrive at a number of alternative specifications of the matrix of weights based on how one may wish to measure reliability. The weights in WGEKS have been constructed in a range of ways in the literature (Rao and Timmer, 2003; Diewert, 2005; Hill and Timmer, 2006).

Rao and Timmer (2003) underlined that a first group of statistical measures of reliability can be defined from a sampling perspective. In this context binary comparisons based on a small number of matched products priced in both countries or periods would be less reliable. Similarly, if the products matched and used in a binary comparison cover only a small proportion of the total output size of the manufacturing sector (or a branch) in the two countries or periods involved then the products considered may not be representative of the whole sector and hence any comparison based on price data for these products would be less reliable.

In this paper we considered two measures used by Melser (2018) and Melser and Webster (2021) which are consistent with this literature. In the first measure, called the average matched expenditure share (AMES) method, the weights are calculated as:

$$w_{rt}^{AMES} = \frac{1}{2}\left[\left(\frac{\sum_{i \in U_{r,t}} p_{i,t} \cdot q_{i,t}}{\sum_{i \in U_t} p_{i,t} \cdot q_{i,t}}\right) + \frac{\sum_{i \in U_{t,r}} p_{i,r} \cdot q_{i,r}}{\sum_{i \in U_r} p_{i,r} \cdot q_{i,r}}\right] \qquad [11]$$

Where *i* referes to products while prices and quantities are denoted p_{it} and q_{it} respectively and $U_{r,t}$ = Ur ∩ Ut is the index set of products available in both periods.

This choice of weights results in periods with large matched expenditure shares receiving higher weights. A key advantage is that this approach closely concords with the weighting structure in the Törnqvist index number formula. Note also that it treats each period's expenditure share symmetrically even if total expenditures are quite different. The matrix considered for weighting purposes is also called the matrix of coverage ratios, following Rao, Selvanathan and Pilat (1995).

A second approach focuses on average matched product shares (AMPS) rather than considering expenditures. According to this method, denotes as AMPS, weights are calculated as follows:

$$w_{rt}^{AMPS} = \frac{1}{2}\left(\frac{N_{r,t}}{N_t} + \frac{N_{t,r}}{N_r}\right) \qquad [12]$$

This is quite similar to AMES except it gives each product equal weight rather than basing it on the expenditures' shares. Melser (2018) found that AMPS declines faster than AMES as the distance between r and t rise because it tends to be the lower-expenditure products which disappear from the market.

Another approach, called the average matched expenditure (AME), has been considered by Melser (2018) where larger weight is given to cases where expenditures are matched and also where expenditures are large. This last feature of AME is both potentially an advantage and a disadvantage. When there are more purchases the individual prices calculated from the data will be more reliable and hence should receive higher weight. However, in a highly inflationary environment AME weights can become distorted and give too much emphasis to periods with high prices unlike AMES weights.

4. The case study

Since 2018, ISTAT has been using scanner data of grocery products (excluding fresh food) to compile CPIs[4]. The sample of outlets is selected through a probabilistic design, in which outlets are stratified according to all Italian provinces (107) and retail trade channels (hypermarket, supermarket, outlets with surface between 100 and 400 s.m., discounts and specialist drugs). Probabilities of selection are assigned to each outlet, which are based on the corresponding turnover value[5]. GTINS are selected on a monthly base, according to the dynamic approach[6]. For each GTIN in each outlet, monthly prices are calculated as the arithmetic mean of weekly prices weighted with quantities[7].

Considering the sample of outlets in the province of Rome, experimental multilateral indices are compiled using data from December 2018 to February 2022. To calculate multilateral indices, turnover and quantities sold in the outlets belonging to the same retail trade channel were aggregated to obtain the average price per GTIN. Table 1 shows the number of outlets sampled in the years considered for analyzes.

Tab. 1. Sample of outlets by retail trade channels. Province of Rome. Years 2019-2022.

Number of outlets sampled in the province of Rome by retail trade channels

Retail trade channels	2019	2020	2021	2022
Hypermarkets (HYP)	13	13	8	9
Supermarkets (SUP)	48	48	56	57
Discounts (DIS)	19	19	20	22
Small sales areas* (SSA)	25	25	25	26
Specialist drug (SD)	14	14	16	16
Total sample	**119**	**119**	**125**	**130**

** outlets with surface between 100 and 400 s.m.*

For our experimental use of $RWGEKS_{25}$ $RWGEKS_{full}$, we considered five different product aggregates (three for food sector and the remaining two for the non-food sector): *Chocolate*, *Packaged ice cream*, *Olive oil*, *Body hygiene products* and *Cosmetic products*.

Concerning *Chocolate* and *Packaged ice cream*, they exhibit some seasonality in sales. The first one is characterized by a much higher turnover in winter months; conversely, the second one has a very high turnover in summer months. As for the two non-food product aggregates, the main feature

[4]At present, scanner data feed the calculation of 84 sub-indices ("Product Aggregates" representing 10dgts of the National classification) belonging to six ECOICOP Divisions (01, 02, 05, 06, 09, 12). In 2022, scanner data for 4,007 outlets of the main 21 Retail Trade Chains (RTC) covering the entire national territory are monthly collected by ISTAT on a weekly basis at item code level.

[5] The sample of outlets is renewed every year.

[6] ISTAT receives detailed information concerning turnover and quantity at weekly frequency, GTIN by GTIN, outlet by outlet. Preliminary formal checks are implemented on weekly data in order to identify macroscopic errors at the beginning of data flow and remove them. Among formal checks, inadmissible prices detection is carried out to alert when a price for a product is too high or too low based on the prices recorded for the same items in other outlets of the same province (moving trimming on price levels).

[7] The first three full weeks of the month are always used.

is the presence of many distinct groups of items[8] with a large number of very similar GTINs whose sales vary significantly from period to period. This leads to a high number of missing values. The following Table 2 shows the number of GTINs sold by retail trade channel in the years considered for each product aggregate.

Tab. 2. Number of GTINs by retail trade channel (province of Rome)

Number of GTINs by retail trade channels (province of Rome)

FOOD

Chocolate

Retail trade channels	2019	2020	2021	2022*
Hypermarkets	1.972	1.938	1.535	1.220
Supermarkets	1.748	1.799	1.711	1.394
Discounts	577	554	539	727
Small sales areas	877	901	895	753
Total	**2.680**	**2.683**	**2.286**	**1.873**

** Jan - Feb*

Packaged ice cream

Retail trade channels	2019	2020	2021	2022*
Hypermarkets	1.017	1.026	862	641
Supermarkets	969	1.023	1.066	793
Discounts	313	311	354	348
Small sales areas	603	642	674	419
Total	**1.457**	**1.481**	**1.449**	**1.095**

** Jan - Feb*

Olive oil

Retail trade channels	2019	2020	2021	2022*
Hypermarkets	580	575	393	338
Supermarkets	501	486	418	365
Discounts	51	57	54	107
Small sales areas	260	260	244	205
Total	**737**	**743**	**545**	**484**

** Jan - Feb*

NON FOOD

Body hygiene products

Retail trade channels	2019	2020	2021	2022*
Hypermarkets	2.618	2.367	1.708	1.476
Supermarkets	1.842	1.826	1.827	1.409
Discounts	399	400	387	467
Small sales areas	673	666	638	498
Specialist drug	1.541	1.590	1.521	1.225
Total	**3.816**	**3.721**	**3.247**	**2.630**

** Jan - Feb*

Cosmetic products

Retail trade channels	2019	2020	2021	2022*
Hypermarkets	3.278	2.927	766	505
Supermarkets	896	800	844	511
Discounts	114	102	89	63
Small sales areas	61	53	49	29
Specialist drug	3.861	3.696	3.936	2.863
Total	**5.752**	**5.356**	**5.016**	**3.583**

** Jan - Feb*

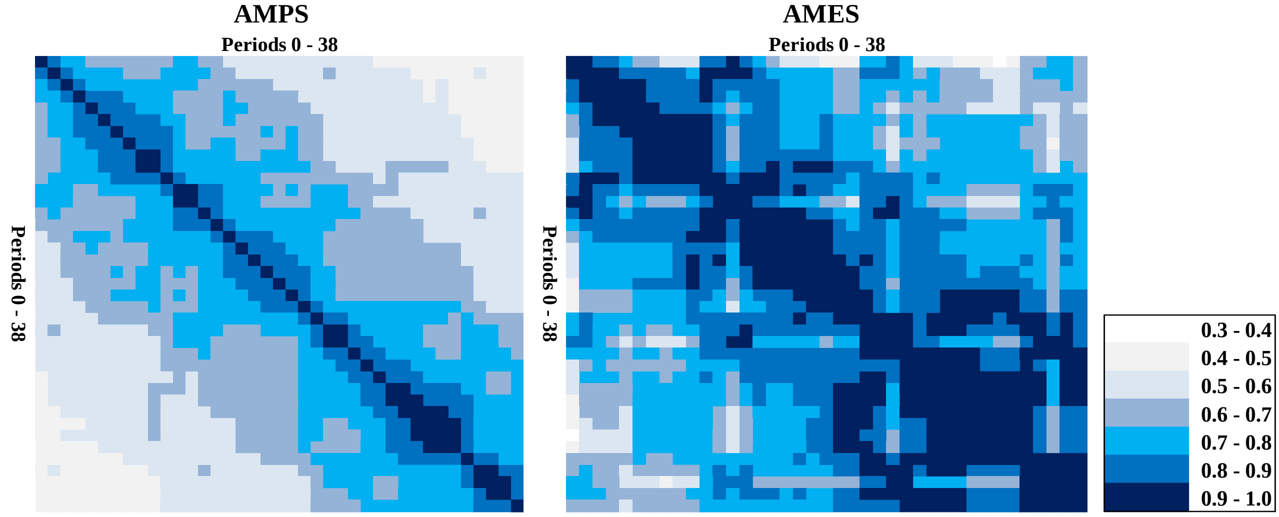
To calculate RYWGEKS we built weight matrices following the two approaches described in the previous paragraph 3: AMES and AMPS methods. The weight matrices appear quite different by applying the two methods for all products considered in the various retail trade channels. In general, as the distance between two periods increases, AMPS tends to decrease much faster than AMES[9]. Consequently, the AMES usually remains on relatively higher levels than the AMPS. To provide few examples of this evidence, Figure 1 shows AMPS and AMES matrices calculated for the whole window (periods 0-38) for the following product aggregates: *Chocolate* sold in the supermarkets (a), *Package ice cream* sold in the hypermarkets (b), *Olive oil* sold in the discounts (c), *Cosmetics products* sold in the hypermarkets (d) and *Body hygiene products* sold in the specialist drug (e).

For products affected by the presence of seasonality, such as *Package ice cream*, the decline of the AMPS is less regular and this is more evident when the AMES is considered. Both matrices decrease much faster for products with a high number of missing values, such as *Cosmetic products*. Comparing the retail trade channels, the shape of the AMPS appears to be quite similar; for the AMES there are differences especially for the discounts.

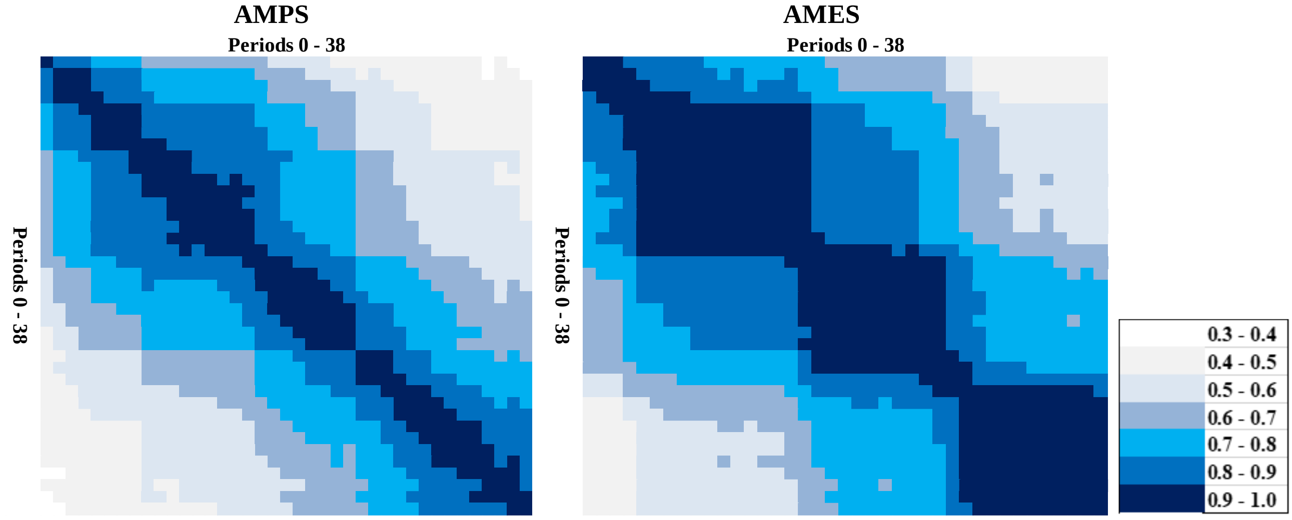
[8] For example, the aggregate “Cosmetic products” includes 22 groups, among which, Lips gloss, Lipsticks, Eye liner, Nail polish remover).
[9] See also Melser (2018).

Fig. 1. AMPS and AMES matrices (period 0-38)

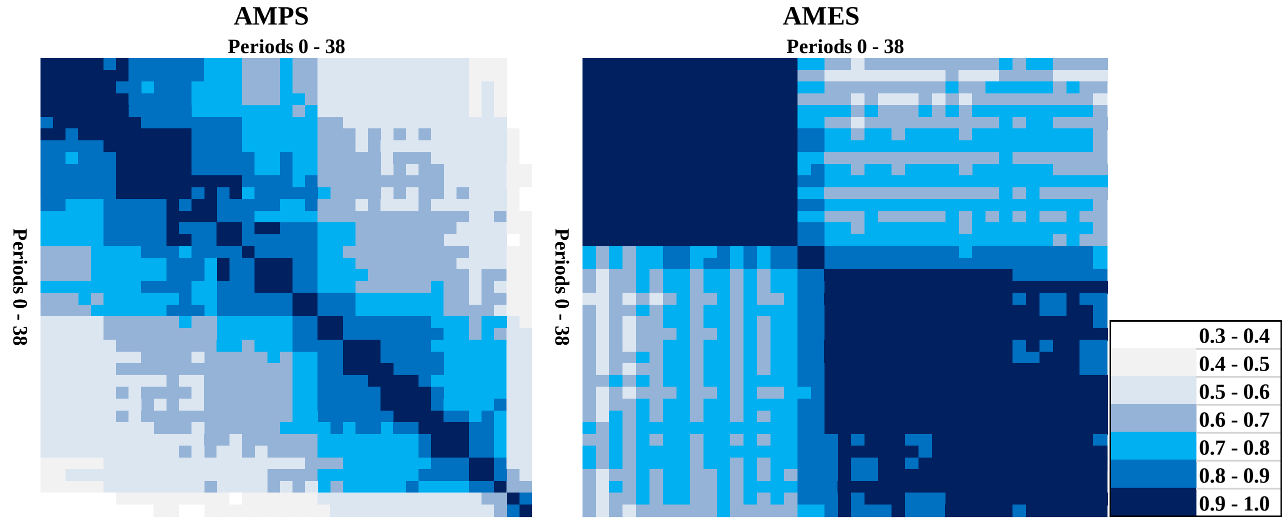
(a) Product aggregate "Chocolate" - Outlet type "Supermarket"



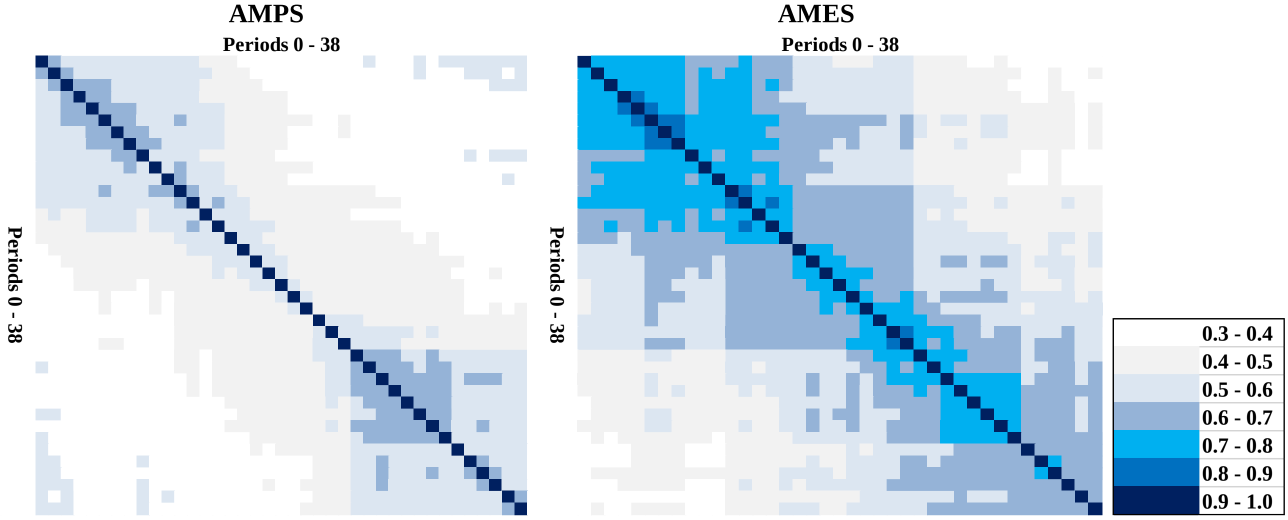
(b) Product aggregate "Packaged ice cream" - Outlet type "Hypermarket"



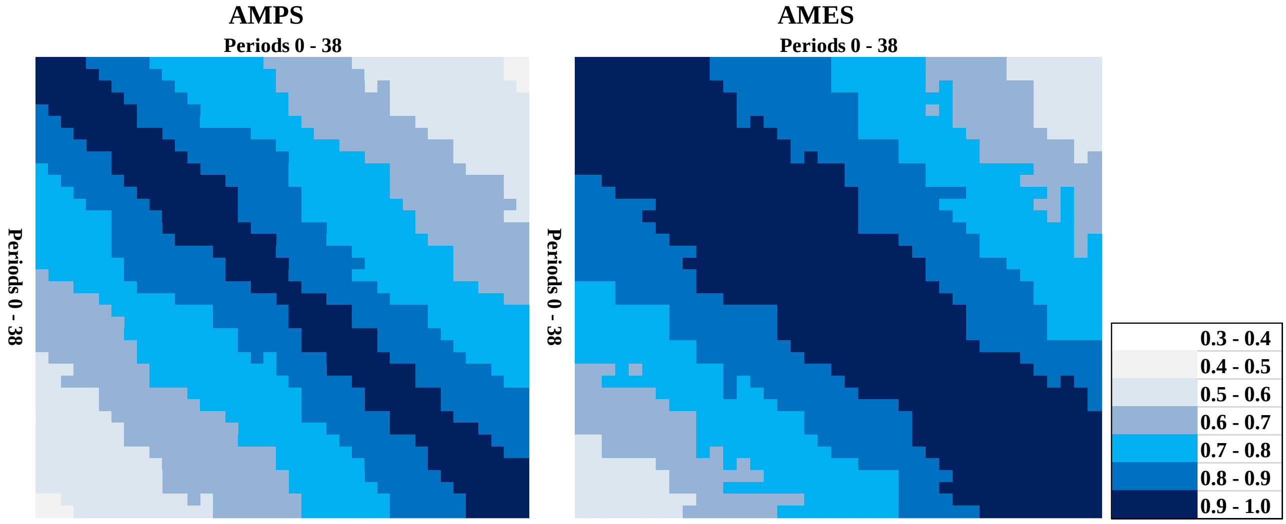
(c) Product aggregate "Olive oil" - Outlet type "Discount"



(d) Product aggregate "Cosmetic products" - Outlet type "Hypermarket"



(e) Product aggregate "Body hygiene products" - Outlet type "Specialist drug"



5. The results of the experimental use of WGEKS

For the experimental use of the weighted and unweighted version of the GEKS, the indices for rolling windows of 25 periods length, named $RWGEKS_{25}$ and $RGEKS_{25}$ respectively, have been compiled with different splicing options[10], together with the full window (39 periods) indices ($WGEKS_{full}$ and $GEKS_{full}$). As for the weights, we used both AMES and AMPS weights. The results of these calculations have then been compared with those obtained with the "standard" GEKS method.

In order to contain the extension of the analysis, we limited the calculation of multilateral indices to the case of Törnqvist-GEKS.

To compile the multilateral indices, we defined the items as the combination of barcode and retailer channel. Accordingly, the average price of the each item is defined as the ratio of the total turnover

[10] In this study, we have considered four different splicing methods: movement splicing, mean splicing, half splicing and windows splicing (the last three, onto published indices).

and quantities sold (monthly) by the shops of the same type[11]. In all, concerning food products (chocolate, ice creams and olive oil) and non-food products (cosmetics and body hygiene products), we have estimated 180 and 150 indices, respectively.

In what follows, the behaviour of the indices is assessed through the comparison of the corresponding annual rates of change. In this perspective, the differences between inflation rates as estimated by WGEKS and GEKS provide a measure of the impact of the introduction of explicit weights in the calculation of the indices. In particular, we are interested in evaluating the size of this impact and in finding evidence on how it modifies when rolling windows indices are considered, also in relation to the different choices of the splicing option. To this aim, it is convenient to start focusing on the indices calculated over the whole time interval covered by our dataset.

Table 3 shows the number of positive and negative differences and their range between the year on year rates of change of WGEKS and GEKS indices, by retailer trade channel and by the system of weights used.

The outcomes of this first set of computations show that the evolution of the indices is very similar. The differences in the inflation rates are generally limited and there seems to be no evidence that the introduction of explicit weights bring to a systematic under or overestimation of the dynamic of the prices of the five products considered, even though for some of them and for some types of shops the discrepancies proved to be relatively large[12]. Moreover, the use of AMPS weights seems to have mostly a bigger impact on the indices than AMES weights have.

Tab. 3. Differences in annual percentage rates of change of full window WGEKS and GEKS – number of positive and negative differences, max and min, by retail trade channel and type of weights.

Product	Retail channel	AMES				AMPS			
		pos	neg	max	min	pos	neg	max	min
choccolate	s.s.a	18	9	0,11	-0,08	21	6	0,07	-0,11
	hyper	13	14	0,14	-0,23	11	16	0,14	-0,21
	super	13	14	0,16	-0,18	11	16	0,18	-0,17
	discount	11	16	0,05	-0,04	10	17	0,06	-0,04
ice cream	s.s.a	14	13	0,14	-0,08	15	12	0,13	-0,09
	hyper	15	12	0,27	-0,21	15	12	0,34	-0,25
	super	15	12	0,12	-0,09	14	13	0,18	-0,11
	discount	12	15	0,03	-0,04	12	15	0,06	-0,05
olive oil	s.s.a	16	11	0,06	-0,05	19	8	0,10	-0,05
	hyper	14	13	0,08	-0,07	15	12	0,14	-0,13
	super	13	14	0,10	-0,09	16	11	0,17	-0,16
	discount	17	10	0,67	-0,37	12	15	0,62	-0,42
cosmetics	s.s.a	8	19	0,06	-0,27	9	18	0,23	-0,39
	hyper	10	17	0,24	-0,36	10	17	0,28	-0,32
	super	12	15	0,10	-0,13	12	15	0,10	-0,14
	discount	9	18	0,16	-0,12	11	16	0,08	-0,13
	s.d.	11	16	0,05	-0,09	10	17	0,04	-0,10
body hygiene products	s.s.a	13	14	0,09	-0,13	13	14	0,10	-0,13
	hyper	11	16	0,16	-0,14	11	16	0,19	-0,16
	super	8	19	0,08	-0,09	11	16	0,09	-0,09
	discount	19	8	0,10	-0,06	18	9	0,11	-0,07
	s.d.	10	17	0,10	-0,22	11	16	0,11	-0,22

[11] The total turnover and quantities at product level have been estimated using the outlets' sample weights.
[12] These results are coherent with the findings of other works in this subject. See Melser (2018).

Looking at the differences between the rolling windows weighted and standard GEKS indices, to some extent they corroborate the previous results. The impact on the annual rates of change is modest (with few exceptions). However, it varies in a range that appears to be affected by the splicing method adopted to link the indices of the rolling windows (Figures 2 and 3).

Fig. 2. Minimum and maximum differences between annual rates of change of weighted and standard RGEKS by product, retailer channel and splicing option. AMES weights

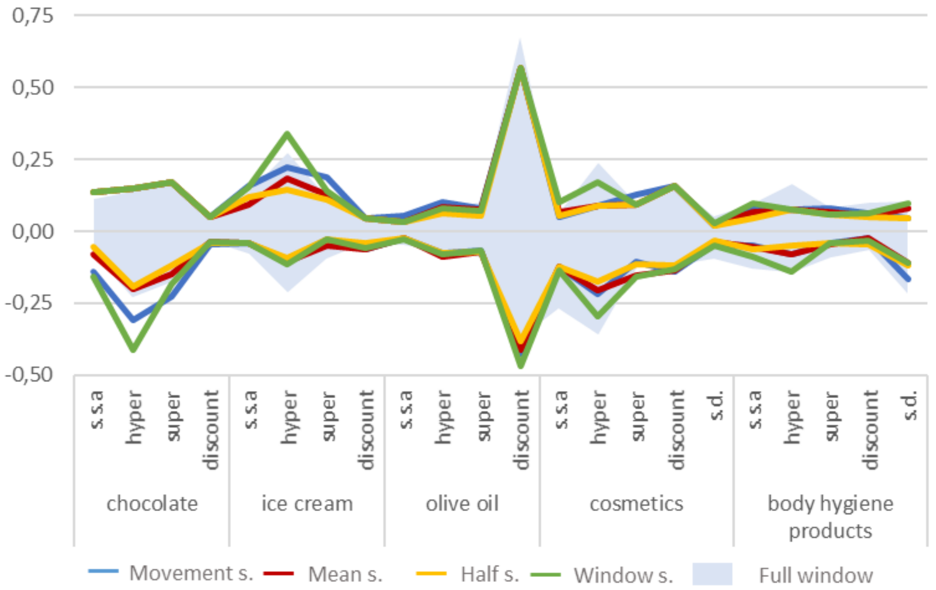
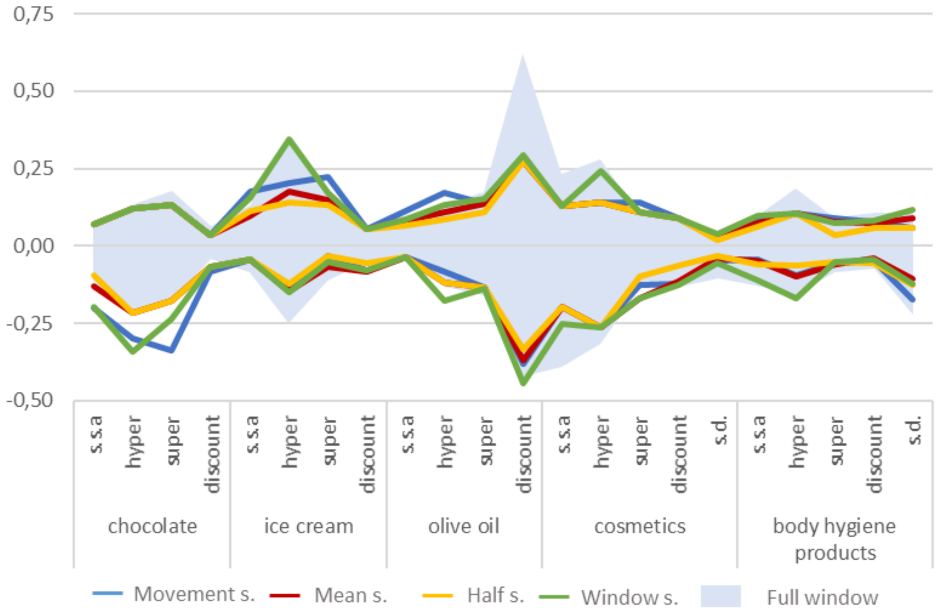


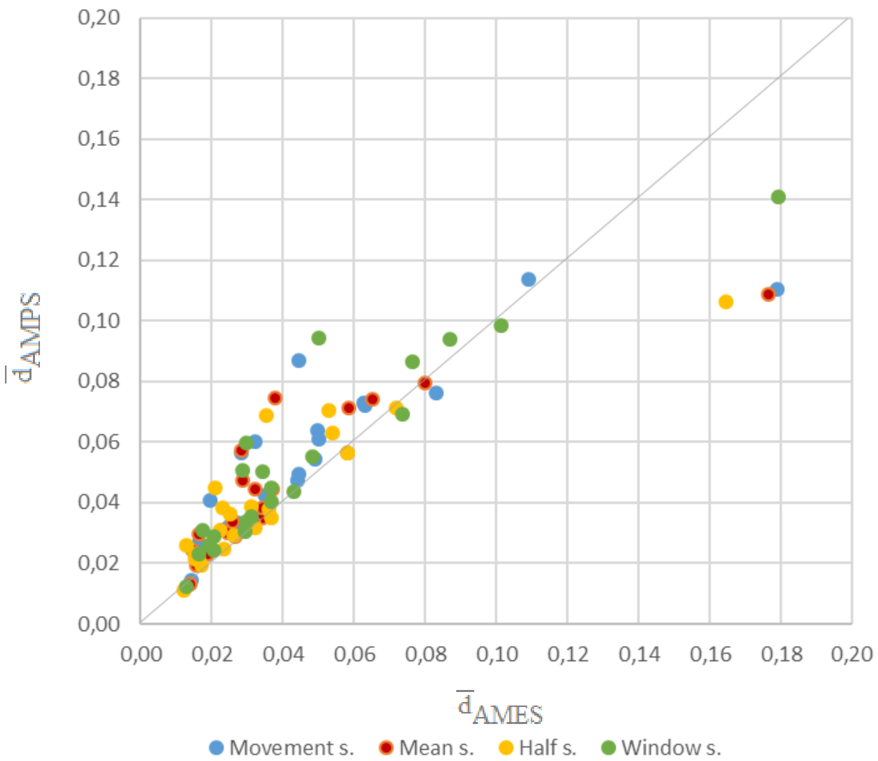
Fig. 3. Minimum and maximum differences between annual rates of change of weighted and standard RGEKS by product, retailer channel and splicing option. AMPS weights



In particular, the adoption of the half splicing and mean splicing seems to have the smaller impact's range in almost all the cases when both AMES and AMPS weights are used. It is worth noting that, as in the case of full window GEKS, the choice of AMPS weights produces relatively wider deviations of the weighted version of RGEKS from the standard counterpart, than AMES weights.

In order to illustrate further this issue, we have calculated the arithmetic mean of absolute differences between the "year on year" rates of change of the weighted and standard version of RGEKS for both the systems of weights. That is, $\bar{d}_{AMES} = \frac{1}{T}\sum_t |x_{AMES,t} - x_t|$, where $x_{AMES,t}$ and x_t represent the annual rate of change of (AMES) weighed RGEKS and the annual rate of its standard correspondent index ($\bar{d}_{AMPS}$ is defined in the analogous way). These are shown in Figure 4. With only few exceptions, the impact of AMES weights is always significantly smaller as compared to AMPS, regardless the splicing method used.

Fig. 4. Arithmetic mean of the absolute differences in inflation rates between weighted RGEKS and standard RGEKS, (AMES and AMPS weights).



Finally, we have analysed the implication of the use of weights in terms of the distance between the full window GEKS and the corresponding rolling window indices.

To this aim, let $\bar{D}_{AMES}$ be the arithmetic mean of the absolute differences between the inflation rates of the full window and of the 25 periods rolling windows weighted GEKS calculated using AMES, that is, $\bar{D}_{AMES} = \frac{1}{T}\sum_t D_{AMES,t} = \frac{1}{T}\sum_t |x^{fw}_{AMES,t} - x^{25w}_{AMES,t}|$. In addition, let $s(D_{AMES,t})$ be the standard deviation of $D_{AMES,t}$.

$\bar{D}_{AMPS}$, $s(D_{AMPS,t})$, $\bar{D}$ and $s(D_t)$ are the equivalent measures for the AMPS weighted GEKS and standard GEKS.

The points in Figure 5 represent the pair $[(\bar{D}_{AMES}$ - $\bar{D})$; $(s(D_{AMES,t}) - s(D_t)]$ calculated on the indices of each product and splicing option.

Interestingly, the scatter graph shows that the $\bar{D}$ is mostly greater than $\bar{D}_{AMES}$: the absolute distance between the rolling windows GEKS and the transitive corresponding index tends to increase when no explicit weights are used for the calculation of the multilateral indices. Moreover the absolute differences D_t tend to be not so close to their average than $D_{AMES,t}$ are. These results are even more evident in the case of AMPS weighted GEKS (Figure 6).

From this point of view, the impact of imposing the non-revisability constraint on the multilateral indices seems to be, generally, less severe when the weighted version (and in particular AMPS weighted version) of GEKS are considered.

Fig. 5. Arithmetic mean and standard deviation of the absolute differences in inflation rates between rolling window and full window AMES weighted and standard GEKS.

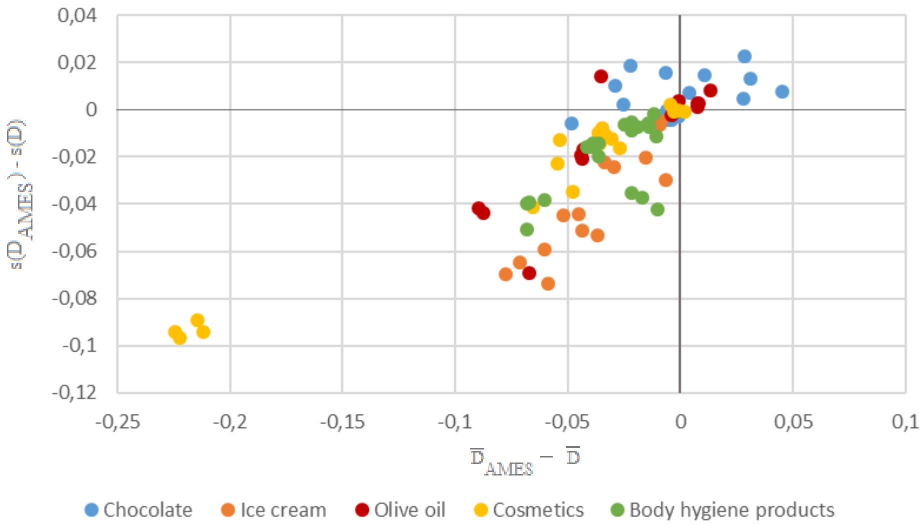
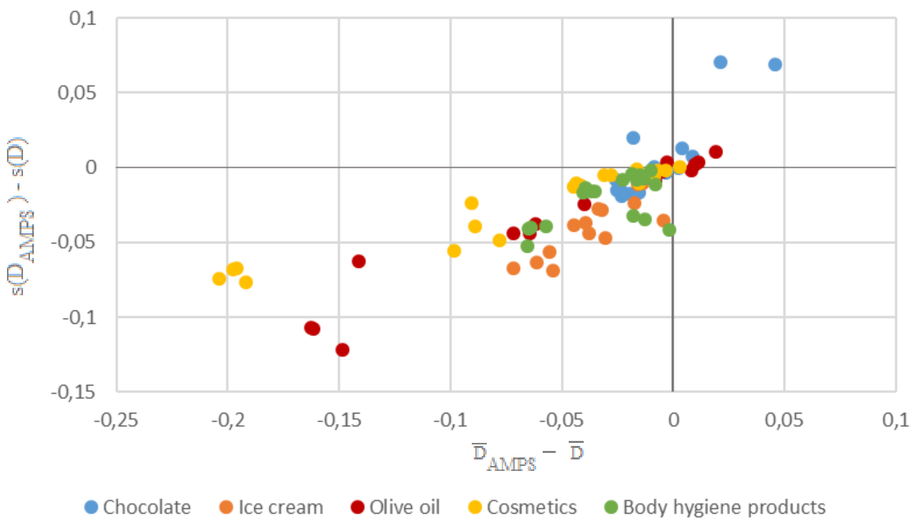


Fig. 6. Arithmetic mean and standard deviation of the absolute differences in inflation rates between rolling window and full window AMPS weighted and standard GEKS.



6. The Simulation framework: CES and Sato-Vartia index

In order to understand which multilateral method should be used to aggregate detailed price and quantity data we explore another approach based on the definition of a “true” cost of living index. We estimate a target cost of living index based on Constant Elasticity of Substitution (CES) purchaser preferences, which has been used in many economics and marketing studies (Diewert and Fox, 2022). By assuming that purchasers have known CES preferences we can construct the corresponding true cost of living indices given our real scanner data set. Then the Törnqvist GEKS and WGEKS will be constructed using the CES simulated data and compared with the corresponding true CES cost of living indices.

The CES unit cost function has the following functional form:

$$C(p_t) = \left(\sum_{i \in U_t} a_i \, p_{it}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \qquad [13]$$

if $\sigma \geq 0$ and $\sigma \neq 1$. When $\sigma = 1$ $C(p_t) = \prod_i p_i^{a_i}$.

where σ and the αn are positive parameters with $\Sigma_i \alpha_i = 1$. The unit cost function defined by [13] corresponds to *a Constant Elasticity of Substitution (CES) utility function* which was introduced into the economics literature by Arrow, Chenery, Minhas and Solow (1961). The parameter σ is the *elasticity of substitution* and governs the degree of substitution between products. When σ = 0, the unit cost function defined by [13] becomes linear in prices and therefore corresponds to a fixed coefficients or Leontief utility function which exhibits 0 substitutability between all commodities. When σ = 1, the corresponding aggregator or utility function is a Cobb-Douglas function. When σ approaches +∞, the corresponding aggregator function *f* approaches a linear aggregator function which exhibits infinite substitutability between each pair of inputs. The CES unit cost function defined by [13] is *not* a fully flexible functional form but it is frequently used to aggregate commodities in a group of commodities which are thought to be highly substitutable with each other (Diewert and Fox, 2022).

Let us consider the price vectors $\mathbf{p}_t \equiv [p_{1t},...,p_{Nt}]$ for t = 1,..,T. If purchasers have CES preferences and are minimizing the costs of achieving their utility levels in each period, it will turn out that the components of their period t expenditure share vectors $\mathbf{s}_t \equiv [s_{1t},...,s_{Nt}]$ for t = 1,..,T will be equal to the following expressions:

$$s_{it} = \frac{a_i p_{it}^{1-\sigma}}{\sum_{i \in U_t} a_i \, p_{it}^{1-\sigma}} \qquad [14]$$

where i = 1,...,N; t = 1,...,T. Thus given the price vectors $\mathbf{p}_t$, the vector of positive parameters $\alpha \equiv [\alpha_1,...,\alpha_N]$ and the nonnegative parameter σ where σ ≠ 1, then the share vectors $\mathbf{s}_t$ can be computed using equations [14] for t = 1,...,T.

Following Melser and Webster (2021) we consider the Sato-Vartia index (Sato, 1976; Vartia, 1976) that takes the form of:

$$I_{0t}^{ST} = \prod_{i \in U_{0t}} \left(\frac{p_{it}}{p_{i0}} \right)^{w_{i0t}} \quad , w_{i0t} = \frac{\left(\frac{\tilde{s}_{it} - \tilde{s}_{io}}{ln\tilde{s}_{it} - ln\tilde{s}_{io}} \right)}{\sum_{i \in U_{0t}} \left(\frac{\tilde{s}_{it} - \tilde{s}_{io}}{ln\tilde{s}_{it} - ln\tilde{s}_{io}} \right)} \quad , \tilde{s}_{it} = \frac{p_{it} q_{it}}{\sum_{i \in U_{0t}} p_{it} q_{it}} \qquad [15]$$

Our expectation is that the Törnqvist WGEKS should perform better than the unweighted GEKS in closely approximating the exact CES results.

6.1 Empirical results

Following previous research studies addressing the issue of comparing multilateral index methods with the corresponding true cost of living indices, often based on the CES function (Diewert and Fox, 2022; Melser and Webster, 2021), we undertake a number of simulations using real data from the 3 product aggregates in the food sector (Chocolate, Packaged ice cream and Olive oil) referred to a specific Italian province (Rome). Given the price vectors p_t, the vector of positive parameters $\boldsymbol{\alpha}$ is set to 1 in our simulations, following Melser and Webster (2022), then the share vectors s_t and $\tilde{s}_{it}$ in equation [15] are computed using prices p_{it} from real scanner data and specifying the nonnegative parameter σ equations equal to 0, 2, 3 and 5 alternatively for t = 1,...,T.

The AMES method is applied to compute the WGEKS based on Törnqvist over the period considered using all the items in the selected product categories for the different type of outlet, that is hypermarket, supermarket, outlets with surface between 100 and 400 s.m. and discounts. The AMES method is selected as it closely concords with the weighting structure in the Törnqvist index number formula.

We estimated the unweighted GEKS and weighted GEKS using AMES as weights over a 25-month window and using as linking method the movement splice. We also computed these three indices over the full window (39 months) As bilateral indices we used the Törnqvist and the Sato-Vartia expressed in [15]. These alternative indices are evaluated for σ equal to 0, 2, 3 and 5 as they are considered a “plausible range” of values of σ in the scanner data context (Ivancic, Diewert and Fox, 2010).

Although this line of analysis warrants further investigation, we illustrate the results for “Olive oil” without distinguishing for outlet type.

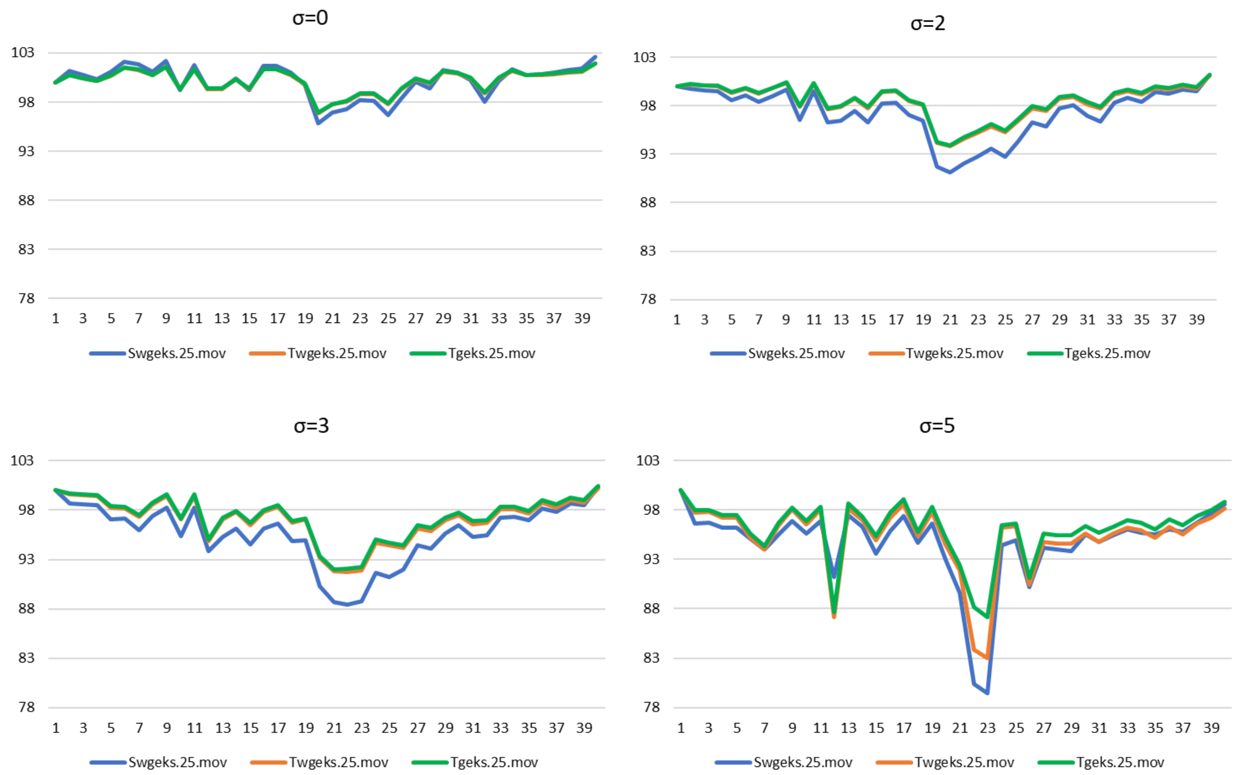
Given that the CES WGEKS Sato-Vartia is our reference point, our expectation is that the Törnqvist WGEKS perform better than the unweighted GEKS in approximating the “true” index.

It can be seen that when σ = 0, the WGEKS and WGEKS cannot be distinguished from each other (Figure 7) and are very close to the CES WGEKS which lie below the other indices in the period 4-9, 19-21 and 31-33. When σ = 2, the WGEKS and WGEKS price levels are slightly above the corresponding the CES price levels. Thus, these indices have little amounts of substitution bias for our real scanner data set. Interestingly, when σ = 5 the WGEKS performs better than the unweighted Törnqvist GEKS with regard to the CES price levels.

This effect is especially evident when hypermarkets are considered (Figure 8).

Fig. 7 Alternative price levels for different methods and elasticities of substitution - Olive Oil - all outlets

Fig. 8 Alternative price levels for different methods and elasticities of substitution - Olive Oil - hyper



7. Conclusions

The generalization of the standard GEKS method, based on Törnqvist indices, used for the compilation of transitive consumer price indices, was tested on real data collected in the province of Rome in the period 2018, December – 2022, February. The groups of products considered are five different product aggregates: Chocolate, Packaged ice cream, Olive oil and Body hygiene products and Cosmetic products.

The different systems of weights considered for introducing information on the reliability of the underlying binaries for the weighting matrix proved to have moderate effect on the dynamic of GEKS but the evidence suggests that the weighted version of rolling windows GEKS, under different splicing options, tend to be slightly closer to the full window counterpart, as compared to standard GEKS.

Therefore, an important result was achieved, even if limited, given the small sample of product and territorial area of application. Indeed, the use of weights in GEKS scheme of indices compilation seems to reduce the impact of the constraint of non-revising the indices themselves. This is evident looking at the absolute distance between the rolling windows GEKS calculated and the transitive corresponding index that tends to increase when no explicit weights are used for the calculation of the multilateral indices.

In general, what emerges from the elaborations carried out is that the impact of imposing the constraint on the multilateral indices of non-revising those calculated for the previous periods, seems to be less severe when the weighted version (and in particular AMPS weighted version) of GEKS are considered.

Another line of research was implemented by comparing standard and weighted GEKS, which is based on the calculation of a target “true” cost of living index under the hypothesis of Constant Elasticity of Substitution (CES) purchaser preferences. The results obtained are interesting but this further line of research is worth to be further deepened.

The evidences coming from the experimentations illustrated in this paper encourage further developments of the studies in the field of the use of multilateral methods and in particular of weighted version of rolling windows GEKS, under different splicing options, to calculate CPI and HICP.

In the last part of 2022 generalization of the experimental compilation of weighted RGEKS for all the grocery products for which scanner data are available and at national level has to be implemented in order to compare the results that will be obtained with other approaches. The objective is to start in Italy, in 2023, a parallel compilation of HICPs and CPIs based on multilateral methods to be compared with the official indicators produced, in order to evaluate the possible adoption of such important methodological innovation starting from 2024 (abandoning the current dynamic approach to the sampling of GTINs).

References

Balk, B.M. (1981), A Simple Method for Constructing Price Indices for Seasonal Commodities, Statistische Hefte 22 (1), 1–8.

Caves D.W., Christensen, L.R. and Diewert, W.E. (1982), Multilateral comparisons of output, input, and productivity using superlative index numbers, Economic Journal 92, 73-86.

Chessa, A.G. (2015), ‘Towards a generic price index method for scanner data in the Dutch CPI’, Room document for the Ottawa Group Meeting, 20–22 May 2015, Urayasu City, Japan.

Chessa, A. G. (2019). A comparison of index extension methods for multilateral methods. In 16th Meeting of the Ottawa Group on Price, Rio de Janeiro, Brazil 08-10 May 2019

Chessa, A. G., Verburg, J., & Willenborg, L. (2017). A comparison of price index methods for scanner data. In 15th Meeting of the Ottawa Group on Price Indices.

De Haan, J., & Van der Grient, H. A. (2011). Eliminating chain drift in price indices based on scanner data. Journal of Econometrics, 161(1), 36-46.

De Haan, J., & Krsinich, F. (2014). Scanner data and the treatment of quality change in nonrevisable price indices. Journal of Business & Economic Statistics, 32(3), 341-358.

Ivancic, L., W.E. Diewert and K.J. Fox (2010), “Using a Constant Elasticity of Substitution Index to estimate a Cost of Living Index: from Theory to Practice”, Australian School of Business Research Paper No. 2010 ECON 15, University of New South Wales, Sydney 2052 Australia

Diewert, W. E., & Fox, K. J. (2022). Substitution Bias in Multilateral Methods for CPI Construction. Journal of Business & Economic Statistics, 40(1), 355-369.

Drechsler, L. 1973. Weighting of Index Numbers in Multilateral International Comparisons. Review of Income and Wealth 19, 17–34.

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

Feenstra, R. C., Ma, H., and Rao, D. P. (2009). Consistent comparisons of real incomes across time and space. Macroeconomic Dynamics, 13(S2):169–193.

Kokoski, M. F., Moulton, B. R., & Zieschang, K. D. (1999). Interarea price comparisons for heterogeneous goods and several levels of commodity aggregation. International and Interarea Comparisons of Income, Output and Prices, Studies in Income and Wealth, 61, 123-169.

Krsinich, F. (2014), 'The FEWS Index: Fixed Effects with a Window Splice — Non-revisable qualityadjusted price indices with no characteristic information', Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 26–28 May 2014, Geneva, Switzerland.

Melser, D. (2018). Scanner data price indices: addressing some unresolved issues. Journal of Business & Economic Statistics, 36(3), 516-522.

Melser, D., & Webster, M. (2021). Multilateral Methods, Substitution Bias, and Chain Drift: Some Empirical Comparisons. *Review of Income and Wealth*, 67(3), 759-785.

Rao, D. S. Prasada, (1997) Aggregation Methods for International Comparison of Purchasing Power Parities and Real Income: Analytical Issues and Some Recent Developments, Bulletin of International Statistical Institute, LVII, 197–200.

Rao, D. S. Prasada (2001). Weighted EKS and generalised CPD methods for aggregation at basic heading level and above basic heading level. In Joint World Bank-OECD seminar on Purchasing Power Parities, Washington DC (Vol. 30).

Rao, D. S. Prasada, & Timmer, M. P. (2003). Purchasing Power Parities for Industry Comparisons Using Weighted Elteto–Koves–Szulc (EKS) Methods. *Review of Income and Wealth,* 49(4), 491-511.

Rao, D. S. Prasada, E. A. Selvanathan and D. Pilat, (1995) Generalized Theil–Tornqvist Indices with Applications to International Comparisons of Prices and Real Output," Review of Economics and Statistics, 352–60.

Silver, M., & Heravi, S. (2005). A failure in the measurement of inflation: Results from a hedonic and matched experiment using scanner data. Journal of Business & Economic Statistics, 23(3), 269-281.

Vartia, Y.O. (1976), "Ideal Log-Change Index Numbers", Scandinavian Journal of Statistics 3, 121-126.