Jacek Białek[1], Anna Bobel[2]

16th Meeting of the Ottawa Group on Price Indices, Rio de Janeiro, 8-10 May 2019

Comparison of Price Index Methods for the CPI Measurement Using Scanner Data

Abstract

Scanner data are a quite new data source for statistical agencies and the availability of electronic sales data for the calculation of the Consumer Price Index (CPI) has increased over the past 16 years. Scanner data can be obtained from a wide variety of retailers (supermarkets, home electronics, Internet shops, etc.) and provide information at the level of the barcode, i.e. the Global Trade Item Number (GTIN, formerly known as the EAN code). One of advantages of using scanner data is the fact that they contain complete transaction information, i.e. prices and quantities for every sold item. It means that we may use expenditure shares of items as weights for calculating price indices at the lowest (elementary) level of data aggregation. One of new challenges connected with scanner data is the choice of the index formula which should be able to reduce the chain drift bias and the substitution bias. In this paper, we compare several price index methods for CPI calculations based on scanner data. In particular, we consider bilateral index methods with chained versions of direct weighted and unweighted indices, and also selected multilateral index methods, i.e. the quality adjusted unit value method (QU method) and its special case (the Geary-Khamis method), the augmented Lehr method, the so called "real time index", the GEKS method and the CCDI method. We also propose some price index modifications. We verify the impact of window updating methods and also different weighting schemes in quantity weights on the price index, i.e. we consider alternatively the QU-TS method and the QU-EW method. We compare all these methods using artificial data sets and real scanner data sets obtained from one supermarket and *allegro.pl.*

Keywords: Scanner data, Consumer Price Index, superlative indices, elementary indices, chain indices, QU-GK index, Geary-Khamis method, real time index, GEKS, bilateral indices, multilateral indices.

JEL Classification: C43, E31

1. Introduction

Scanner data mean transaction data that specify turnover and numbers of items sold by GTIN (barcode, formerly known as the EAN code). Scanner data have numerous advantages compared to traditional survey data collection because such data sets are much bigger than traditional ones and they contain complete transaction information, i.e. information about prices and quantities.

[1] Department of Statistical Methods, University of Lodz, Lodz, Poland, jacek.bialek@uni.lodz.pl.
Central Statistical Office in Poland, Department of Trade and Services, Poland, J.Bialek@stat.gov.pl

[2] Central Statistical Office in Poland, Department of Trade and Services, Poland, A.Bobel@stat.gov.pl

In other words, scanner data contain expenditure information at the item level (i.e. at the barcode or the GTIN level), which makes it possible to use expenditure shares of items as weights for calculating price indices at the lowest (elementary) level of data aggregation.

Scanner data from two supermarkets were introduced in the Dutch CPI in 2002 and, in January 2010, the number of supermarkets providing the scanner data was extended to six. The Dutch CPI was re-designed (de Haan (2006), van der Grient & de Haan (2010), de Haan and van der Grient (2011)). In 2017, scanner data of ten supermarkets chains were used and at present surveys are not carried out anymore for supermarkets, i.e. scanner data from other retailers (for instance, from do-it-yourself stores or from travel agencies) are used in the Dutch CPI (Chessa (2015)). Until 2015, four EU countries were using scanner data (the Netherlands, Norway, Sweden, and Switzerland). The number of countries that make use of scanner data in their CPI has been growing, i.e. in April 2016, the number of EU countries increased to seven (Belgium, Denmark and Iceland started to use such data sets) and at present, some of national statistical institutes (NSIs) consider starting to use scanner data. Some other countries consider using scanner data in their CPI calculation in the nearest future (or have just started using it), for instance: the French National Statistical Institute (INSEE) launched in 2010 a pilot project in order to get some insights into the suitability of these data for CPI purposes, the Statistics Portugal was awarded in 2011 a Eurostat grant to undertake the initial research on the exploitation of scanner data, in Luxembourg, collaboration was put in place with several retailers who agreed to transmit every month their data to the IT system (STATEC) and scanner data can be introduced in the regular production from January 2018. In January 2018, in Poland, the project titled "INSTATCENY" began and its main aim is to create the new methodology of CPI measurement based on data from different (traditional and untraditional) sources, including scanner data and web-scraped data. In 2017, the Eurostat provided *Practical Guide for Processing Supermarket Scanner data,* which is commonly available on website: *https://ec.europa.eu/eurostat/web/hicp/overview*). In the above-mentioned guide, we can read: "This guide describes the situation in 2017. It will need to be updated as the use of scanner data develops and broadens". In fact, the methodology for CPI (or HICP) construction using scanner data has strongly evolved over the last few years (see for instance: Ivancic et. al. (2011), Krsnich (2014), Griffioen & Bosch (2016), de Haan at. al. (2016), Chessa & Griffioen (2016), Chessa (2017), Diewert & Fox (2017)). One of new challenges connected with scanner data is the choice of the index formula which should be able to reduce the chain drift bias and the substitution bias.

In this paper, we compare several price index methods for CPI calculations based on scanner data. The paper is organised as follows: Section 2 describes a selected bilateral and multilateral index method which can be used in the case of scanner data and this Section also discusses updating and weighting problem connected with multilateral methods; Section 3 proposes some price index modifications; Section 4 presents the results from our simulation study and examines the influence of price and quantity dispersions on the characteristic of bilateral and multilateral index methods; Section 5 presents the empirical study based on real scanner data sets obtained from one supermarket and the e-commerce platform *allegro.pl*; Section 6 lists the main conclusions.

2. Index methods for CPI calculations using scanner data

Most statistical agencies use bilateral index numbers in the CPI measurement, i.e. they use indices which compare prices and quantities of a group of commodities from the current period with the corresponding prices and quantities from a base (fixed) period. In multilateral methods, we collect information about prices and quantities of a group of commodities from T periods and next we calculate a sequence of price indices for these T periods. Although Ivancic, Diewert and Fox (2011) have suggested that the use of multilateral indices in the scanner data case can solve the chain drift problem, most statistical agencies using scanner data still make use of the monthly chained Jevons index (Chessa et. al. (2017)). Since the elementary Jevons price index belongs to bilateral (direct) index methods, we start our description of possible methods with these methods. Following Chessa et. al. (2017), let us denote the sets of homogeneous products belonging to the same product group in months 0 and t by G_0 and G_t respectively, and let $G_{0,t}$ denote the set of matched products in both moments 0 and t. A product may refer to a single item (GTIN) or to a sub-group of items (GTINs) having the same characteristics, and thus being in the same homogeneity group. In the next part of the paper, we consider the second scenario, i.e. a homogeneous group of different GTINs but having identical characteristics. We also consider a month as a time period over which scanner data are aggregated. In fact, one month is the longest interval among time intervals recommended by Eurostat for the scanner data aggregation (see *Practical Guide for Processing Supermarket Scanner data* (2017), page 13) although, the same document on the same page states: “Most commonly, scanner data are collected weekly, i.e. all transactions taking place during a week are aggregated”.

2.1. **Bilateral index methods**

2.1.1. *Unweighted formulas*.

A recommendation of the European Commission concerning the choice of the elementary formula at the lowest level of data aggregation can be found on website: *http://www.ilo.org/public/english/bureau/stat/download/cpi/corrections/annex1.pdf* and it is as follows: "For the HICPs the ratio of geometric mean prices or the ratio of arithmetic mean prices are the two formulae which should be used within elementary aggregates. The arithmetic mean of price relatives may only be applied in exceptional cases and where it can be shown that it is comparable". In other words, if expenditure information is not available, the European Commission recommends the Jevons (1865) price index (see also Diewert (2012) or Levell (2015)), which can written as follows

$$P_J^{0,t} = \prod_{i \in G_{0,t}} \left(\frac{p_i^t}{p_i^0}\right)^{\frac{1}{N_{0,t}}} , \tag{1}$$

where p_i^τ denotes the price of the i-th product at the time $\tau \in \{0,t\}$ and $N_{0,t} = card\ G_{0,t}$. On the other hand, the same recommendation takes also into consideration ("in exceptional cases") the Carli (1804) price index, which can be written as follows

$$P_C^{0,t} = \frac{1}{N_{0,t}} \sum_{i \in G_{0,t}} \frac{p_i^t}{p_i^0} . \tag{2}$$

In our research, we consider only the first formula (1) together with its monthly chained version which is denoted here by $P_{CH-J}^{0,t}$.

2.1.2. Weighted formulas

Since scanner data contain information about the expenditure, it is possible in their case to calculate weighted bilateral indices. *Superlative* price indices, firstly proposed by Diewert (1976), are the most recommended index formulas for the scanner data case (as base formulas). Following Chessa et. al. (2017), we consider the Törnqvist (1936) price index, which is given by

$$P_T^{0,t} = \prod_{i \in G_{0,t}} \left(\frac{p_i^t}{p_i^0}\right)^{\frac{s_i^0 + s_i^t}{2}}, \tag{3}$$

where s_i^0 and s_i^t denote the expenditure shares of matched products in months 0 and t.

Other commonly known superlative price indices are the Fisher price index (1922) and the Walsh price index (1901). Their formulas, denoted by $P_F^{0,t}$ and by $P_W^{0,t}$ respectively, can be written as follows:

$$P_F^{0,t} = \sqrt{P_{La}^{0,t} \cdot P_{Pa}^{0,t}}, \tag{4}$$

$$I_W^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t \cdot \sqrt{q_i^0 q_i^t}}{\sum_{i \in G_{0,t}} p_i^0 \cdot \sqrt{q_i^0 q_i^t}}, \tag{5}$$

where q_i^0 and q_i^t denote quantities of matched products in months 0 and t, $P_{La}^{0,t}$ and $P_{Pa}^{0,t}$ denote the Laspeyres price index (1864) and the Paasche price index (1874) respectively (see Section 2.2.1). In the next part of the paper only the Fisher and the Törnqvist price indices are taken into consideration.

2.2. **Multilateral index methods**

Multilateral index methods have their genesis in comparisons of price levels across countries or regions. These methods satisfy the transitivity, which is a desirable property for spatial comparisons due to the fact that the results are independent of the choice of base country (region). Commonly known methods are the GEKS method (also known as the EKS method – see Gini (1931), Eltetö and Köves (1964), Szulc (1964), the Geary-Khamis (GK) method (Geary (1958), Khamis (1972)), the CCDI method (Caves, Christensen and Diewert (1982), Inklaar and Diewert (2016)) or the *real time index method* (Chessa (2015)). In this paper, we consider most of these methods but the problem of the best choice of the multilateral formula seems to be still open.

2.2.1. The quality adjusted unit value index and the Geary-Khamis (GK) method

The term "Quality adjusted unit value method" (shortened to the "QU method") was introduced by Chessa (see, for instance, Chessa (2015, 2016)). The QU method is a family of unit value based index methods with the above-mentioned Geary-Khamis (GK) method as a special case. According to the QU method, the price index $P_{QU}^{0,t}$ which compares the period t with the base period 0 is defined as follows

$$P_{QU}^{0,t} = \frac{\sum_{i \in G_t} p_i^t q_i^t / \sum_{i \in G_0} p_i^0 q_i^0}{\sum_{i \in G_t} v_i q_i^t / \sum_{i \in G_0} v_i q_i^0}, \quad (6)$$

where the numerator in (6) is the measure of the turnover (expenditure) change between the two considered months and the denominator in (6) is a weighted quantity index. Note that both the turnover index and the weighted quantity index are transitive, and thus the price index $P_{QU}^{0,t}$ is also transitive (Chessa at. al. (2017)). Note also that the quantity weights v_i are the only unknown factors in formula (6) and these factors convert sold quantities q_i^0 and q_i^t into "common units" $v_i q_i^0$ and $v_i q_i^t$. Prices of products, p_i^0 and p_i^t, are converted into "quality adjusted prices" p_i^0 / v_i and p_i^t / v_i. If the considered consumption segment is homogeneous, then product quantities can be summed (factors v_i are equal for all products) and the index $P_{QU}^{0,t}$ simplifies to the unit value index (the nominator of (6)). If the above-mentioned consumption segment is not homogeneous, then the unit value index must be adjusted. Note also that the formula $P_{QU}^{0,t}$ defines a family of price indices. In fact, limiting considerations to products sold in both moments 0 and t, and setting v_i equal to the product prices in the current period t, the formula (6) leads to the Laspeyres index

$$P_{La}^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t q_i^0}{\sum_{i \in G_{0,t}} p_i^0 q_i^0}. \quad (7)$$

Similarly, if we consider the group of products $G_{0,t}$ and if the quantity weights v_i are set equal to the prices in the base period (month) 0, then the formula $P_{QU}^{0,t}$ simplifies to the Paasche price index, i.e.

$$P_{Pa}^{0,t} = \frac{\sum_{i \in G_{0,t}} p_i^t q_i^t}{\sum_{i \in G_{0,t}} p_i^0 q_i^t} . \tag{8}$$

In other words, different choices of factors v_i lead to different prices index formulas. In the GK method, the weights v_i are defined as follows

$$v_i = \sum_{z=0}^{T} \varphi_{i,GK}^z \frac{p_i^z}{P_{QU}^{0,z}} , \tag{9}$$

where

$$\varphi_{i,GK}^z = \frac{q_i^z}{\sum_{\tau=0}^{T} q_i^\tau} , \tag{10}$$

and $[0,T]$ is the entire time interval of the product observations (typically $T = 12$, see Diewert & Fox (2017)). Please note that formulas (6), (9) and (10) lead to a set of equations which should be solved simultaneously. The above-mentioned solution can be found iteratively (Maddison and Rao (1996), Chessa (2016)) or as the solution to an eigenvalue problem (Diewert (1999)). An interesting alternative method for obtaining this solution can be also found in Diewert & Fox (2017).

2.2.2. The augmented Lehr index

The Lehr method is similar to the Geary-Khamis method (see Section 3.2.1, formula (6) with weights defined in (9)) but it does not use the complex iterative method. The quality adjusted factors v_i are defined here as follows

$$v_i = \frac{p_i^0 q_i^0 + p_i^T q_i^T}{q_i^0 + q_i^T} . \tag{11}$$

The immediate conclusion from (11) is that the Lehr index uses only data from months 0 and T, and in fact this is a bilateral index. Nevertheless, we can change the formula of the quality adjustment factors, and thus, similarly to multilateral methods, we take into considerations all available information from the interval , i.e. (see Loon & Roels (2018))

$$v_i = \frac{\sum_{\tau=0}^{T} p_i^\tau q_i^\tau}{\sum_{\tau=0}^{T} q_i^\tau} \quad . \tag{12}$$

In the next part of the paper, the augmented Lehr index, i.e. the index constructed as in (6) with quantity weights defined in (12), will be denoted by $P_{AL}^{0,t}$ and the above-mentioned factors will be signified by v_i^{AL}. In other words, the considered augmented Lehr index can be written as follows

$$P_{AL}^{0,t} = \frac{\sum_{i \in G_t} p_i^t q_i^t / \sum_{i \in G_0} p_i^0 q_i^0}{\sum_{i \in G_t} v_i^{AL} q_i^t / \sum_{i \in G_0} v_i^{AL} q_i^0} \, . \tag{13}$$

2.2.3. The real time index

Let us note that price imputations are not needed when prices from each month of the current year are included in weights v_i. Taking typically value $T = 12$, Chessa (2015) suggests defining these weights by including product prices and quantities from each month of the current year and the base month December of the previous year (there are 13 months together). However, as the same author admits, in practice, we can use prices and quantities of all 13 months only in the final month of the year, and thus some updating method is needed for v_i calculations each month. Although there are several methods for updating quantity weights (see for instance Krsinich (2014)), we focus on an interesting and quite easy for implementation method proposed by Chessa (2015). He suggests the following procedure of calculating *the real time index*: (1) For the current year, we use a time window with December of the previous year as the fixed base month and the window is enlarged each month with the current month; (2) The price index of the current month t is calculated by using the updated quantity weights according to a special algorithm. In particular, this algorithm needs some initial values of price indices $P_{QU}^{0,\tau}$: $0 \le \tau \le t$ and it repeats updating weights $v_i = \sum_{z=0}^{t} \varphi_{i,GK}^z \frac{p_i^z}{P_{QU}^{0,z}}$ and next updating values of price indices $P_{QU}^{0,\tau}$: $0 \le \tau \le t$ (according to (6)) until the difference between indices from the last two iterations is small enough. Chessa (2015) recommends a method for calculating initial indices. Moreover, he sets the stop criterion at 0.001 and assumes the maximum absolute difference between the price index vectors as a distance measure. Nevertheless, in our study, we set the stop criterion at 0.0001 and we use the Euclidean distance for comparisons of two successive iterations. Steps (1) and (2) are repeated until December of the current year and after that the base month is shifted to December of the current year. In this way, the whole procedure may be repeated in the subsequent year. For more details, see also Chessa (2016).

2.2.4. The GEKS method

Let us consider a time interval $[0,T]$ of observations of prices and quantities which will be used for the GEKS index construction. The GEKS price index between months 0 and t is an unweighted geometric mean of $T+1$ ratios of bilateral price indices $P^{\tau,t}$ and $P^{\tau,0}$ which are based on the same price index formula. The bilateral price index formula should satisfy the time reversal test, i.e. it should satisfy the condition $P^{a,b} \cdot P^{b,a} = 1$. Typically, the GEKS method uses the superlative Fisher price index and in such case the GEKS formula can be written as follows

$$P_{GEKS}^{0,t} = \prod_{\tau=0}^{T} \left(\frac{P_F^{\tau,t}}{P_F^{\tau,0}}\right)^{\frac{1}{T+1}}. \qquad (14)$$

2.2.5. The CCDI method

The GEKS method for making international index number comparisons between countries comes from Gini (1931) but it should be mentioned that it was derived in a different manner by Eltetö and Köves (1964) and Szulc (1964). Feenstra, Ma and Rao (2009), and also De Haan and var der Grient (2011) suggested that the Törnqvist price index formula (see (3)) could be used instead of the Fisher price index in the Gini methodology. Caves, Christensen and Diewert (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) and Inklaar & Diewert (2016) extended the CCD methodology to making price comparisons across production units. Thus, in the paper of Diewert and Fox (2017), the multilateral price comparison method that uses the GEKS method based on the Törnqvist price index is called the CCDI method. The corresponding CCDI price index can be expressed as follows

$$P_{CCDI}^{0,t} = \prod_{\tau=0}^{T} \left(\frac{P_T^{\tau,t}}{P_T^{\tau,0}}\right)^{\frac{1}{T+1}}. \quad (15)$$

2.2.6. Other methods

In the literature, we can find some other multilateral index methods which are not considered in this paper. The Country-Product Dummy (CPD) method proposed by Summers (1973) has been adapted for spatial price comparisons to the time domain and now it is known as the Time Product Dummy (TPD) method (de Haan & Krsinich (2014)). The multilateral hedonic method is closely related to the TPD method, i.e. its model parameters (known as "item fixed effects") are not estimated for items (as in the TPD method) but they are estimated for the characteristics of items (attributes). Both the TPD method and the above-mentioned hedonic method do not simplify to a unit value index when all products are homogeneous and they are flawed with regard to their use of turnover in constructing weights (Chessa (2015)). Some other methods can be encountered in the paper of Haan at al. (2016), for instance, the so-called "Cycle Method" (see also Willenborg (2010, 2017), Willenborg and van der Loo (2016)). In Section 5, we propose some price index modifications instead of presenting these above-mentioned and omitted here methods.

2.3. **Alternative weighting schemes in the QU method**

In the classical form, the GK method uses quantity shares as weight in the construction of v_i. In the literature, we can find at least two other weighting schemes in quantity weights for the GK price index. The first variant was proposed by Hill (2000) and it assumes that deflated prices, i.e. $p_i^z / P_{QU}^{0,z}$, are weighted by the ratio of the turnover share of the i-th product in the month z (denoted here by s_i^z) and the sum of turnover shares of the same product over different months. In the paper of Chessa (2016), this variant is referred to as the "QU-TS" method but we use here the shortened notation "TS", i.e. we denote the above-mentioned weights for deflated prices as $\varphi_{i,TS}^z$. In other words, in the TS method, weights $\varphi_{i,GK}^z$ are replaced by weights calculated as follows

$$\varphi_{i,TS}^z = \frac{s_i^z}{\sum_{\tau=0}^{T} s_i^\tau}, \quad (16)$$

and the final quantity weights are computed as follows

$$v_i = \sum_{z=0}^{T} \varphi_{i,TS}^{z} \frac{p_i^z}{P_{QU}^{0,z}}. \tag{17}$$

The other weighting scheme assumes that deflated prices in months with sales receive equal weight, and thus it is denoted here by the EW method (in Chessa (2016), this method is referred to as the "QU-EW" method). In other words, in the considered weighting scheme, we use the following weights for deflated prices

$$\varphi_{i,EW}^{z} = \frac{\delta_i^z}{\sum_{\tau=0}^{T} \delta_i^{\tau}}, \tag{18}$$

where $\delta_i^z = 1$ if $q_i^z > 0$ and $\delta_i^z = 0$ otherwise. Analogically to (17), in the EW method, the final quantity weights can be written as

$$v_i = \sum_{z=0}^{T} \varphi_{i,EW}^{z} \frac{p_i^z}{P_{QU}^{0,z}}. \tag{19}$$

In the next part of the paper, we will use different notations for quantity weights defined in (10), (17) and (19), i.e. these weights, connected with the GK, TS and EW methods, will be signified by v_i^{GK}, v_i^{TS} and v_i^{EW} respectively. Similarly, the corresponding multilateral indices, which compare the time moment t with the time moment 0, will be denoted by $P_{GK}^{0,t}$, $P_{TS}^{0,t}$ and $P_{EW}^{0,t}$ respectively.

2.4. Updating problem and window updating methods

In the case of bilateral methods, a fixed base month (period) is used and the current period is shifted each month. In monthly chained index methods, the base and the current month are both moved one month. The problem with proceeding with the next month arises in the case of multilateral index methods. Adding information from a new month may influence the values of quality adjustment parameters and values of the corresponding multilateral indices. In this paper, we consider four commonly used rolling-window updating methods which shift the estimation window (often 13 months) forward each period (a month as a rule) and then splice the new indices onto the existing time series. The considered methods are as follows:

2.4.1. The movement splice method

According to the movement splice method (de Haan & van der Grient (2011)), a price index for the new month is calculated by chaining the month-on-month index for the last month of the shifted window to the index of the previous month (the last month of the previous window). The movement splice method can be described by the following recursive formula

$$P_{MS}^{0,t} = P_{MS}^{0,t-1} \cdot P_{t-T,t}^{t-1,t}, \tag{20}$$

where P is any multilateral price index formula, the subscript makes reference to the window period and the superscript indicates the period for which the index is calculated (see Loon & Roels (2018)).

2.4.2. The window splice method

The window splice method proposed by Krsinich (2014) calculates the price index for the new month by chaining the indices of the shifted window to the index of T months ago (i.e. to the index of 12 months ago for windows of 13 months). It can be written by the following general formula

$$P_{WS}^{0,t} = P_{WS}^{0,t-1} \cdot \frac{P_{t-T,t}^{t-T,t}}{P_{t-T-1,t-1}^{t-T,t-1}}. \tag{21}$$

2.4.3. The half splice method

De Hann (2015) suggested that the link period t_0 should be chosen to be in the middle of the first time window and the Australian Bureau of Statistics (2016) called this the half splice method for linking the results of two time windows. In other words, according to this method, the half splice happens at $t_0 = \frac{T+1}{2}$ if T is an odd integer and at $t_0 = \frac{T}{2}$ if T is an even integer. A recursive formula for the half splice method is as follows

$$P_{HS}^{0,t} = P_{HS}^{0,t-1} \cdot \frac{P_{t-T,t}^{t-t_0,t}}{P_{t-T-1,t-1}^{t-t_0,t-1}}. \tag{22}$$

2.4.4. The mean splice method

The mean splice method (Diewert & Fox (2017)) uses the geometric mean of all possible choices of splicing, i.e. all months $\{1,2,...,T\}$ which are included in the current window and the previous one. The general formula for the mean splice method can be written as

$$P_{GMS}^{0,t} = P_{GMS}^{0,t-1} \cdot \prod_{t_0=1}^{T} \left(\frac{P_{t-T,t}^{t-t_0,t}}{P_{t-T-1,t-1}^{t-t_0,t-1}}\right)^{\frac{1}{T}} . \tag{23}$$

It should be mentioned here that the above-presented method of calculating the real time index (see Chessa (2016)) is also a rolling-window updating method and it is called the fixed base monthly expanding window method. The corresponding *real time index* is often denoted by $P_{FBEW}^{0,t}$. In the literature, we can also encounter some other more or less popular window methods, such as the fixed base moving window method (Lamboray (2017)), but they are not considered in our work.

3. Propositions of price index modifications

As it was mentioned above, most statistical agencies using scanner data still make use of the monthly chained Jevons index (Chessa et. al. (2017)). Due to the fact that for some reasons (such as the well-known and broad list of axiomatic properties) some countries want to stay with the Jevons formula, let us consider its modification for the scanner data case (see Section 3.1.). As a consequence, we will consider also a modification of the GEKS index where the base superlative price index formula is replaced by the proposed modified Jevons index (see Section 3.2.). Finally, we will also verify an alternative system of weights in the Geary-Khamis method (Section 3.3).

3.1. Modification of the Jevons formula

Let us consider a homogeneous group of products, i.e. a group of different items (GTINs) but having identical characteristics. The standard matched-model Jevons formula $P_J^{0,t}$ (see (1)) treats each matched item singularly, i.e. it does not take into account the fact that scanner data provide also information about quantities of these matched items. However, let us note that if the total sales of the i-th matched item at the time moment τ is described by the available quantity information q_i^τ, then we could treat this item at the considered moment not as a singular product but rather as q_i^τ identical products (or product's units) sold at the same price p_i^τ. In other words, we could consider a homogeneous group of $\sum_{i \in G_\tau} q_i^\tau$ products (product's units) instead of a homogeneous group of $N_\tau = cardG_\tau$ products. Our considerations lead to the following modification of the classical Jevons price index described in (1):

$$P_{MJ}^{0,t} = \sqrt[\sum_{j \in G_{0,t}} \frac{q_j^0 + q_j^t}{2}]{\prod_{i \in G_{0,t}} (\frac{p_i^t}{p_i^0})^{\frac{q_i^0 + q_i^2}{2}}} = \prod_{i \in G_{0,t}} (\frac{p_i^t}{p_i^0})^{\frac{q_i^0 + q_i^t}{\sum_{j \in G_{0,t}} (q_j^0 + q_j^t)}} . \quad (24)$$

In the next part of the paper, we consider the index $P_{MJ}^{0,t}$ together with its monthly chained version which is signified here by $P_{CH-MJ}^{0,t}$.

3.2. **Modifications of the GEKS index**

The standard GEKS index uses the superlative Fisher price index as a base price index formula in its body (see (14)). In fact, the minimum requirement for the above-mentioned base bilateral index is the time reversal test (Chessa et. al. (2017)). It is easy to verify that the modified Jevons price index (24) satisfies the time reversal test, and thus we decided to use it (in the place of the Fisher price index) in the case of the GEKS formula (14). We introduce the multilateral JGEKS index in the following form

$$P_{JGEKS}^{0,t} = \prod_{\tau=0}^{T} (\frac{P_{MJ}^{\tau,t}}{P_{MJ}^{\tau,0}})^{\frac{1}{T+1}} . \quad (25)$$

Justifications for using the JGEKS index are as follows: (a) from the "definition", we can sum up quantities at the lowest level of data aggregation; (b) weights which are used in the JGEKS index reflect a pure consumer reaction to price changes, i.e. weights do not depend directly on artificially fixed prices (such as special discounts or promotions). It may be important when the consumers' reaction to price changes is delayed. (c) finally, the difference between the GEKS and JGEKS indices may serve as a measure of rationality of consumers, i.e. it seems that when prices and quantities are strongly correlated, then the difference between these two indices is very small.

3.3. **Modification of the Geary-Khamis index**

In the classical form, the GK method uses quantity shares as weight in the construction of v_i. In Section 2.3, we mentioned some other discussed weighting schemes which could be interesting alternatives in the GK index construction. Now we suggest considering a different system of weights based on observed and available expenditures, namely

$$\varphi_{i,EX}^{z} = \frac{p_i^z q_i^z}{\sum_{\tau=0}^{T} p_i^\tau q_i^\tau}, \tag{26}$$

which allows us to calculate the final quantity weights in the QU method as follows

$$v_i = \sum_{z=0}^{T} \varphi_{i,EX}^{z} \frac{p_i^z}{P_{QU}^{0,z}}. \tag{27}$$

We will denote these quantity weights by v_i^{EX} and the corresponding QU index, i.e. the index defined in (6) but using weights v_i^{EX} instead of weights v_i, by $P_{EX}^{0,t}$.

4. Simulation study

Case 1

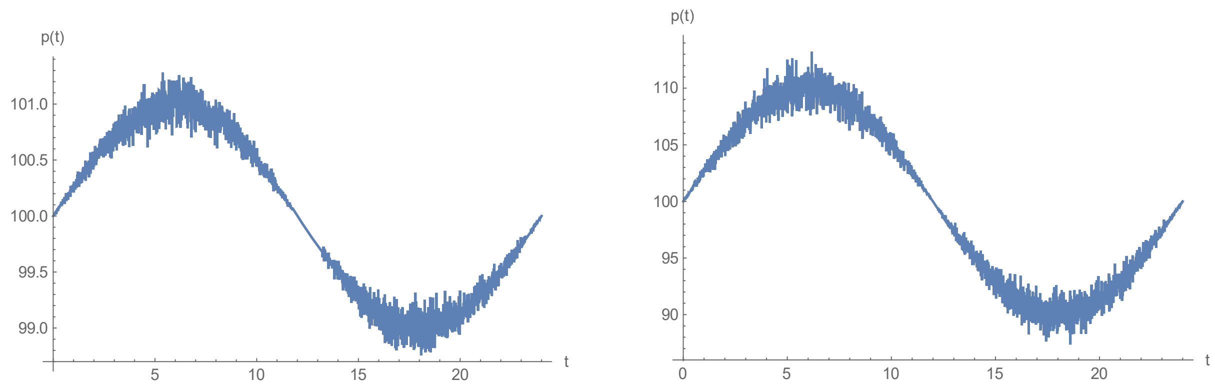
In the first experiment, we are going to verify the chain drift effect in the case of bilateral and multilateral indices. Chain drift occurs when an index does not return to unity when prices in the current period return to their levels in the base period (ILO 2004, p. 445). For instance, Szulc (1983), (1987) demonstrated how big the chain problem could be with chained Laspeyres indices but also, as it is commonly known, chain drift can also be a problem with chained superlative indices. Some authors consider the chain drift problem more narrowly, i.e. they assume that only when both prices and quantities in the current period revert back to their levels in the base period, a corresponding price index should indicate that no price change occurred (Diewert and Fox (2017), von Auer (2019)). Potentially, multilateral methods should deal with the chain problem in this "narrow" sense. In particular, in the paper of Diewert and Fox (2017), we can read about GEKS indices: "they satisfy Walsh's multi-period identity test so (…) the above indices are free from chain drift". Nevertheless, the problem with scanner data is that quantity vectors may strongly differ in compared months, i.e. we may observe seasonal goods, disappearing or new goods or products which are deleted from the sales offer for some reasons. In particular, when prices return to their base level, we often observe quantities with totally different values compared to the base period. The set of matched items for months *t-1* and *t* may differ from the set of matched items for months *t* and *t+1*. As a consequence, even multilateral indices may not return to unity when prices revert back to the levels in the base period but quantities do not. It will be now illustrated. Let us consider a group of $N = 40$ matched items observed during two years, i.e. each month during the time interval [0, 24]. Let us assume that the price of k-th item can be described by the following stochastic process:

$p_k^t = 100 + k \cdot \sin(x \cdot t) \cdot Y$,

where $x = \frac{2\pi}{24}$ and the random variable Y is normally distributed, i.e. $Y \sim N(1;0.1)$. Thus, we have: $E(p_k^t) = 100 + k \sin(x \cdot t)$, $D(p_k^t) = k \sin(x \cdot t)$ and $p_k^0 = p_k^{12} = p_{24}^{24}$. Sample realisations of price processes (for $k = 1$ and $k = 10$) are presented in Fig. 1.

Fig. 1. Sample realisations of price processes for $t \in [0,24]$

a) $k = 1$ b) $k = 10$



To take into considerations a wide spectrum of quantity cases, we consider the following deterministic quantity processes:

Case 1.1. (periodic quantities negatively correlated with prices)

$q_k^t = 1000 - k \cdot \sin(x \cdot t)$

Case 1.2. (periodic quantities positively correlated with prices)

$q_k^t = 1000 + k \cdot \sin(x \cdot t)$

Case 1.3. (strongly decreasing quantities uncorrelated with prices)

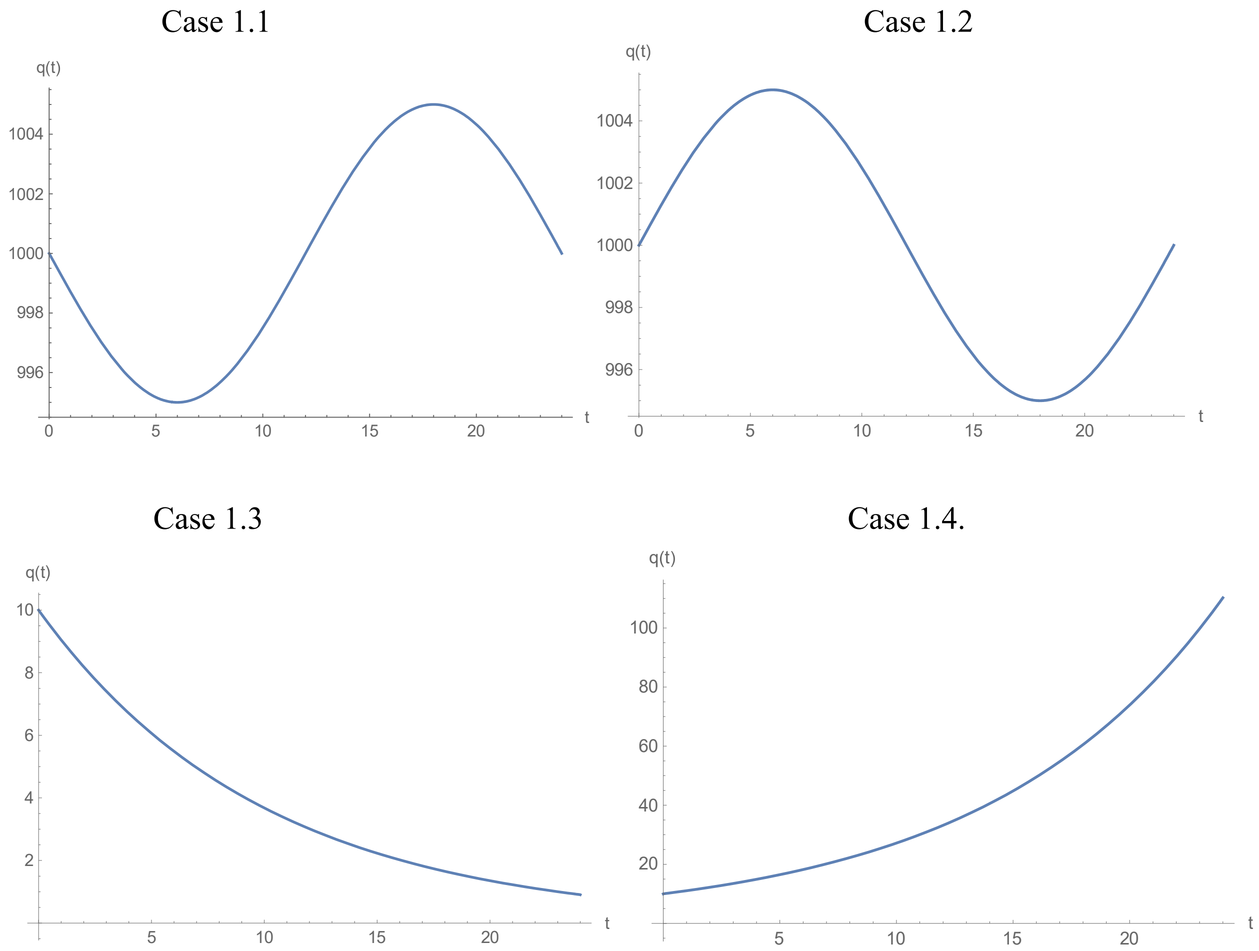
$$q_k^t = 10 \cdot \exp(-\frac{k}{50} \cdot t)$$

Case 1.4. (strongly increasing quantities uncorrelated with prices)

$$q_k^t = 10 \cdot \exp(\frac{k}{50} \cdot t).$$

Sample realisations of quantity processes for $k = 5$ and for all Cases 1.1 – 1.4 are presented in Fig. 2.

Fig. 2. Sample realisations of quantity processes for $t \in [0,24]$ and for $k = 5$.



Now we measure the price dynamics comparing the given month t to the base month 0. In Cases 1.1 and 1.2, when quantity processes are not strongly fluctuated and are correlated with price movements, all indices (unweighted, weighted, including multilateral ones) equal 1 for $t \in \{12,24\}$. In these cases, quantities revert to the starting level after one and two years. The differences between indices are negligible (see sample Fig. 3, Fig. 4 and Fig. 5).

Fig. 3. Values of selected indices (Case 1.1)

(in the case of multilateral indices, a 13-month window is considered, $T = 12$, $t \in [0,12]$)

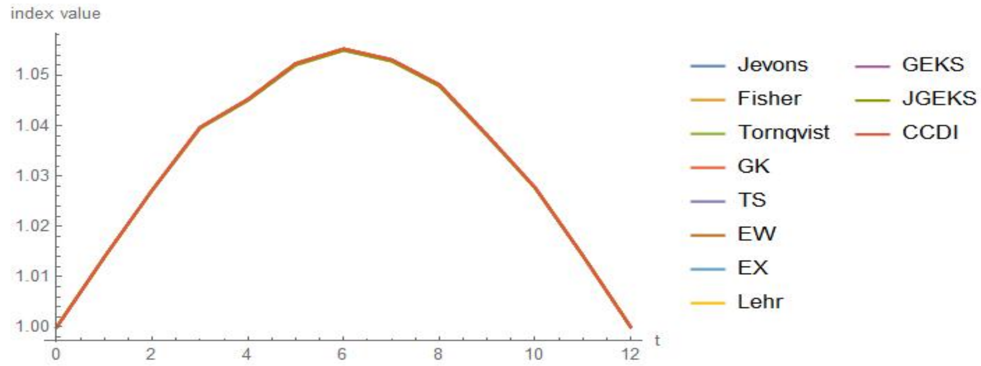


Fig. 4. Values of selected multilateral indices (Case 1.1)

(the whole time window is available, $T = 24$, $t \in [0,24]$)

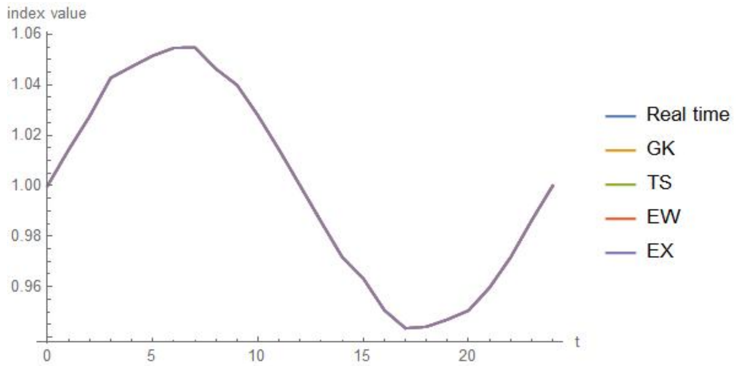
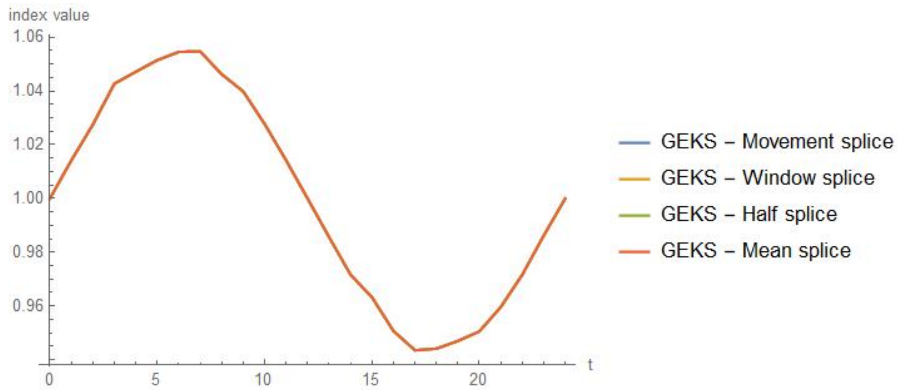


Fig. 5. Values of the GEKS index (Case 1.1)

(a 13-month window is considered, $T = 12$, $t \in [0,24]$)



The situation in Cases 1.3 and 1.4 is different. When quantities strongly decrease (Case 1.3), chained superlative indices and multilateral indices seem to slightly overestimate the real price change for time intervals [0,12] and [0,24] (see Fig. 6, Fig. 7, Fig. 8 and Tab.1). The

Lehr index is now the most sensitive in the case of the choice of the window updating method (see Fig. 9).

Fig. 6. Values of superlative indices and their chained versions (Case 1.3, $t \in [0,12]$)

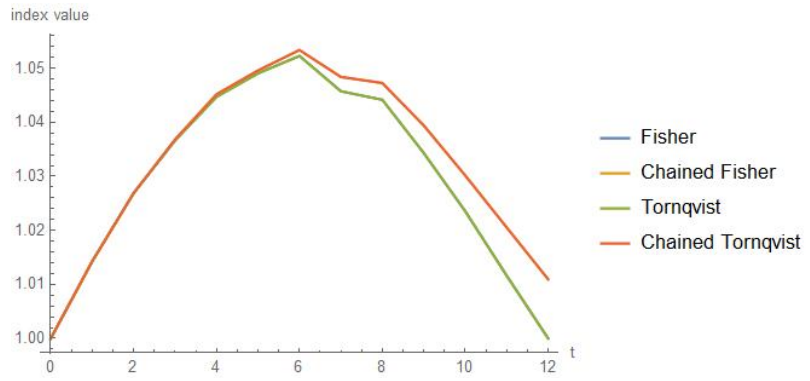


Fig. 7. Values of selected multilateral indices (Case 1.3)

(a 13-month window is considered, $T = 12$, $t \in [0,12]$)

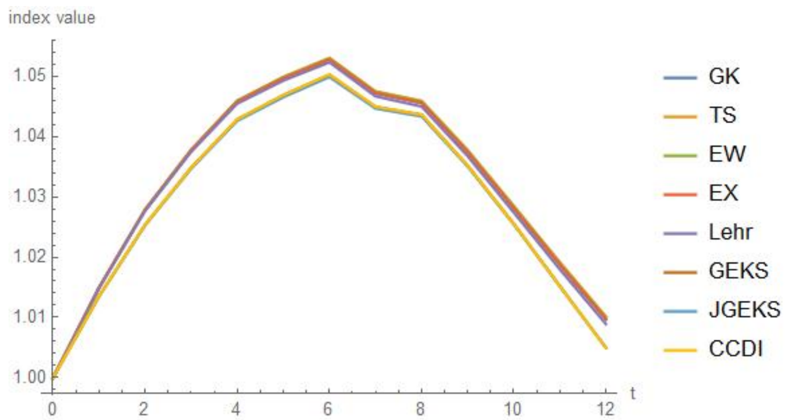


Fig. 8. Values of selected multilateral indices (Case 1.3)

(the whole time window is available, $T = 24$, $t \in [0,24]$)

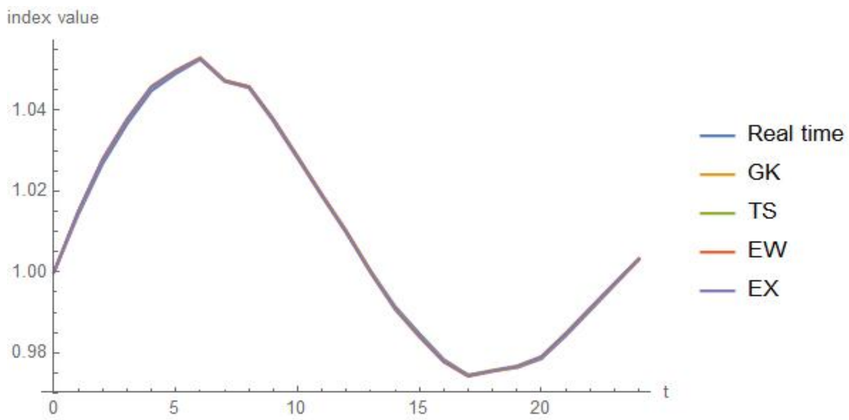
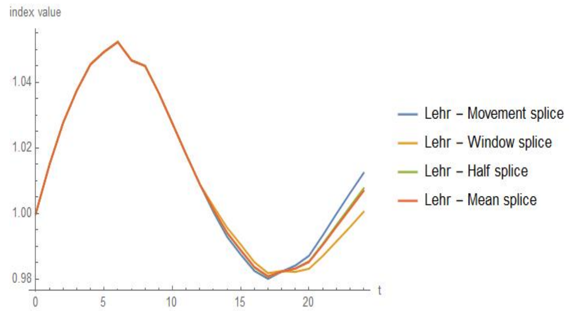


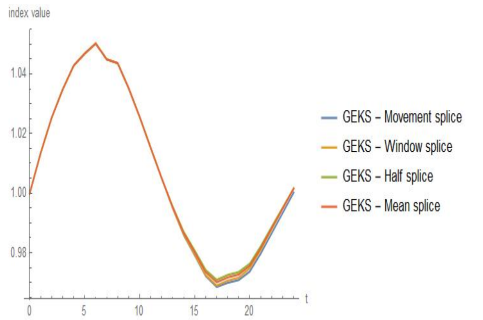
Fig. 9. Values of selected multilateral indices for different window updating methods

(Case 1.3, a 13-month window is considered, $T = 12$, $t \in [0,24]$)

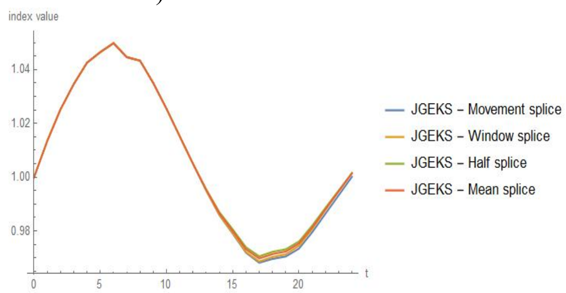
a) Lehr index



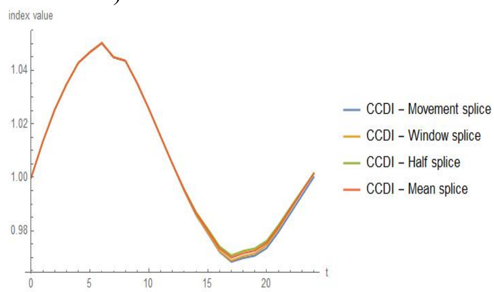
b) GEKS index



c) JGEKS index



d) CCDI index



Tab. 1. Values of considered indices
(Case 1.3, a 13-month window is considered, $T = 12$, $t \in \{12,24\}$)

Index formula	**Time interval**	
	[0,12]	[0,24]*
Classical indices		
Jevons	1.00000	1.00000
Chained Jevons	1.00000	1.00000
Fisher	1.00000	1.00000
Chained Fisher	1.01096	1.00351
Törnqvist	1.00000	1.00000
Chained Törnqvist	1.01096	1.00350
Multilateral indices		
GK	1.00965	1.00894
TS	1.01017	1.00359
EW	1.01003	0.99933
EX	1.00977	1.00939
Real Time	1.00965	1.00894
GEKS	1.00500	1.00153
JGEKS	1.00498	1.00143
CCDI	1.00498	1.00140
Lehr	1.00891	1.00680

(*) the mean splice method is used

When quantities strongly decrease (Case 1.4), as a rule chained superlative indices and multilateral indices seem to be slightly below the real price change for time intervals [0,12] and [0,24] (see Fig. 10, Fig. 11, Fig. 12 and Tab.2). The Lehr index is now the most sensitive in the case of the choice of the window updating method (see Fig. 13).

Fig. 10. Values of superlative indices and their chained versions (Case 1.4, $t \in [0,12]$)

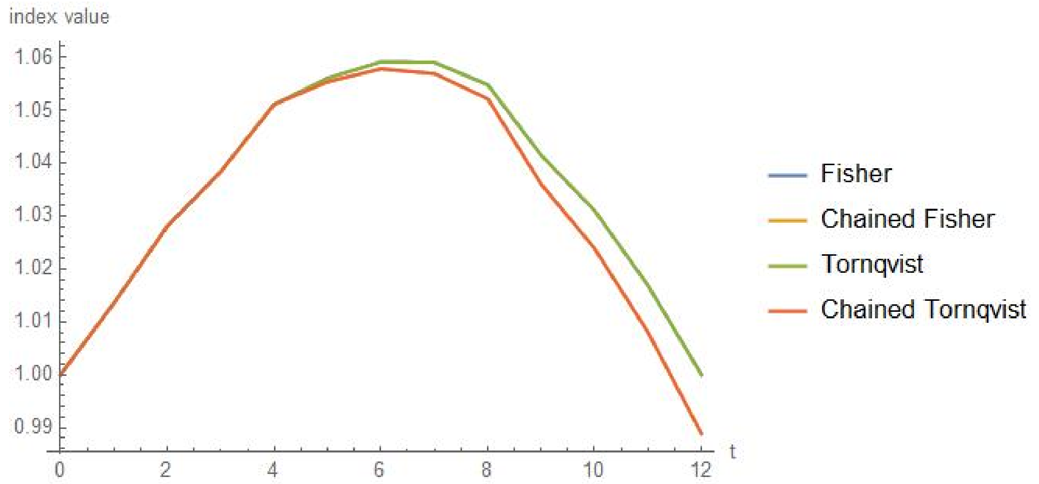


Fig. 11. Values of selected multilateral indices (Case 1.4)

(a 13-month window is considered, $T = 12$, $t \in [0,12]$)

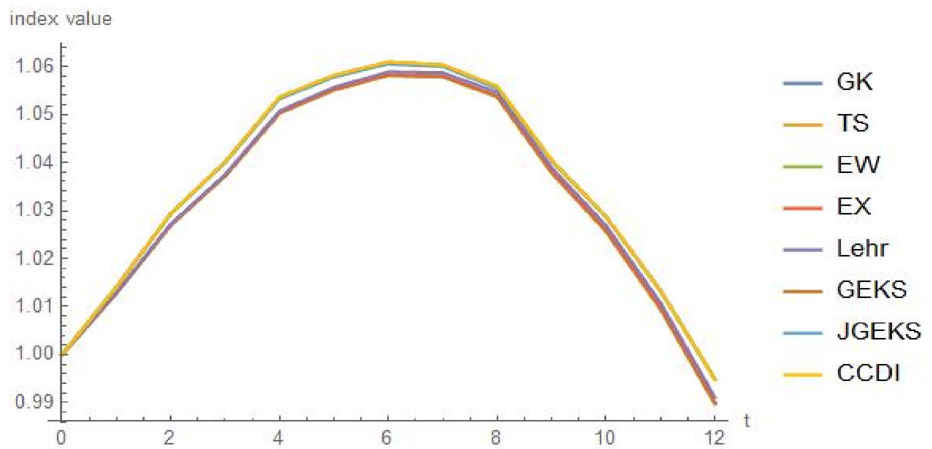


Fig. 12. Values of selected multilateral indices (Case 1.4)

(the whole time window is available, $T = 24$, $t \in [0,24]$)

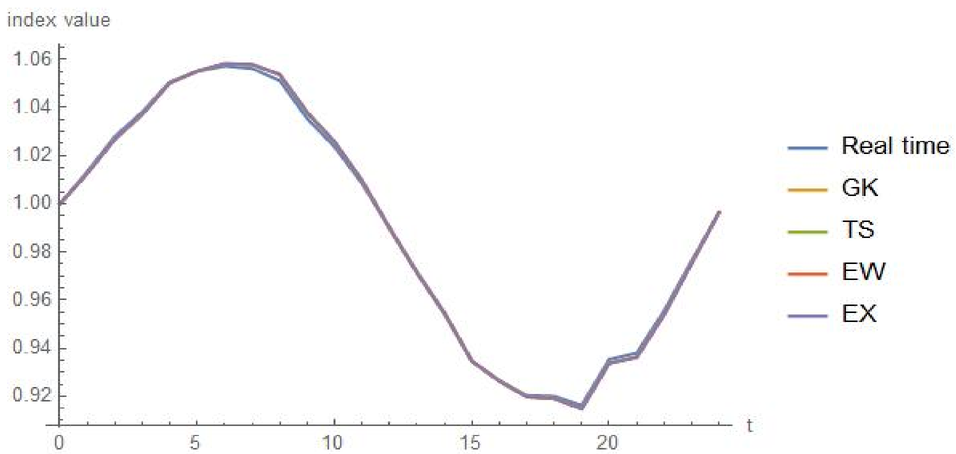
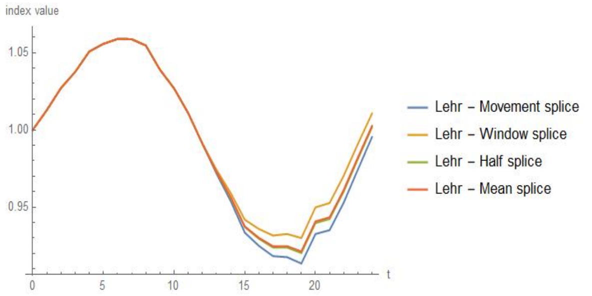


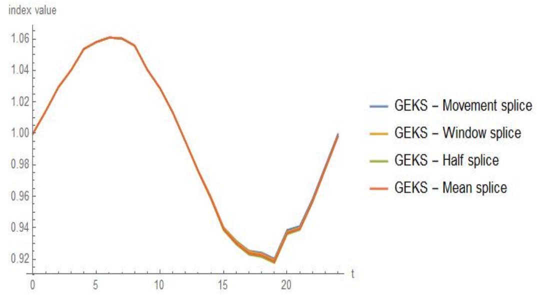
Fig. 13. Values of selected multilateral indices for different window updating methods

(Case 1.4, a 13-month window is considered, $T = 12$, $t \in [0,24]$)

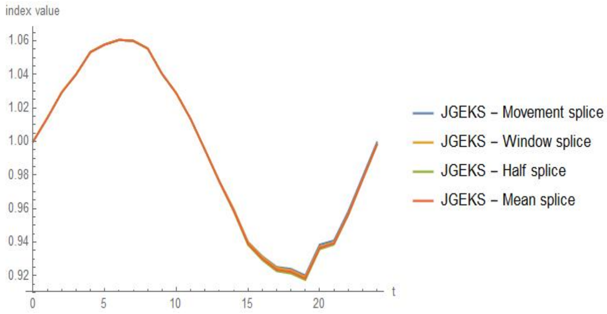
a) Lehr index



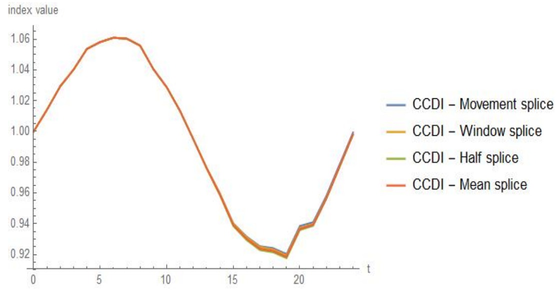
b) GEKS index



c) JGEKS index



d) CCDI index



Tab. 2. Values of considered indices
(Case 1.4, a 13-month window is considered, $T = 12$, $t \in \{12,24\}$)

Index formula	**Time interval**	
	[0,12]	[0,24]*
Classical indices		
Jevons	1.00000	1.00000
Chained Jevons	1.00000	1.00000
Fisher	1.00000	1.00000
Chained Fisher	0.98885	0.99609
Törnqvist	1.00000	1
Chained Törnqvist	0.98885	0.99608
Multilateral indices		
GK	0.99012	1.00882
TS	0.98967	0.99648
EW	0.98975	1.00021
EX	0.99000	1.00840
Real Time	0.99012	1.00882
GEKS	0.99486	0.99815
JGEKS	0.99485	0.99806
CCDI	0.99482	0.99803
Lehr	0.99112	1.00257

(*) the mean splice method is used

Case 2

A geometric Brownian motion (GBM) (also known as exponential Brownian motion) is a continuous-time stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion (also called a Wiener process) with drift (see Oksendal, 2002; Privault, 2012). *The main arguments for using the GBM price model are as follows: (a) t*he expected returns (relative price changes) are independent of the value of the process (price), which is consistent with what we would expect in reality; (b) the GBM process only assumes positive values, just like real commodity prices; (c) the GBM process shows the same kind of 'roughness' in its paths as we see in real prices; (d) estimations of its parameters are relatively easy. In our simulation study, we use the GBM model for generating price processes and, having known the expected value of obtained price shares, we compare values of calculated multilateral indices with these theoretical ones. We assume that the given $i-$th price process satisfies the following stochastic differential equation

$$dp_i^t = \alpha p_i^t dt + \beta p_i^t dW_i^t, \tag{28}$$

where the percentage drift α and the percentage volatility β are constant, and $\{W_i^t : 0 \le t < \infty, i = 1,2,...,N\}$ are independent Wiener processes. The solution for the stochastic differential (7) is as follows (Oksendal, 2002, Jakubowski et al., 2003):

$$p_i^t = p_i^0 \exp((\alpha - \frac{\beta^2}{2})t + \beta W_i^t), \tag{29}$$

and we assume that all initial prices p_i^0 are deterministic. As a consequence, we obtain

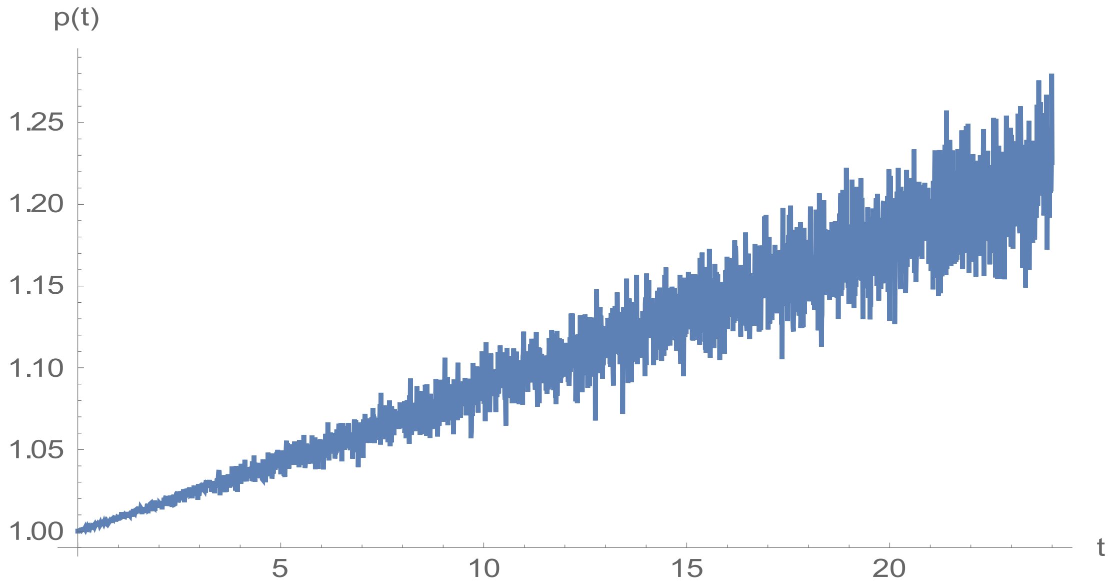
$$E(P^t) = E(P_i^t) = \exp(\alpha t), \tag{30}$$

and

$$Var(P^t) = Var(P_i^t) = \exp(2\alpha t)[\exp(\beta^2 t) - 1], \tag{31}$$

where P_i^t is the $i-$th price relative and $P^t = \frac{p^t}{p^0}$ denotes the (unknown) population price index that we want to estimate. Since we assume that all price processes have the same probability distribution at each time moment (the drift and the volatility parameter are identical for all price processes), we could expect that any measure of price dynamics in the time interval [0,t] will provide the value $\exp(\alpha t)$. In our study, we generate $N = 1000$ price processes to exclude accidental results. Sample realisation of price process (for $\alpha = 0.1$ and $\beta = 0.02$) is presented in Fig. 14.

Fig.14. Sample realisation of price process (for $\alpha = 0.1$ and $\beta = 0.02$)



We consider the following cases of quantity processes:

Case 2.1 (periodic quantities)

$$q_k^t = 1000 + k \cdot \sin(x \cdot t)$$

Case 2.2. (strongly decreasing quantities)

$$q_k^t = 10 \cdot \exp(-\frac{k}{50} \cdot t)$$

Case 2.3. (strongly increasing quantities)

$$q_k^t = 10 \cdot \exp(\frac{k}{50} \cdot t).$$

We consider the following values of parameters: $\alpha = 0.1$ and $\beta \in \{0.02,\ 0.05,\ 0.1\}$. As a consequence, the expected price changes for time intervals [0,12] and [0,24] (normalised into [0,1] and [0,2]) are 1.10517 and 1.2214 respectively. Values of selected multilateral indices calculated for these two time intervals are presented in Tab.3 – 6 and in Fig. 15 (a 13-month window and the mean splice method are used).

Tab.3. Values of selected multilateral indices for considered time intervals – **Case 2.1**

Index value	Time interval [0,12] (theoretical index value = 1.10517)			Time interval [0,24] (theoretical index value = 1.2214)		
	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$
GK	1.10492	1.10064	1.13713	1.22484	1.21391	1.26426
TS	1.10492	1.10064	1.13713	1.22484	1.21391	1.26426
EW	1.10492	1.10064	1.13713	1.22484	1.21391	1.26426
EX	1.10492	1.10064	1.13713	1.22484	1.21391	1.26426
Real Time	1.10492	1.10064	1.13713	1.22484	1.21391	1.26426
GEKS	1.10495	1.10072	1.13721	1.22497	1.21390	1.26400
JGEKS	1.10471	1.09907	1.13096	1.22426	1.20550	1.23890
CCDI	1.10495	1.10073	1.13726	1.22497	1.21392	1.26387
Lehr	1.10492	1.10064	1.13713	1.22487	1.21389	1.26405

Tab.4. Values of selected multilateral indices for considered time intervals – **Case 2.2**

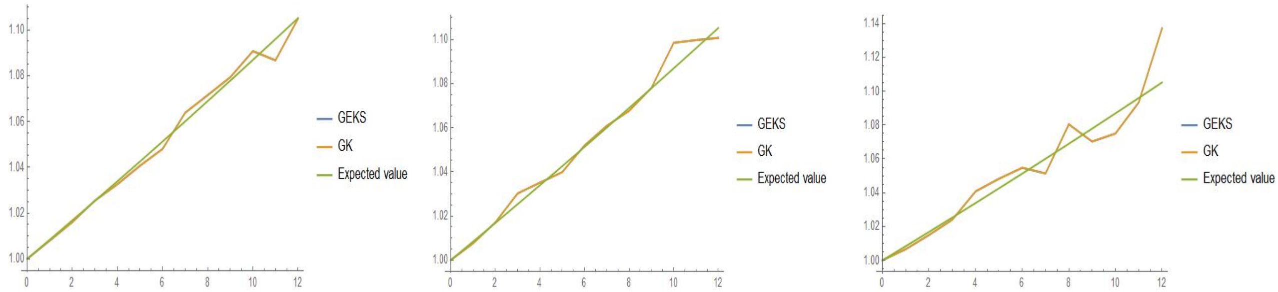
Index value	Time interval [0,12] (theoretical index value = 1.10517)			Time interval [0,24] (theoretical index value = 1.2214)		
	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$
GK	1.09738	1.09406	1.09444	1.21575	1.27987	1.32131
TS	1.09738	1.09447	1.09450	1.21527	1.27571	1.30300
EW	1.09782	1.09469	1.09556	1.21516	1.27580	1.31908
EX	1.09803	1.09387	1.09345	1.21566	1.27758	1.30853
Real Time	1.09837	1.09406	1.09444	1.21575	1.27987	1.32131
GEKS	1.09786	1.09475	1.09825	1.21402	1.27322	1.30844
JGEKS	1.09766	1.09358	1.09539	1.21404	1.26845	1.28782
CCDI	1.09786	1.09467	1.09810	1.21403	1.2739	1.30823
Lehr	1.08685	1.08296	1.08447	1.20766	1.25881	1.30362

Tab.5. Values of selected multilateral indices for considered time intervals – **Case 2.3**

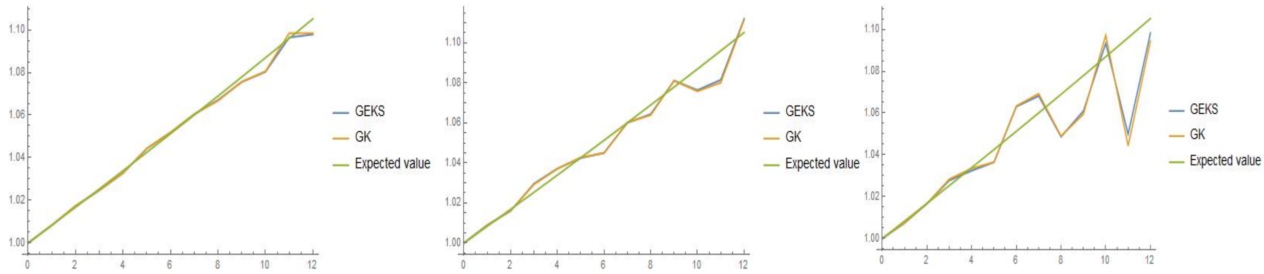
Index value	Time interval [0,12] (theoretical index value = 1.10517)			Time interval [0,24] (theoretical index value = 1.2214)		
	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$	$\beta = 0.02$	$\beta = 0.05$	$\beta = 0.1$
GK	1.10087	1.10428	1.09444	1.22130	1.23098	1.32131
TS	1.10173	1.10570	1.09450	1.22237	1.25137	1.30300
EW	1.10194	1.10632	1.09556	1.22226	1.25209	1.31908
EX	1.10091	1.10428	1.09345	1.22148	1.23256	1.30853
Real Time	1.10087	1.10428	1.09444	1.22130	1.23098	1.32131
GEKS	1.10232	1.10570	1.09825	1.22118	1.24958	1.30844
JGEKS	1.10205	1.10432	1.09539	1.22059	1.24529	1.28782
CCDI	1.10229	1.10561	1.09810	1.22107	1.24885	1.30823
Lehr	1.09109	1.09422	1.08447	1.22743	1.23190	1.30362

Fig. 15. Comparison of values of the Geary-Khamis and GEKS indices with the theoretical price dynamics ($t \in [0,12]$)

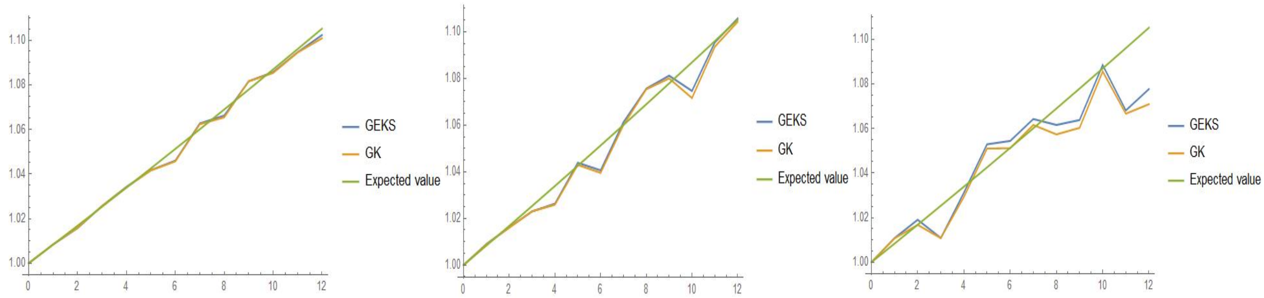
Case 2.1. $\beta = 0.02$ $\beta = 0.05$ $\beta = 0.1$



Case 2.2. $\beta = 0.02$ $\beta = 0.05$ $\beta = 0.1$



Case 2.3. $\beta = 0.02$ $\beta = 0.05$ $\beta = 0.1$



The level of the "average error" of the given multilateral price index $P^{0,t}$ in the time interval $[0,T]=[0,24]$ is measured by the root mean square deviation:

$$RMSD = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(P^{0,t} - \exp(\frac{\alpha \cdot t}{12}))^2}\ , \qquad ()$$

which describes the level of matching of obtained index values to theoretical values in points $1,2,...,T$ (under the GBM price model). The values of *RMSDs* calculated for all considered indices in Cases 2.1 – 2.3 and for two sample values of β are presented in Tab. 6.

Tab. 6. Values of *RMSDs* calculated for considered multilateral indices in Cases 2.1 – 2.3 and for $\beta \in \{0.02;\ 0.1\}$

Index formula	**Case 2.1**		**Case 2.2**		**Case 2.3**	
	$\beta = 0.02$	$\beta = 0.1$	$\beta = 0.02$	$\beta = 0.1$	$\beta = 0.02$	$\beta = 0.1$
GEKS MS	0.00334	0.02024	0.00821	0.05531	0.00675	0.04122
GEKS WS	0.00334	0.02023	0.00808	0.05468	0.00664	0.04045
GEKS HS	0.00334	0.02025	0.00816	0.05575	0.00671	0.04093
GEKS GMS	0.00334	0.02025	0.00815	0.05534	0.00673	0.04086
JGEKS MS	0.00343	0.02213	0.00825	0.04564	0.00669	0.04321
JGEKS WS	0.00343	0.02212	0.00812	0.04512	0.00658	0.04240
JGEKS HS	0.00343	0.02215	0.00820	0.04599	0.00665	0.04291
JGEKS GMS	0.00343	0.02215	0.00819	0.04564	0.00667	0.04283
CCDI MS	0.00334	0.02020	0.00822	0.05581	0.00675	0.04132
CCDI WS	0.00334	0.02019	0.00809	0.05527	0.00664	0.04057
CCDI HS	0.00334	0.02021	0.00817	0.05633	0.00671	0.04102
CCDI GMS	0.00334	0.02021	0.00816	0.05591	0.00673	0.04095
Lehr MS	0.00335	0.02024	0.01132	0.05747	0.00799	0.04685
Lehr WS	0.00335	0.02021	0.01123	0.05635	0.00795	0.04013
Lehr HS	0.00335	0.02025	0.01125	0.05725	0.00799	0.04397
Lehr GMS	0.00335	0.02024	0.01123	0.05709	0.00796	0.04356
GK MS	0.00334	0.02021	0.00832	0.05583	0.00665	0.04142
GK WS	0.00334	0.02017	0.00808	0.05426	0.00624	0.04057
GK HS	0.00334	0.02022	0.00827	0.05630	0.00661	0.04133
GK GMS	0.00334	0.02021	0.00826	0.05631	0.00663	0.04125
Real Time	0.00334	0.02023	0.00897	0.05712	0.00643	0.03807

5. Empirical study

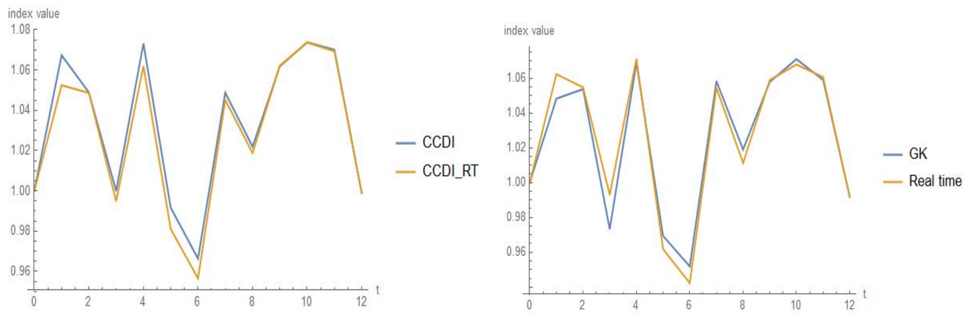
Poland is at the beginning of the way to the regular and official use of scanner data in the CPI measurement. Statistics Poland has started to cooperate with three supermarkets but they do not provide scanner data in a regular way. Moreover, there is no IT system for combining and analysing different data sources from different retailers (supermarkets) written in different file formats. Nevertheless, some experiments on real scanner data sets are being done by using the R package and Mathematica software. In the following empirical study, we consider two scanner data sources: (a) the first is "classical", i.e. data sets come from one supermarket and they concern the following group of products: plain flour (COICOP group: 011121), milk 3.2% (COICOP group: 011411) and rice (COICOP group: 011111). In this case, we have only a 13-month time series (Dec. 2014 – Dec. 2015), so our analysis is limited here; (b) the other scanner data source is *allegro.pl,* which is one of the biggest online e-commerce platform in Poland. We use transaction data on mountain bikes, touring bicycles and children's bicycles from the group "bicycles" (COICOP group: 071301). This time, the length of the considered time interval is 25 months (Dec. 2016 – Dec. 2018), and thus window updating methods (for a 13-month window) can be used here. In both cases (a) and (b), we use data aggregated to one month and products are defined by using EAN codes and retailers' internal product codes (only in the "a" case). EANs that share the same characteristics are combined into the same homogeneous group of products. Matching products to the proper group is supported by using some text mining methods and also some manual verification is made to avoid the "re-launch problem". To be included in the calculations, a product has to have a turnover above a minimum threshold. Products that show extreme pricing changes from one month to another are also excluded from the sample (outlier filter), i.e. we exclude 5% of the most extreme price changes. Our results are as follows:

Case A (data from a supermarket)

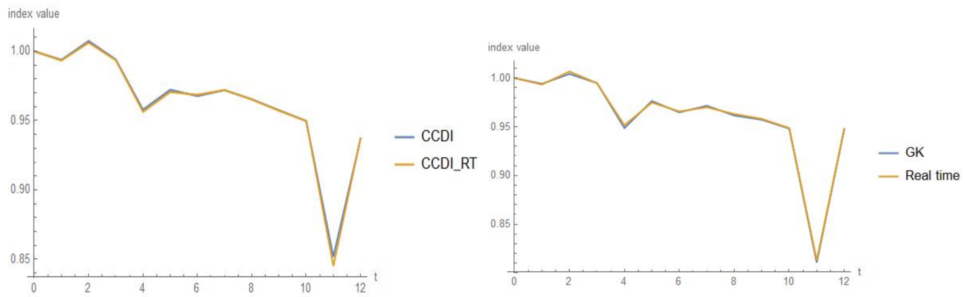
Fig. 16 presents a comparison of two selected multilateral indices calculated over the whole period of 13 months (i.e. the **CCDI** and **GK** indices when a full window is available) with the corresponding indices calculated over the "currently" available window (i.e. for the current time moment *t*, the available time window is $[0,t]$ – see the **CCDI_RT** and the **real time** indices). Fig. 17 presents a comparison of the **GEKS** index with the **CCDI** and **JGEKS** indices calculated over the whole period of 13 months. Fig. 18 presents all considered multilateral indices together with the **chained Jevons** index calculated for the fully available time window.

Fig. 16. Comparison of selected multilateral indices (CCDI, GK) for fully and "currently" available time windows (calculated for plain flour, milk and rice).

a) plain flour



b) milk



c) rice

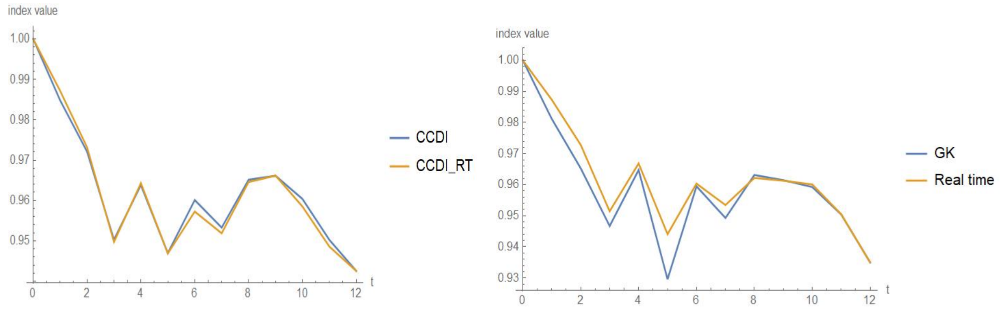
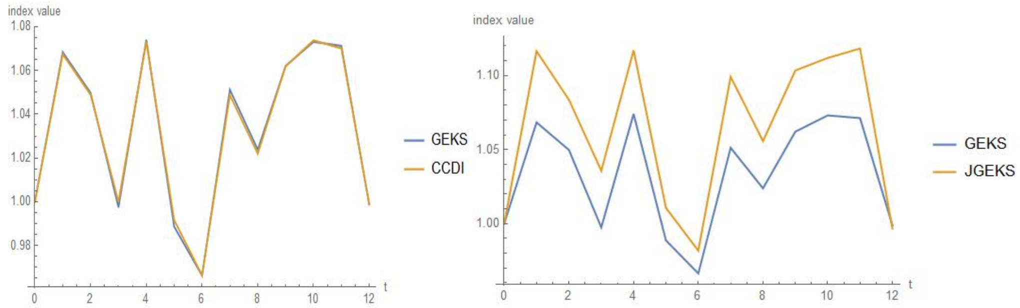
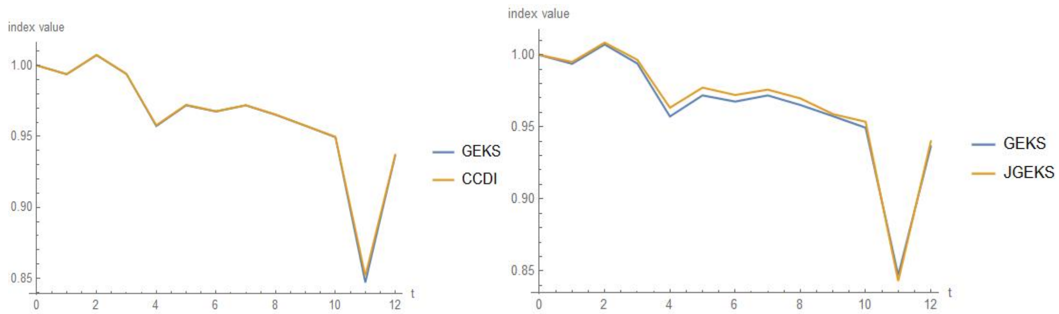


Fig. 17. Comparison of the GEKS index with the CCDI and JGEKS indices calculated over the whole period of 13 months for plain flour, milk and rice.

a) plain flour



b) milk



c) rice

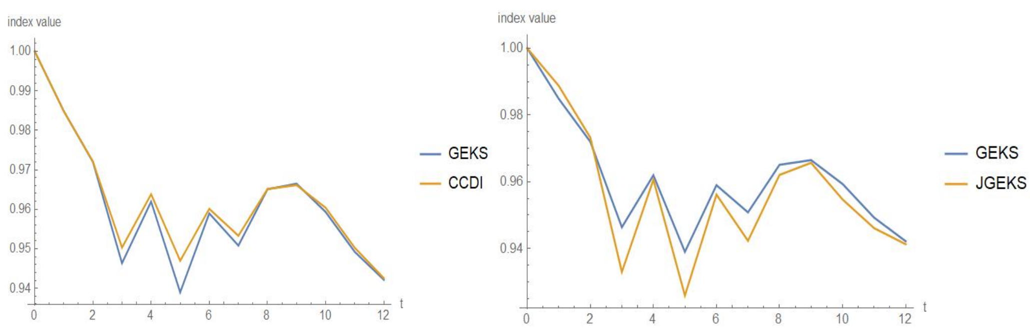
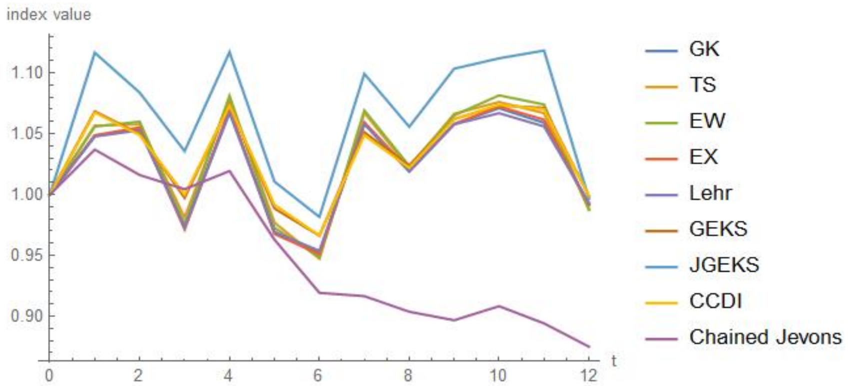
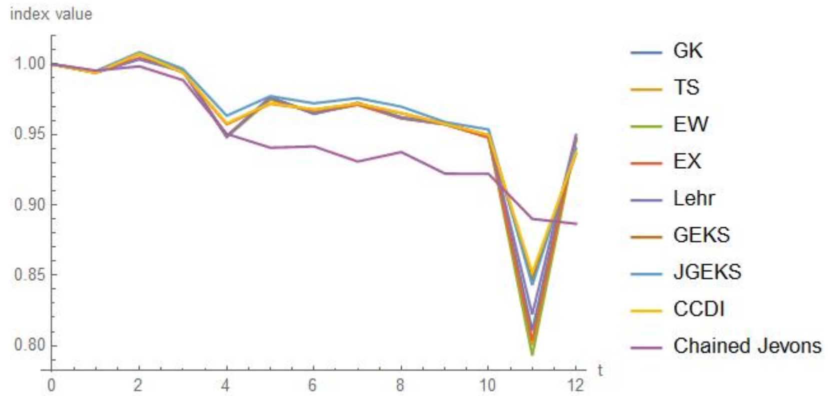


Fig. 18. All considered multilateral indices together with the chained Jevons index calculated over the whole period of 13 months for plain flour, milk and rice.

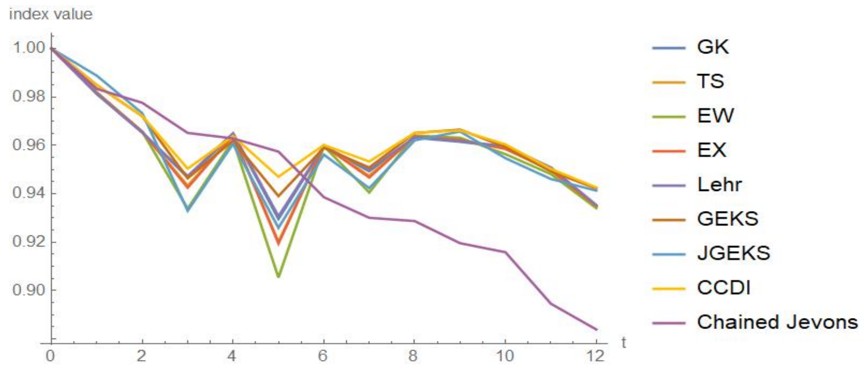
a) plain flour



b) milk



c) rice

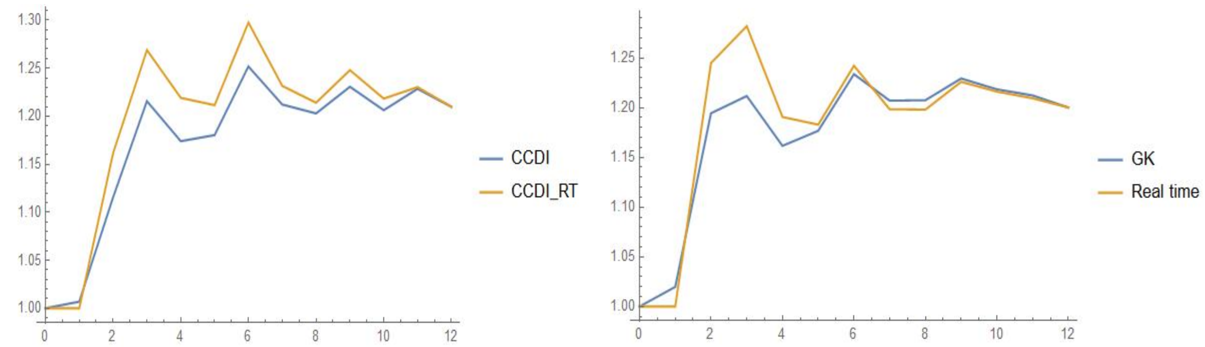


Case B (data from allegro.pl)

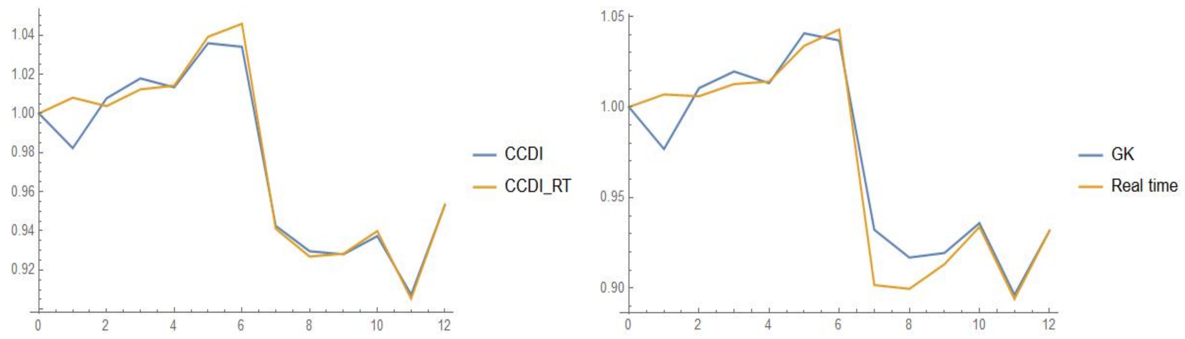
Fig. 19 presents a comparison of two selected multilateral indices calculated over the whole period of 13 months for the year: 2018 (i.e. the **CCDI** and **GK** indices when a full window is available) with the corresponding indices calculated over the "currently" available window (i.e. for the current time moment t, the available time window is $[0,t]$ – see the **CCDI_RT** and the **real time** indices). Fig. 20 presents a comparison of the **GEKS** index with the **CCDI** and **JGEKS** indices calculated over the whole period of 13 months for the year 2018 (a full time window is available). Fig. 21 shows differences in the window updating methods used in the case of the CCDI, GEKS, JGEKS and Lehr indices (T=12 and thus the splicing indices are calculated for the year 2018). Fig. 22 presents differences between the GEKS index and the corresponding splice indices, i.e. differences between the GEKS index calculated over the whole time window of 25 months ("GEKS Full": Dec. 2016 – Dec. 2018) and the GEKS index updated after the Dec. 2017 (a 13-month time window) by using the movement splice, the window splice, the half splice, and the mean splice methods (the chain drift effect is tested). Fig. 23 presents a comparison of weighting schemes in the QU method in two variants: (A) the comparison among the GK, TS, EW and EX indices; (B) differences between the GK index and the TS, EW, EX indices (a 13-month time window is considered, year: 2018). Fig. 24 presents a comparison of all discussed multilateral indices with the **chained Jevons** index based on data collected in a traditional way by Statistics Poland (it is denoted by **JEV (SP))** and calculated for all groups of bicycles sold in 2018.

Fig. 19. Comparison of selected multilateral indices (CCDI, GK) for fully and "currently" available time windows (for mountain bikes, touring bicycles and children's bicycles sold in 2018).

a) mountain bikes



b) touring bicycles



c) children's bicycles

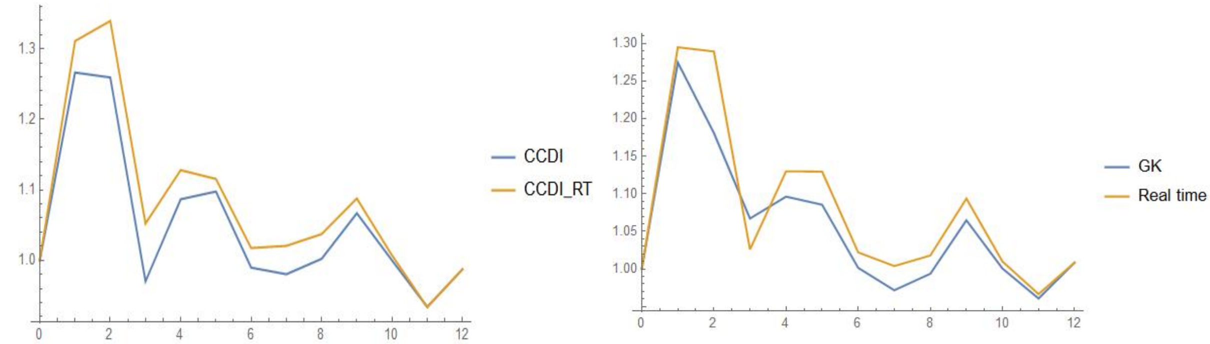
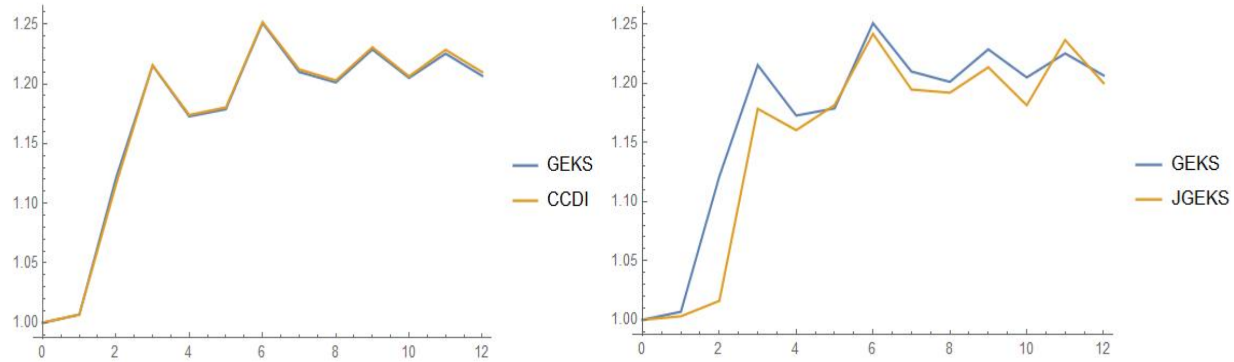
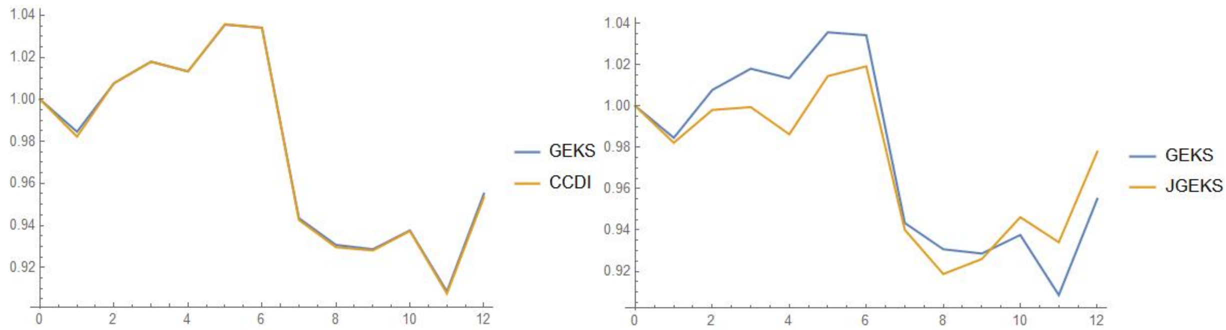


Fig. 20. Comparison of the GEKS index with the CCDI and JGEKS indices (a full window of 13 months is available) for mountain bikes, touring bicycles and children's bicycles sold in 2018.

a) mountain bikes



b) touring bicycles



c) children's bicycles

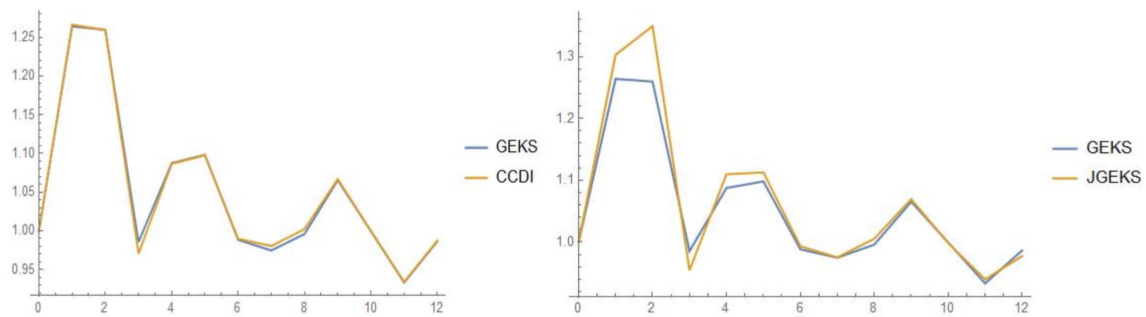
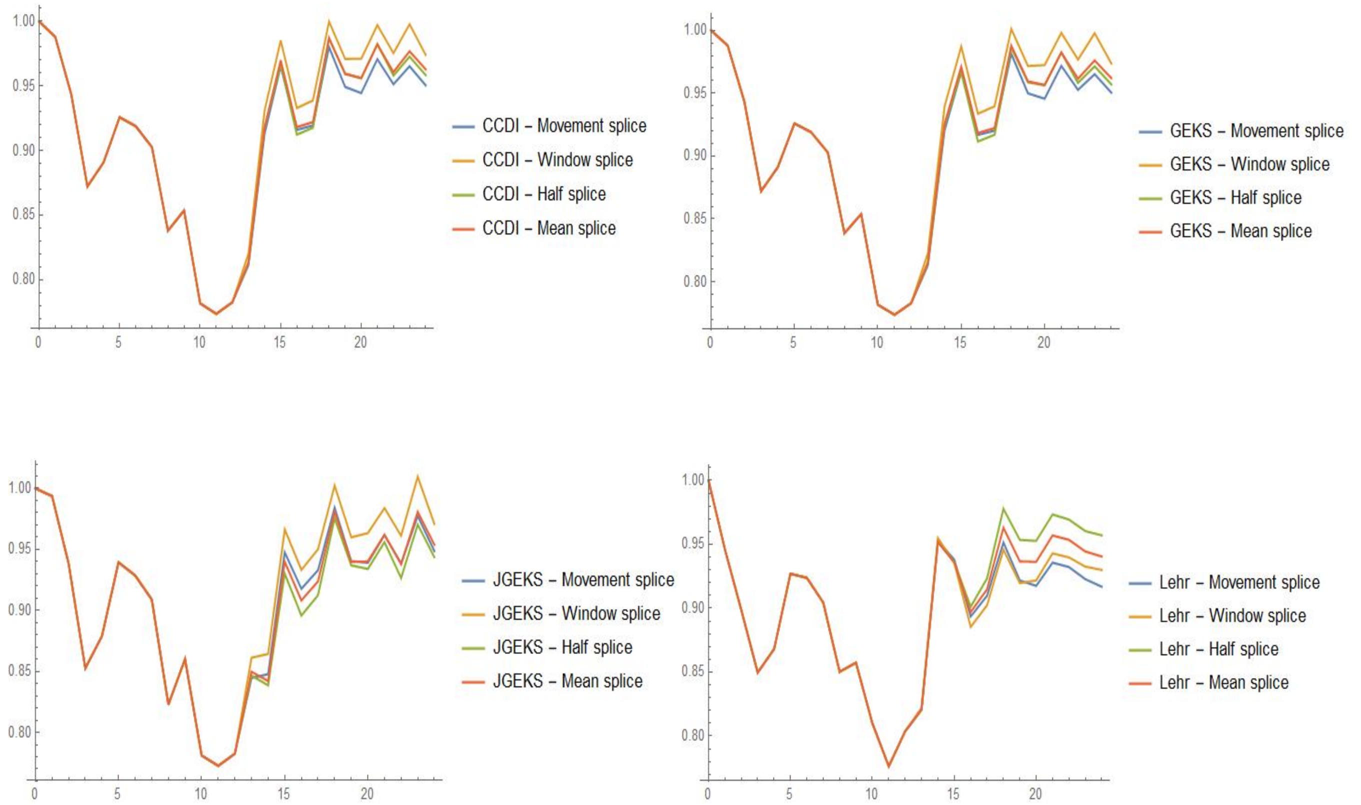
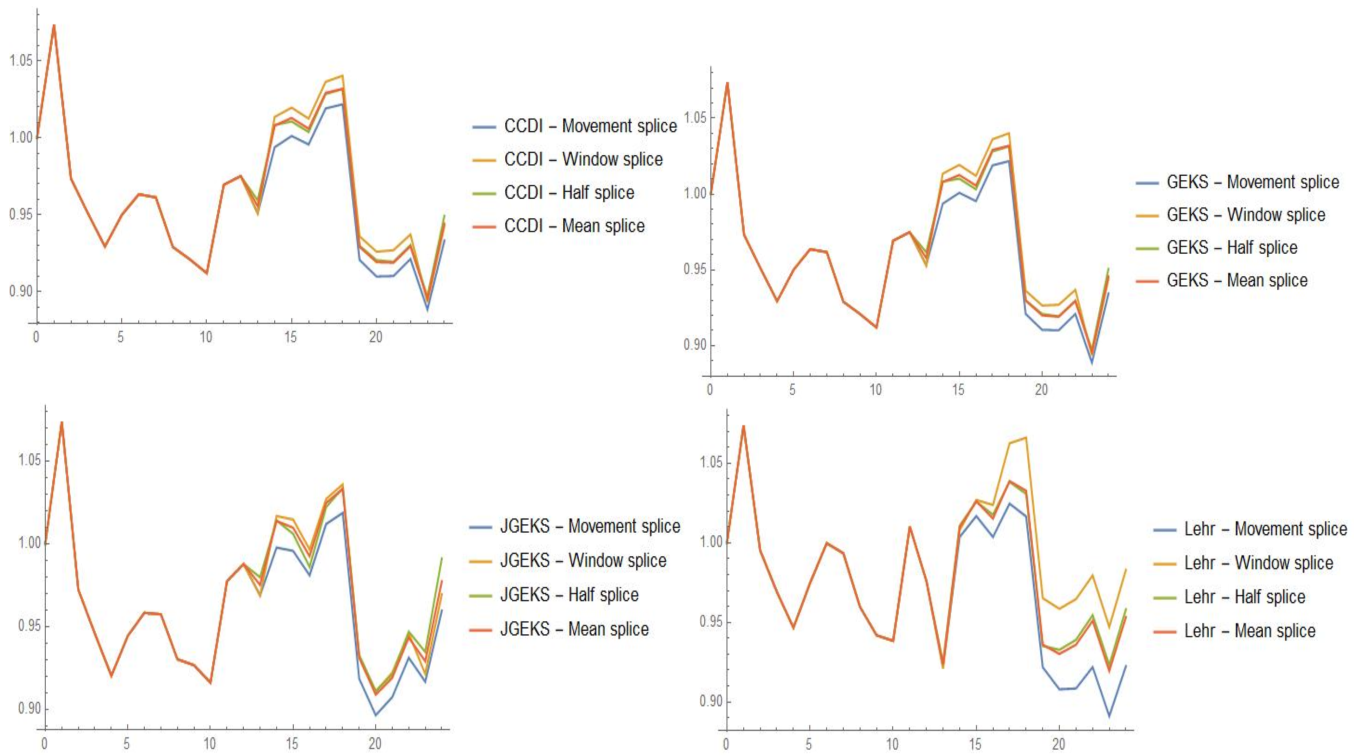


Fig. 21. Window updating methods in the case of the CCDI, GEKS, JGEKS and Lehr indices (a 13-month time window is considered)

a) mountain bikes



b) touring bicycles



c) children's bicycles

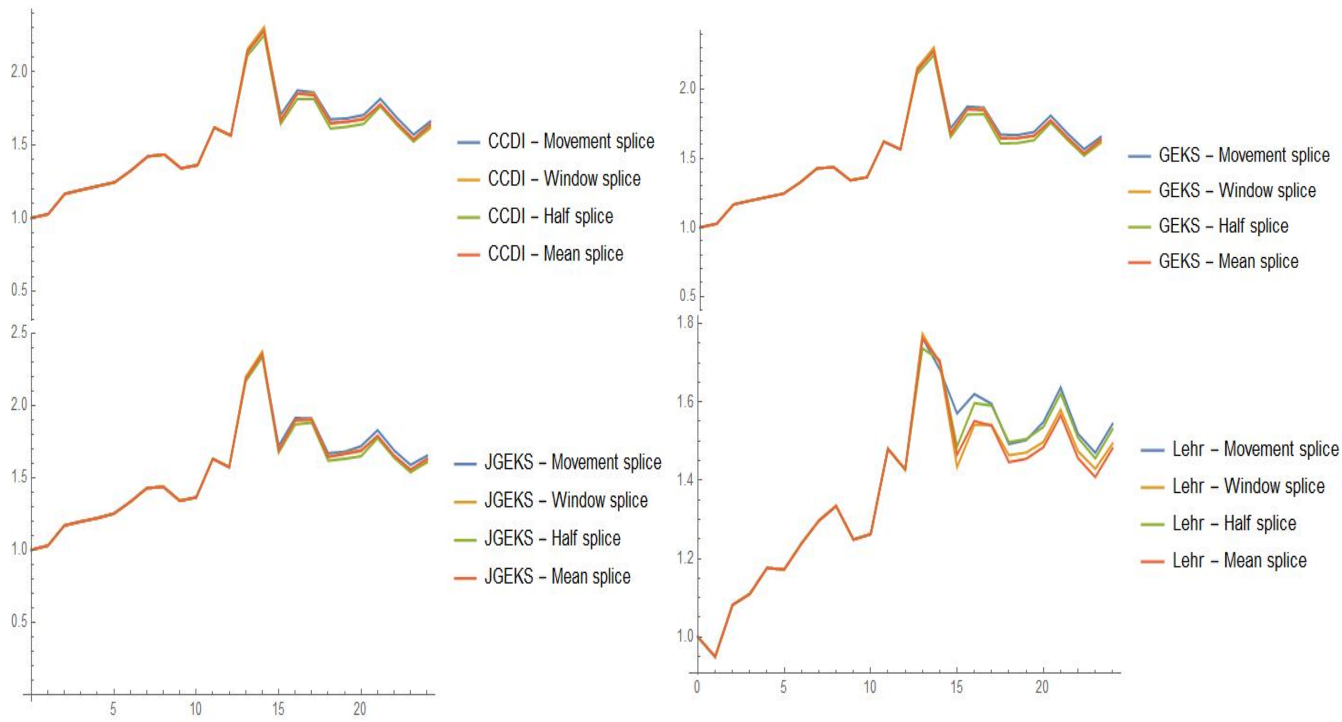
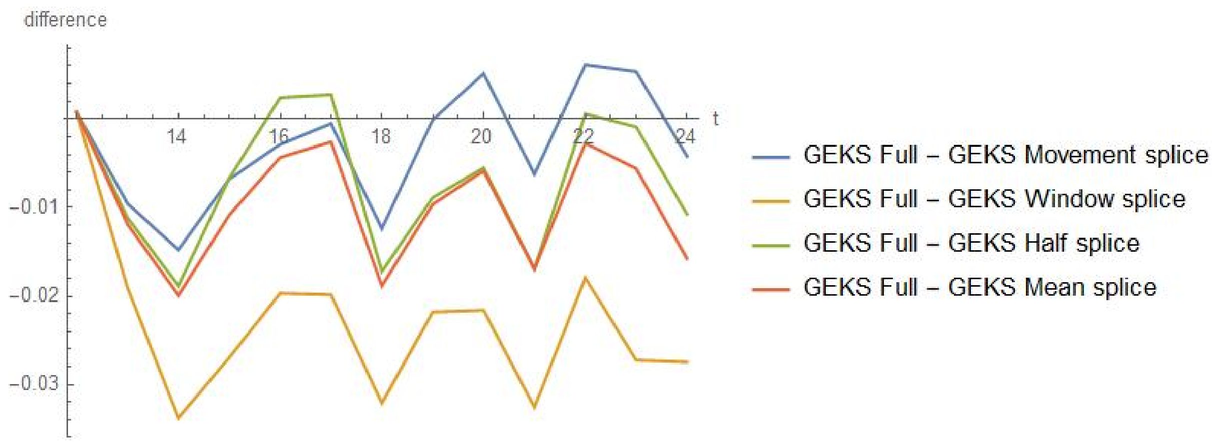
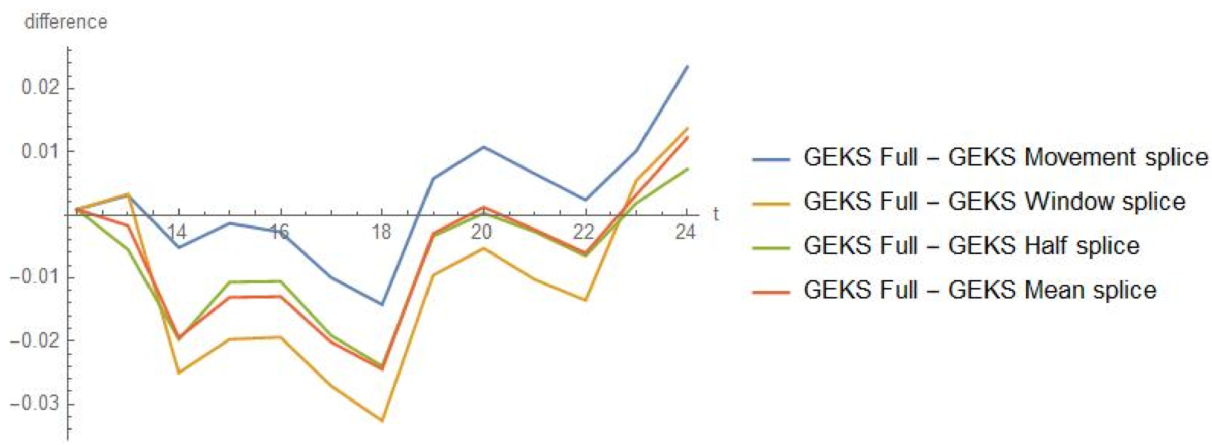


Fig. 22. Differences between the GEKS index and the corresponding splice indices (year: 2018)

a) mountain bikes



b) touring bicycles



c) children's bicycles

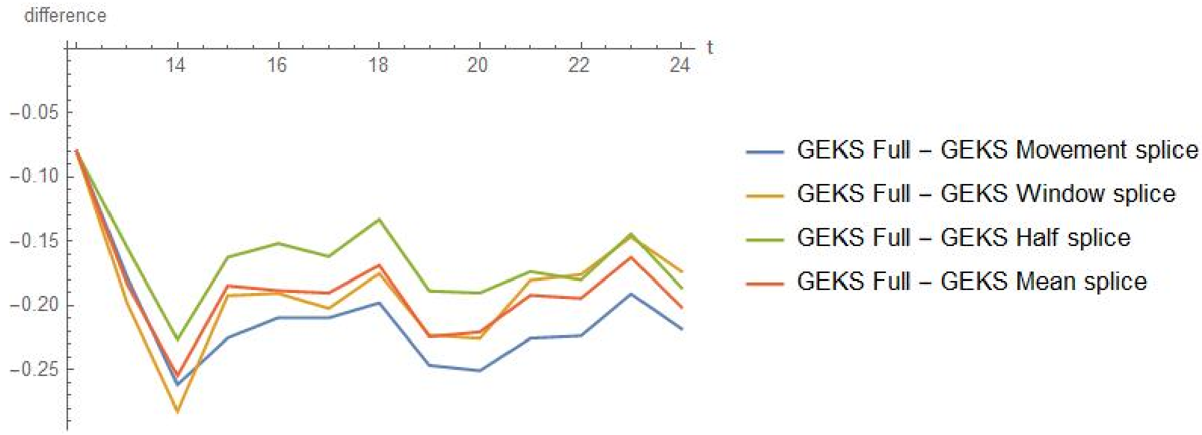
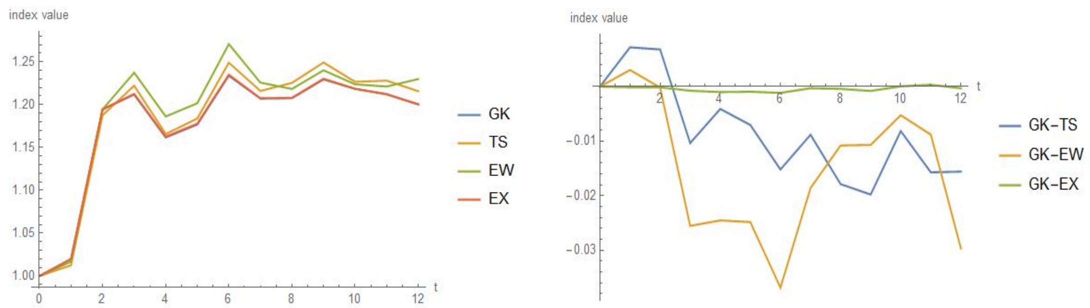
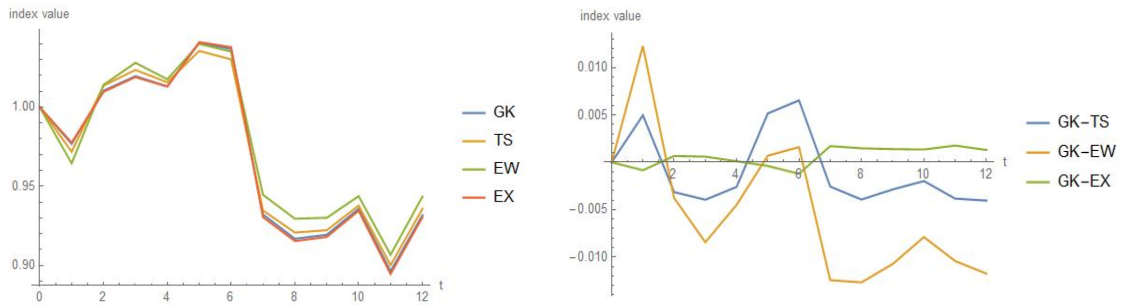


Fig. 23. Comparison of weighting schemes in the QU method
(year: 2018, a 13-month time window is used)

a) mountain bikes



b) touring bicycles



c) children's bicycles

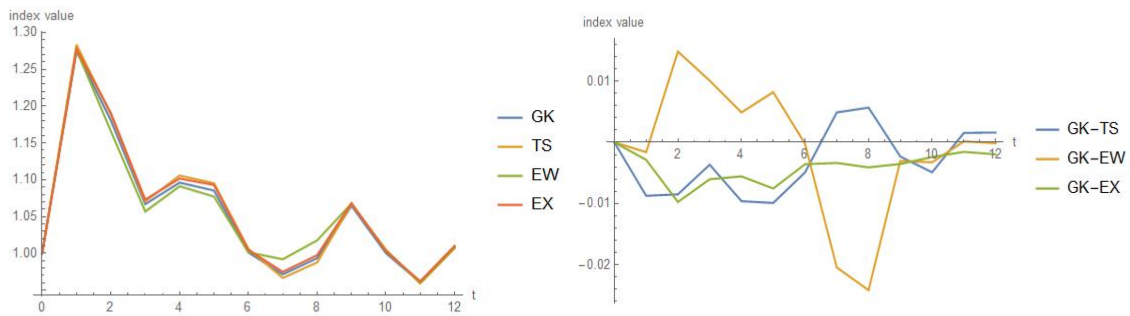
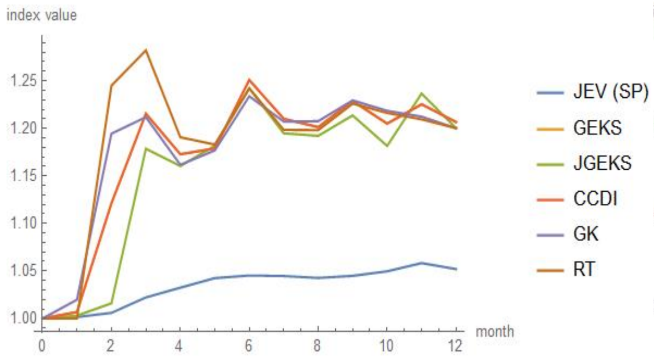
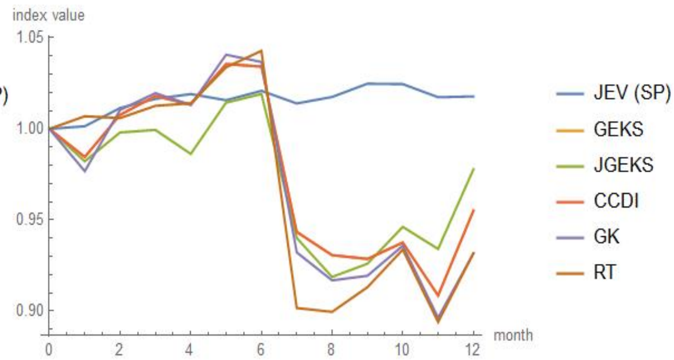


Fig. 24. Comparison of all discussed multilateral indices with the **chained Jevons** based on data collected in a traditional way by Statistics Poland (a full window of 13 months, 2018 year).

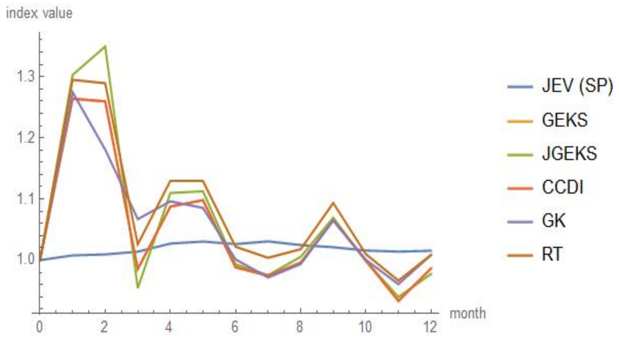
a) mountain bikes



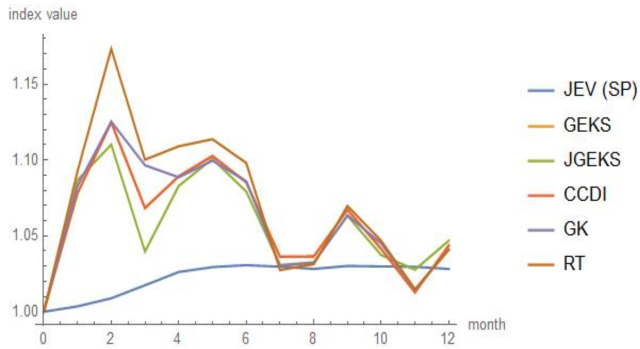
b) touring bicycles



c) children's bicycles



d) all groups of bicycles



6. Conclusions

The general conclusions from the **Simulation Study** are: (a) Even multilateral indices may differ from unity if only prices revert back to their levels in the base period. In this sense, they may suffer from chain drift. In our study, chained superlative indices and multilateral indices seem to slightly overestimate the real price change when quantities strongly decrease. When quantities strongly decrease, as a rule chained superlative indices and multilateral indices seem to be slightly below the real price change. In our experiments (which are not presented here), we observe that the monotonicity of quantities (in particular those connected with new and disappearing goods) has a much more bigger impact on differences among multilateral indices than the level of price volatilities. However, when both prices and quantities in the current period revert back to their levels in the base period, multilateral indices indicate that no price change occurred; (b) When prices change periodically and quantities are correlated with price movements (positively or negatively) and there are no temporary unavailable products, then differences between multilateral indices are negligible. In the same case, when quantities start to strongly decrease or increase, there may appear some differences among multilateral indices. These differences will rise if temporary unavailable products are included in the sample (see also Empirical Study); (c) When prices can be described by a geometric Brownian motion (they follow a trend instead of displaying periodical changes), differences between the theoretical (known) value of the price change and any multilateral price index (for a given time moment) are the biggest in the case of strongly decreasing quantities. In general, these differences will rise if the volatility of prices increases. In particular, for any considered cases of quantity changes, the measured root mean square error seems to be comparable for considered multilateral indices (GEKS, JGEKS, CCDI, Lehr, Real time) but it seems to be the smallest in the case of window splice method.

Our **Empirical Study** provides the following conclusions: (a) When we have no historical data from supermarkets and we start using scanner data sets, then the application of multilateral indices for the "currently" available time window (from the beginning of cooperation with supermarkets till the current month) is justified since differences between selected indices (CCDI, GK) for the fully and "currently" available time window are not too big, i.e. these differences are decreasing functions of time and, as a rule, after 6 – 8 months they are negligible. Nevertheless, in the case of very dynamic scanner data sets (such as those connected with bicycles, where there are many new and disappearing bike models during a year – see Case B in the Simulation Study), these differences may equal several percentage points (b) In practice, there are no substantial differences between the GEKS and CCDI indices and it is not surprising since superlative indices (Fisher, Törnqvist) approximate each other (Diewert (1976)). Nevertheless, the differences between the GEKS and JGEKS indices are crucial and, in our opinion, it confirms that the movements of quantities may not be (rationally) correlated with price movements; (c) Differences between multilateral indices and the chained Jevons index may be very big (see Fig. 18 for plain flour or rice), and as a rule they are. Thus, switching the chained Jevons index to one of multilateral indices does matter in the CPI measurement; (d) The chain drift bias may be substantial when using splice indices (in our case, it has the biggest value for *mountain bikes* and *touring bicycles*, where it exceeds 3 percentage points). In our study, the best result (i.e. the smallest chain drift effect) as a rule is obtained by using the movement splice method and the worst result (the biggest chain drift bias) is obtained by using the window splice method. The mean splice method provides index values being somewhere in the middle, i.e. in our study, such obtained values are between the values obtained with the use of the movement splice method and the window splice method; (e) The differences between the GEKS index and splice indices as a rule are negative; (f) The choice of the weighting schemes in the QU method does matter – differences in results may be crucial (in our study time moments for which the differences between the TS, EW and EX indices exceeded 3 percentage points were observed). The EW index differs the most in relation to the Geary-Khamis index; (g) The results of price dynamics obtained by using alternative data sources (e.g.: allegro.pl) may be completely different in comparison to those obtained by using traditionally collected data sets (see Fig. 23); (h) The Lehr price index seems to be the most sensitive in the case of the choice of window updating method (see also our Simulation Study).

References

(ABS) Australian Bureau of Statistics. (2016). Making Greater Use of Transactions Data to Compile the Consumer Price Index, Information Paper 6401.0.60.003, 29 November 2016, Canberra.

Carli, G. (1804). Del valore e della proporzione de'metalli monetati. In: Scrittori Classici Italiani di Economia Politica, 13, 297-336.

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 Ottawa Group Meeting, 20-22 May 2015, Urayasu City, Japan.

Chessa, A.G. (2016), "A New Methodology for Processing Scanner Data in the Dutch CPI", Eurona 1/2016, 49-69.

Chessa, A.G., and Griffioen, R. (2016). Comparing Scanner Data and Web Scraped Data for Consumer Price Indices. Report, Statistics Netherlands.

Chessa, A.G. (2017). Comparisons of QU-GK Indices for Different Lengths of the Time Window and Updating Methods. Paper prepared for the second meeting on multilateral methods organised by Eurostat, Luxembourg, 14-15 March 2017. Statistics Netherlands.

Chessa, A.G., Verburg, J., and Willenborg, L. (2017). A comparison of price index methods for scanner data. Paper presented at the 15th Meeting of the Ottawa Group on Price Indices, Eltville am Rhein, Germany, 10-12 May 2017

Consumer Price Index Manual. Theory and practice. (2004). ILO/IMF/OECD/UNECE/Eurostat/The World Bank, International Labour Office (ILO), Geneva.

Diewert, W. E., 1976. Exact and superlative index numbers. Journal of Econometrics 4, 114-145.

Diewert, W. E. (2012). Consumer price statistics in the UK. Office for National Statistics, Newport.

Diewert, W.E., and Fox, K.J. (2017). Substitution Bias in Multilateral Methods for CPI Construction using Scanner Data. Discussion paper 17-02, Vancouver School of Economics, The University of British Columbia, Vancouver, Canada.

de Haan, J. (2006). The re-design of the Dutch CPI. Statistical Journal of the United Nations Economic Commission for Europe, 23, 101-118.

de Haan, J., van der Grient, H.A. (2011). Eliminating chain drift in price indexes based on scanner data. Journal of Econometrics, 161, 36-46.
de Haan, J., and Krsinich, F. (2014). Time Dummy Hedonic and Quality-Adjusted Unit Value Indices: Do They Really Differ? Paper presented at the Society for Economic Measurement Conference, 18-20 August 2014, Chicago, U.S.
de Haan J.(2015). A Framework for Large Scale Use of Scanner Data in the Dutch CPI. Paper presented at the 14th Ottawa Group meeting, Tokyo, Japan.
de Haan, J., Willenborg, L., and Chessa, A.G. (2016). An Overview of Price Index Methods for Scanner Data. Paper presented at the Meeting of the Group of Experts on Consumer Price Indices, 2-4 May 2016, Geneva, Switzerland.
Eltetö, Ö., and Köves, P. (1964), "On a Problem of Index Number Computation Relating to International Comparisons", (in Hungarian), Statisztikai Szemle 42, 507-518.
Eurostat (2017), Practical Guide for Processing Supermarket Scanner Data, [in:] Harmonised Index of Consumer Prices (available on website https://ec.europa.eu/eurostat/web/hicp/overview)
FEENSTRA, R. C., MA, H., and RAO, D. S. PRASADA. (2009). "Consistent comparisons of real incomes across time and space". Macroeconomic Dynamics, 13(S2), pp.169-193.
Fisher I. (1922), The Making of Index Numbers, Boston: Houghton Mifflin.
Geary, R.G. (1958), "A Note on Comparisons of Exchange Rates and Purchasing Power between Countries", Journal of the Royal Statistical Society Series A 121, 97-99.
Gini, C. (1931), On the Circular Test of Index Numbers, Metron 9:9, 3-24.
Griffioen, A.R., and ten Bosch, O. (2016). On the Use of Internet Data for the Dutch CPI. Paper presented at the UNECE-ILO Meeting of the Group of Experts on Consumer Price Indices, 2-4 May 2016, Geneva, Switzerland.
Hill, R.J. (2000), 'Measuring substitution bias in international comparisons based on additive purchasing power parity methods', European Economic Review, 44, pp. 145–162.
Inklaar, R. and W.E. Diewert (2016), "Measuring Industry Productivity and Cross-Country Convergence", Journal of Econometrics 191, 426-433.
Ivancic, L., Diewert, W.E., and Fox, K.J. (2011). Scanner Data, Time Aggregation and the Construction of Price Indices. Journal of Econometrics, 161 (1), 24-35.
Jevons, W.S. (1865). The variation of prices and the value of the currency since 1782. J. Statist. Soc. Lond., 28, 294-320.
Khamis, S.H. (1972), A New System of Index Numbers for National and International Purposes, Journal of the Royal Statistical Society Series A 135, 96-121.
Krsinich, F. (2014). The FEWS Index: Fixed Effects with a Window Splice – Non-Revisable Quality-Adjusted 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.
Lamboray C. (2017). The Geary Khamis index and the Lehr index: how much do they differ?. Paper presented at the 15th Ottawa Group meeting, 10-12 May 2017, Elville am Rhein, Germany.
Laspeyres E. (1871), Die Berechnung einer mittleren Waarenpreissteigerung, Jahrbücher für Nationalökonomie und Statistik, 16, 296—314.
Levell, P. (2015). Is the Carli index flawed?: assessing the case for new retail price index RPIJ. J. R. Statist. Soc. A, 178(2), 303-336.
Loon, K. V., Roels, D. (2018). Integrating big data in the Belgian CPI. Paper presented at the meeting of the group of experts on consumer price indices, 8-9 May 2018, Geneva, Switzerland.

Maddison, A., and Rao, D.S.P. (1996). A Generalized Approach to International Comparison of Agricultural Output and Productivity. Research memorandum GD-27, Groningen Growth and Development Centre, Groningen, The Netherlands.
Paasche H. (1874), Über die Preisentwicklung der letzten Jahre nach den Hamburger Borsennotirungen, Jahrbücher für Nationalökonomie und Statistik, 12, 168—178.
Summers, R. (1973). International Price Comparisons Based Upon Incomplete Data. Review of Income and Wealth, 19, 1-16.
Szulc, B. (1964). Indices for Multiregional Comparisons. (in Polish). Przeglad Statystyczny 3, 239-254.
Szulc, B. (1983). Linking Price Index Numbers, in: Price Level Measurement, W. E. Diewert and C. Montmarquette (eds.), 537 – 566.
Törnqvist, Leo. 1936. The Bank of Finland's Consumption Price Index, Bank of Finland Monthly Bulletin 10, 1-8.
Walsh C. M. (1901). The Measurement of General Exchange Value, The MacMillan Company:
New York.
Willenborg, L. (2010). Chain Indexes and Path Independence. Report, Statistics Netherlands.
Willenborg, L., and van der Loo, M. (2016). Transitivizing Price Index Numbers Using the Cycle Method: Some Empirical Results. Report, Statistics Netherlands.
Willenborg, L. (2017). Transitivizing Elementary Price Indexes for Internet Data using the Cycle Method. Discussion Paper, Statistics Netherlands
Van der Grient, H.A., de Haan, J. (2010). The use of supermarket scanner data in the Dutch CPI. Paper presented at the Joint ECE/ILO Workshop on Scanner Data, 10 May 2010, Geneva.
Von Auer L. (2017). Processing scanner data by an augmented GUV index. Eurostat Review of National Accounts and Macroeconomic Indicators, 1/2017, 73-91.
Von Auer L. (2019). The Nature of Chain Drift. Paper presented at the 17th Meeting of the Ottawa Group on Price Indices, 8 – 10 May 2019, Rio de Janerio, Brasil.