Discussion Paper

From GEKS to cycle method

The views expressed in this paper are those of the author(s) and do not necessarily reflect the policies of Statistics Netherlands

2017 | 11

Leon Willenborg

Content

Summary

The paper starts with the standard GEKS method, that is used to transitivize nontransitive price indices. Then several generalizations of the standard method are investigated. These generalizations imply that the GEKS method can be used in more cases than in which the standard case can be applied. For instance, the standard GEKS method requires that for each pair of months in the reference period a price should be known. It is shown that this condition can be relaxed if the definition of the GEKS is slightly modified. The standard GEKS method uses geometric averaging. In the paper it is shown that other averaging methods can be used to obtain similar averaged results. Another generalization is to use different weights for the various months in the reference period. This result implies a local form of GEKS, in which only the prices of a neighbourhood of months of a reference month are used instead of the entire reference window. A local GEKS method is thought to be more attractive and flexible than a global one, such as the standard method. GEKS is a bit of a trick that works to produce transitive price indices. But there is no control over the goodness of fit. This would require the use of some kind of metric. In the paper it is shown that it is possible to give such a formulation, however. In fact this leads to the cycle method, another transitivizing method, proposed earlier by the present author.

Keywords

Transitive price indices, transitivizing method, transitivation, averaging, GEKS method, cycle method.

1. Introduction

The aim of this paper is to consider several generalizations of the well-known GEKS method. This method is used to produce a transitive price index from a nontransitive one. The GEKS method was proposed by several individuals, starting with Gini from Italy, whose suggestion apparently was subsequently forgotten until it was rediscovered (much later) by Köves and Eltetö from Hungary and Sczulc from Poland. The GEKS method is named after these four individuals.

In the recent comparative study of several competing price index methods (cf. De Haan et. al (2016) and Chessa et al. (2016)) also the GEKS method was used. It was not immediately clear how to implement this method in Stata. The author studied this problem, and proposed the method in Excel that is explained in Section 2. When writing up his method, he found that certain of the constraints used can in fact be removed. They led to various variants that seem all to be useful in practice. The current document records these results. This exploration also showed that the GEKS method can be generalized to the cycle method. In the author's opinion this method is superior to GEKS, but not as simple to apply as the GEKS method (which can be applied easily in Excel, as the present report shows). But with special software (in R, for instance[1]), it should also be within reach of interested researchers.

We start this paper with an exposition of the GEKS method. We also discuss how it can be computed using Excel. By considering the computational aspects the GEKS method (and the variants to be discussed) become more tangible. They also allow to actually compute the adjustments, rather than to fantasize on them when reading the descriptions of the methods.

We state various conditions that are necessary to be able to apply the method, or several choices that have been made in defining the method. The variants we discuss later in the paper are all inspired by relaxing one or several of these conditions or properties, but keeping the main ideas. We thus obtain a number of GEKS-like methods.

There exist other methods to transitivize a price index. We consider one of these to contrast with the GEKS method, namely the cycle method (CM).[2] We do not give a full exposition of the cycle method here; only a brief description, to 'prepare the mind'. It is intended as a gentle introduction to the cycle method for those coming from price index theory. The interested reader is referred to Willenborg (2017) for an extensive discussion of the cycle method. An earlier reference is Willenborg (1993)[3], but this may not be so easy for price index specialists, who are probably not familiar with land surveying, which is the original inspiration for the present author (in a time when he was completely ignorant of price index number theory).

[1] The present author has been developing such software in R. The key step - calculating a cycle matrix - can be done in a surprisingly short script.

[2] It would be possible to consider another one based on optimal spanning trees (proposed by Hill). But as Willenborg (2016) tries to show, this method is not superior to the cycle method. It therefore makes no sense to consider this method here as well, as a contrast to the GEKS method.

[3] Uploaded to ResearchGate "the social networking site for scientists and researchers to share papers, ask and answer questions and find collaborators" (description from en.Wikipedia.org)

The paper is organized as follows. In Section 2 the standard GEKS method is discussed. It is shown how the definition of the method leads to formulations of the method that are attractive computationally (it is easy to apply in a spreadsheet like Excel, for instance) on the one hand and suitable for generalizations on the other. Numerical examples are given to demonstrate the working of the GEKS method.

The standard GEKS method gives all indices that are used equal weight. This is not always attractive. One would perhaps like to give an index with base month and reporting month far apart a lesser weight than one where these months are close, say adjacent. In Section 3 such an extension of GEKS is discussed. In fact the section starts with a solution for another problem with GEKS, which is that the indices for all pairs of months should be provided. It is shown how one can easily side-step this restriction by suitably modifying the standard GEKS method. Also in that section it is shown that there is no need to used geometric averaging as in standard GEKS. With suitably chosen (small) adaptations one can use other averaging methods just as well. Some numerical examples are given to illustrate this point. And also that the calculations are still easy to perform in a spreadsheet program such as Excel. Furthermore in that section it is shown how one can deal with situation where not all price indices are known; some are missing, for whatever reason. This situation may apply in the spatial case, or the spatio-temporal case. This is discussed in a separate subsection, showing that GEKS (and the variants provided in the present paper) can be applied in a temporal, spatial or spatio-temporal setting.

In Section 4 updating is briefly discussed for the GEKS method through an example, that should convey the underlying ideas so that it should be possible to apply the method in a general case. The ideas used are similar to those used earlier for the cycle method (cf. Willenborg, 2015). Updating is also a topic discussed more extensively in Willenborg (2017).

To the present author, before he started writing the present paper, GEKS was a trick to produce transitive indices. But why it actually worked was a bit of a mystery (to him). By writing the present paper, his insight into this method has increased substantially. Variants and generalizations more or less suggested themselves. But also the feeling that the method would need a solid basis grew while writing the paper. What he was looking for was some kind of model, in particular an optimization model, from which the GEKS method would follow. In fact, while writing the report it became apparent that the formulation of the standard GEKS method given in Section 2 can be used for such a foundation. In the final section, Section 5, optimization models are given that generalize a certain property of the standard GEKS method. This new method is not precisely the same as standard GEKS but is a straightforward generalization (in fact yielding certain loglinear models). Also it became apparent, this optimization model is a step in the direction of the cycle method. It is shown how the cycle method generalizes the optimization models derived from the (standard) GEKS method.[4]

[4] The present document was reviewed by Sander Scholtus.

2. Standard GEKS method

2.1 Preliminary remarks

In the present document the GEKS method is the basic method that is considered. The variants to be discussed later use some of the basic aspects of the method, but discard certain others. So the variants can be considered different generalizations of the GEKS method.

Not only the method is discussed in the present section but also its computation in Excel.[5]

2.2 Description

In this section we give a general description of the GEKS method, geared at its computation. In the remainder of this section we present concrete examples, showing how the GEKS index can be very easily calculated in Excel.

We assume that a full price index matrix (PIM)

$$P = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & \ddots & \vdots \\ p_{n1} & \cdots & p_{nn} \end{pmatrix} \tag{2.1}$$

is given for a period $W = \{1, \dots, n\}$, where p_{ij} is the price index with base month i and reporting month j. P is supposed not to have missing entries. Furthermore we assume that:

- P satisfies the property that $p_{ji}p_{ij} = 1$, for all $i, j \in \{1, \dots, n\}$.
- $p_{ii} = 1$ for all $i \in \{1, \dots, n\}$.
- The price index is nontransitive, that is, there are $i, j, k \in W$ such that $p_{ij}p_{jk} \neq p_{ik}$.

The latter requirement is only to give the method a purpose. But it works when the input is a transitive matrix (and the adjusted PIM is equal to the original one).

The output of the GEKS adjustment process is a PIM denoted by P^G that is transitive. That is, if the PIM is given by

$$P^G = \begin{pmatrix} p_{11}^G & \cdots & p_{1n}^G \\ \vdots & \ddots & \vdots \\ p_{n1}^G & \cdots & p_{nn}^G \end{pmatrix}, \tag{2.2}$$

then $P_{ij}^G P_{ik}^G = P_{ik}^G$ for all $i, j, k \in W$. P^G is obtained as follows:

- Calculate the geometric mean for each column in P. Let the result for column i be denoted by π_i

[5] In R would also be another option. But Excel is probably more easily accessible.

– Using these averages one can calculate the elements of P as follows: $p_{ij}^G = \frac{\pi_j}{\pi_i}$, for all $i, j \in \{1, \ldots, n\}$.

It is immediately clear that $p_{ij}^G p_{ji}^G = 1$ for all $i, j \in W$, and that $p_{ii}^G = 1$ all $i \in \{1, \ldots n\}$. Transitivity follows from the special form of p_{ij}^G , viz as a ratio of two separate quantities: $p_{ij}^G p_{jk}^G = \frac{\pi_j}{\pi_i}\frac{\pi_k}{\pi_j} = \frac{\pi_k}{\pi_i} = p_{ik}^G$, for all $i, j, k \in W$.

These computations can easily be done in Excel, as the following example shows. The example is given in Excel not to promote the use of this package per se. It is only a very convenient as a tool to demonstrate small examples easily and conveniently (including the output).

Remark The GEKS estimator is often defined in another way, namely as follows:

$$p_{ij}^G = \prod_{k=1}^n \left(p_{ik} p_{kj}\right)^{\frac{1}{n}}, \tag{2.3}$$

which is the geometric average of 'indirect' price comparisons $p_{ik}p_{kj}$, rather than the direct ones p_{ij}. This definition tacitly assumes that the price indices for each pair (i, j) exist. We can rewrite (2.3) as follows:

$$p_{ij}^G = \prod_{k=1}^n \left(p_{ik} p_{kj}\right)^{\frac{1}{n}} = \prod_{k=1}^n \left(\frac{p_{kj}}{p_{ki}}\right)^{\frac{1}{n}} = \frac{\left(\prod_{k=1}^n p_{kj}\right)^{\frac{1}{n}}}{\left(\prod_{k=1}^n p_{ki}\right)^{\frac{1}{n}}} = \frac{\kappa_j}{\kappa_i}, \tag{2.4}$$

which is a ratio of the column averages κ_i and κ_j, or of price indices which have the same reporting month. Alternatively we can rewrite (2.3) as follows:

$$p_{ij}^G = \prod_{k=1}^n \left(p_{ik} p_{kj}\right)^{\frac{1}{n}} = \prod_{k=1}^n \left(\frac{p_{ik}}{p_{jk}}\right)^{\frac{1}{n}} = \frac{\left(\prod_{k=1}^n p_{ik}\right)^{\frac{1}{n}}}{\left(\prod_{k=1}^n p_{jk}\right)^{\frac{1}{n}}} = \frac{\rho_i}{\rho_j}, \tag{2.5}$$

which is a ratio of the row averages ρ_i and ρ_j, or of the price indices which have the same base month. Now in case of geometric averaging we have that

$$\kappa_i \rho_i = 1, \tag{2.6}$$

for $i \in \{1, \ldots n\}$. That is, the column and row (geometric) averages of the PIM, are reciprocal. This is a property of the geometric averages. In Section 3.5 we shall use other forms of averaging, which lack this property. But still a GEKS-like method can be defined.

Using (2.6) we can derive a third, symmetric, form for the standard GEKS price index, which is convenient in spreadsheet calculations:

$$p_{ij}^G = \rho_i \kappa_j, \tag{2.7}$$

for all $i, j \in \{1, \ldots, n\}$.The identities (2.5), (2.6) and (2.7) form the basis of the calculations that we perform in the present report. And they are also the basis for generalizing the GEKS method.
■

Example 2.2.1: Computation of GEKS

Consider the 6×6 PIM implicit in Table 2.2.1. It is not transitive, as e.g. $p_{24} = 1.35 \neq p_{23}p_{34} = 1.4 \times 1.05 = 1.47$. The numbers in the blue fields are the geometric averages of the corresponding columns. The numbers the green fields are the geometric averages of the corresponding rows. Note that a column (geometric) average and its corresponding row (geometric) average are reciprocal values. This holds in general. It will be used here to calculate the adjusted price index.

2.2.1 The input PIM.

Input PIM		1	1,1	1,2	1,15	1,3	1,2	1,154515
		0,909091	1	1,4	1,35	1,23	1,3	1,183458
		0,833333	0,714286	1	1,05	1,11	0,9	0,924501
		0,869565	0,740741	0,952381	1	1,2	0,98	0,947031
		0,769231	0,813008	0,900901	0,833333	1	1,05	0,888804
		0,833333	0,769231	1,111111	1,020408	0,952381	1	0,940523
		0,866165	0,844981	1,081664	1,055931	1,125107	1,063238	

Table 2.2.2 shows the resulting table, with the adjusted PIM matrix included. It should be noted that taking a geometric average in Excel is a standard function.

2.2.2 The resulting GEKS PIM matrix.

GEKS PIM		1	0,975543	1,248797	1,219088	1,298953	1,227523	1,154515
		1,02507	1	1,280104	1,24965	1,331517	1,258298	1,183458
		0,800771	0,781186	1	0,97621	1,040163	0,982965	0,924501
		0,820285	0,800224	1,02437	1	1,065512	1,00692	0,947031
		0,769851	0,751023	0,961388	0,938516	1	0,94501	0,888804
		0,814648	0,794725	1,01733	0,993128	1,05819	1	0,940523
		0,866165	0,844981	1,081664	1,055931	1,125107	1,063238	

The formulas used to calculate the elements of the adjusted PIM from the geometric column averages in Table 2.2.2 are shown in Table 2.2.3. The row averages (in green) are in column K, and the column averages (in blue) are in row 19. ■

This small example is only to illustrate the computation of the GEKS method. In the next subsection we consider an example with some real data.

Example 2.2.2: Illustrating standard GEKS with real data

The example we consider here is a bit bigger and contains 'real' data, that have been used in a study to compare various index formulas for various types of data. In fact the data used in the following example are about t-shirts, and the index used is a Törnquist index. For the computation of the GEKS index such details are of no importance, of course. Only the results are more realistic. The data in this example are also used for other examples in the present paper.

In Table 2.2.5 the GEKS index derived from Table 2.2.4 is presented.

The GEKS estimate in Figure 2.2.6 is actually very close to the original price index. The differences are exaggerated by the choice of scale of the vertical axis.

2.2.3 Formulas to compute the elements of the adjusted PIM from the geometric averages.

GEKS PIM		=E$19*$K13	=F$19*$K13	=G$19*$K13	=H$19*$K13	=I$19*$K13	=J$19*$K13	1,1545145085696
		=E$19*$K14	=F$19*$K14	=G$19*$K14	=H$19*$K14	=I$19*$K14	=J$19*$K14	1,18345821988097
		=E$19*$K15	=F$19*$K15	=G$19*$K15	=H$19*$K15	=I$19*$K15	=J$19*$K15	0,924501423631008
		=E$19*$K16	=F$19*$K16	=G$19*$K16	=H$19*$K16	=I$19*$K16	=J$19*$K16	0,947081451222124
		=E$19*$K17	=F$19*$K17	=G$19*$K17	=H$19*$K17	=I$19*$K17	=J$19*$K17	0,898804207313588
		=E$19*$K18	=F$19*$K18	=G$19*$K18	=H$19*$K18	=I$19*$K18	=J$19*$K18	0,940523359016882
		0,866164948623267	0,844981244958635	1,08166409963164	1,05593114010049	1,12510718532994	1,0632378137267	

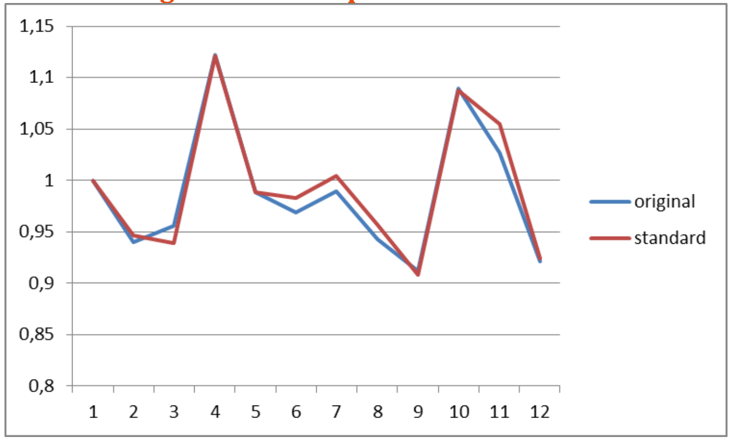
2.2.4 Example with real data. In the right-most column and the bottom row are geometric averages of the corresponding rows and columns, respectively.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,940320492	0,889363527	0,992824078	0,96563071	0,954406142	0,961536407	0,929931164	0,847754061	0,938004494	0,974491835	0,913626909	0,941385532
m2	1,063467199	1	0,955660046	1,059086442	1,042946219	1,023345947	1,034205914	0,978198886	0,892849386	0,897777379	1,040803194	0,969707966	0,99487724
m3	1,124399606	1,046397205	1	1,122143269	1,110308886	1,089186907	1,09895587	1,025155067	0,942518711	1,047036886	1,103998423	1,025477886	1,059889753
m4	1,007227788	0,944209991	0,891151806	1	0,988918722	0,964200199	0,970998943	0,920599818	0,843748391	0,934152424	0,980191886	0,911265552	0,945212268
m5	1,035592582	0,958822211	0,900650272	1,011205449	1	0,969246089	0,97066474	0,928774893	0,852098167	0,947910964	0,993439674	0,919789433	0,956060811
m6	1,047771966	0,977186652	0,918116068	1,037129013	1,031729724	1	0,989437461	0,943633318	0,864836454	0,959210088	1,000628471	0,923845172	0,973018278
m7	1,040002222	0,966925432	0,909954647	1,029867239	1,030221826	1,010675297	1	0,943199754	0,861360312	0,958180249	0,987227857	0,907737195	0,968906777
m8	1,075348412	1,022286995	0,975462183	1,086248314	1,076687158	1,059733671	1,060220803	1	0,912085772	0,925342441	1,035881042	0,944715738	1,012724028
m9	1,179587389	1,120009731	1,060986894	1,185187446	1,173573702	1,156287984	1,160954349	1,096388115	1	1,089084506	1,137510538	1,04006003	1,115123695
m10	1,066092974	1,113861881	0,95507619	1,070489113	1,054951401	1,042524536	1,043644973	1,080681006	0,918202393	1	1,026698709	0,947102368	1,024981849
m11	1,026175864	0,960796437	0,905798396	1,020208404	1,006603648	0,999371923	1,012937381	0,965361812	0,879112735	0,973995576	1	0,921391606	0,971530696
m12	1,09453869	1,031238306	0,975155109	1,097374961	1,087205358	1,082432458	1,101640436	1,058519467	0,961492208	1,055852075	1,085314858	1	1,051472415
	1,062264042	1,005149138	0,943494356	1,057963416	1,045959119	1,027729923	1,032091087	0,987435838	0,896761583	0,975627083	1,029303662	0,951047298	

2.2.5 GEKS estimate based on the geometric averages in Table 2.2.4.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,946232855	0,888191936	0,995951453	0,984650781	0,96749008	0,97159557	0,929557812	0,84419838	0,918441173	0,968971575	0,895302167	0,941385532
m2	1,056822319	1	0,938661061	1,052543723	1,040600921	1,02246511	1,026803883	0,982377442	0,892167689	0,97062913	1,024030786	0,946175312	0,99487724
m3	1,125882773	1,065347271	1	1,121324584	1,108601352	1,089280415	1,093902715	1,046573127	0,950468413	1,034057095	1,090948404	1,008005286	1,059889753
m4	1,004065005	0,950079296	0,89180244	1	0,988653391	0,971422932	0,97554511	0,933336469	0,84763005	0,922174641	0,972910449	0,898941574	0,945212268
m5	1,01558849	0,960983197	0,902037507	1,011476832	1	0,98257179	0,986741278	0,944048215	0,857358158	0,932758284	0,984076379	0,909258576	0,956060811
m6	1,033602329	0,978028483	0,918037253	1,029417741	1,01773734	1	1,004243444	0,960793119	0,872565411	0,949302985	1,001531276	0,925386405	0,973018278
m7	1,029234829	0,973895811	0,914158075	1,025067923	1,013436878	0,995774487	1	0,956733275	0,868878375	0,945291643	0,997299293	0,921476172	0,968906777
m8	1,07578032	1,017938684	0,955499405	1,071424972	1,059267932	1,040806788	1,045223393	1	0,908172003	0,988040989	1,042400651	0,963148451	1,012724028
m9	1,184555697	1,12086552	1,052112818	1,179759967	1,166373692	1,146045886	1,150909068	1,101113002	1	1,087944724	1,147800799	1,060535282	1,115123695
m10	1,088801362	1,030259622	0,967064589	1,084393298	1,072089111	1,053404517	1,05787458	1,012103812	0,919164345	1	1,05501757	0,974806218	1,024981849
m11	1,032022018	0,976533141	0,916633634	1,027843828	1,016181286	0,998471065	1,002708021	0,959324129	0,871231315	0,947851513	1	0,923971549	0,971530696
m12	1,116941338	1,056886591	0,992058289	1,112419348	1,099797161	1,080629665	1,085215256	1,038261546	0,942920068	1,025844913	1,082284407	1	1,051472415
	1,062264042	1,005149138	0,943494356	1,057963416	1,045959119	1,027729923	1,032091087	0,987435838	0,896761583	0,975627083	1,029303662	0,951047298	

2.2.6 Original data compared with standard GEKS estimates.



2.3 Properties and conditions

The example of the GEKS method should give a clear idea of the working of the method for periods other than the 6 months period. The method is pretty straightforward to apply. But in order to apply it some formal conditions need to be met:

- P should be complete, that is, p_{ij} should be known for any $i, j \in \{1, \dots, n\}$,
- P satisfies the property that $p_{ji}p_{ij} = 1$, for all $i, j \in \{1, \dots, n\}$,
- $p_{ii} = 1$ for all $i \in \{1, \dots, n\}$.

We briefly comment on these requirements. The first one can in fact be discarded while still being able to define a GEKS-like estimator. But this requirement can in fact be relaxed. It is still possible to define GIKS like method in case some of the p_{ij}'s are not known, as is shown in the sequel (see sections 3.2, 3.3 and 3.6). The latter two requirements are not very severe, and even pretty 'normal'. These conditions can easily be met. However, the Carli price index is an example of a price index which typically violates the second requirement.

An abstract view on GEKS

The GEKS method is applied to produce transitive price indices from nontransitive ones. But if it is applied to a transitive index, it produces the index itself. See Section 3.4.3 for a proof.

Looking at transitivizing methods like GEKS in a more abstract way we can remark the following. We start with considering an index at some interval of length n. We assume that the price index is represented by a square matrix P, its PIM, as in (2.1).

The transitivization of an price index with PIM P is a mapping of P to a PIM $\tau(P)$, which belongs to a transitive price index. The GEKS method provides a particular mapping τ, which we shall denote by $\mathcal{G}: \mathcal{C} \to \mathcal{C}$, where $\mathcal{C}$ is the set of complete PIM matrices, that is, the matrices with the following properties:

- $P > 0$, that is, $p_{ij} > 0$, for all $i, j \in \{1, \dots, n\}$,
- $p_{ii} = 1$, for $i \in \{1, \dots n\}$,
- $p_{ji} = \frac{1}{p_{ij}}$, for all $i, j \in \{1, \dots, n\}$.

If we look at

$$\log P \triangleq \begin{pmatrix} \log p_{11} & \cdots & \log p_{1n} \\ \vdots & \ddots & \vdots \\ \log p_{n1} & \cdots & \log p_{nn} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & \log p_{1n} \\ \vdots & \ddots & \vdots \\ -\log p_{1n} & \cdots & 0 \end{pmatrix}, \tag{2.8}$$

we find that it is an $n \times n$ skew symmetric matrix, that is $(\log P)' = -\log P$, where ' denotes matrix transposition.

A subset of $\mathcal{C}$ is the set of transitive PIM matrices $\mathcal{T}$, that is $\mathcal{T} \subseteq \mathcal{C}$. They satisfy the additional property that

$$p_{ij} = \prod_{k=i}^{j-1} p_{k,k+1}, \qquad (2.9)$$

for $i < j$.

If $P \in \mathcal{C}$ then $\mathcal{G}(P) \epsilon \mathcal{T} \subseteq \mathcal{C}$. In fact, from (2.3) it follows immediately that a transitive PIM is a fixed point under this mapping, that is $\mathcal{G}(T) = T$, for $T \epsilon \mathcal{T}$. For we have

$$p_{ij}^G = \prod_{k=1}^{n} (p_{ik} p_{kj})^{\frac{1}{n}} = \left(\prod_{k=1}^{n} p_{ik} p_{kj}\right)^{\frac{1}{n}} = \left(\prod_{k=1}^{n} p_{ij}\right)^{\frac{1}{n}} = p_{ij},$$

where it was used that $p_{ik} p_{kj} = p_{ij}$, because of the transitivity of the price index.

Therefore, for any $P \epsilon \mathcal{C}$, $\mathcal{G}(\mathcal{G}(P)) = \mathcal{G}(P)$, which can be rephrased by saying that $\mathcal{G}$ is idem-potent. In plain words, it means that the GEKS method applied to a GEKS index yields the same GEKS index. This property holds for the generalizations of GEKS that we shall consider below as well.

The extensions of GEKS considered below relax the condition that the PIM should be complete. For certain incomplete PIMs (i.e. with some elements missing), it is also possible to apply a GEKS-like transitivization.

3. Weighted GEKS methods

3.1 Preliminary remarks

In the GEKS method the averaging period for each month is taken to be the entire time window that is being considered. This is not necessary, and probably not desirable in case this period is rather long. It is more attractive to look at shorter periods, that also can be accommodated to the individual months. This is a first example of using weights in GEKS (other than 1). Now the weights are 1 or 0. But one can even go beyond this and use more general weights. These weights can be used to indicate that price indices about months widely separated should be treated differently from those that are closer. Because this separation reflects itself in the amount of overlapping articles, it is reasonable to give indices concerning nearby months a higher weight than those concerning more widely separated months.

3.2 Averaging periods

In the GEKS method averages are taken of the indices in the entire period. It is, however, easy to produce a variant in which the averaging period is not fixed but variable. Important is that for each month i in the time window a unique averaging period A_i is associated. For simplicity we will assume that these averaging periods are – in the typical case - of the same length and that the reference month is part of the associated averaging period. So we assume for typical months i and j that $|A_i| = |A_j|$, where $|B|$ indicates the size of set B. Typical months can be find in the middle of a time window. At the beginning or end of the window, we find atypical months, in the sense that the averaging periods associated with these months are different from those of the typical months. They tend to be shorter, although that is a matter of choice. For a long period the influence of the typical months are dominating. We now present some examples illustrating various GEKS variants.

Example: Four periods GEKS

In Table 3.2.1 an example with averaging periods within a time window of one year are shown, as an example. The two months at the beginning and end have shorter averaging periods due to 'transient phenomena' at the beginning and end of the time window. But these are considered as atypical. For the remaining months in the middle of the period (months 3 – 10) the averaging periods are of the same length and the reference month for such a period is in the middle. In fact, what the best length is of the 'ideal' averaging period is hard to say. One would not like to take a period that is too long, as a lot of information from far removed months is then taken into account. It is unclear that the shortest sensible period (of length 3, except at the boundaries of the time window considered) is the best option. In fact, numerical results obtained below seem to contradict this. So one should probably look for periods that are neither too short nor too long. What periods to take requires some experimentation with real data.

3.2.1 Averaging periods. Reference months in red.

	m1	m2	m3	m4	m5	m6	m7	m8	m10	m11	m12
m1	x	x	x								
m2	x	x	x	x							
m3	x	x	x	x	x						
m4		x	x	x	x	x					
m5			x	x	x	x	x				
m6				x	x	x	x	x			
m7					x	x	x	x	x		
m8						x	x	x	x	x	
m10							x	x	x	x	x
m11								x	x	x	x
m12									x	x	x

Table 3.2.3 contains the GEKS estimates of the price indices based on the geometric averages of Table 3.2.2.

The ability to use the means in the marginal to give new estimates for price index pairs that were originally known stretches to the possibility to provide it also for pairs for which the index was not known. In fact this is simply a result of applying transitive closure to an index that is transitive. If known for at least a spanning tree in the original price index digraph (PIDG)[6], it allows to give values for all remaining arcs (ordered pairs of months) as well.

[6] The pairs of months in the time window for which price indices have been calculated form a directed graph, which I called the PIDG. The underlying graph is called the price index graph (PIG). These and other graph inspired concepts are more fully explained in Willenborg (2017).

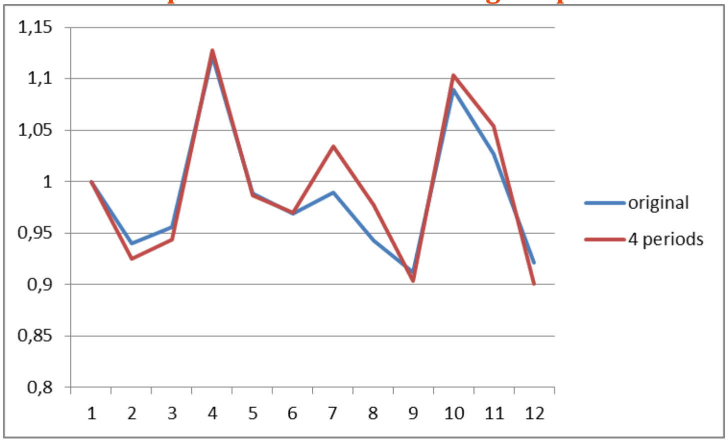
3.2.2 Example with averaging periods as in Table 2.2.4, using geometric averages.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,940320492	0,889363527										0,942146426
m2	1,063467199	1	0,955660045	1,059086442									1,018567311
m3	1,124399606	1,046397205	1	1,122143269	1,110308886								1,079498084
m4		0,944209991	0,891151806	1	0,988918722	0,964200199							0,956906753
m5			0,900650272	1,011205449	1	0,969246089	0,97066474						0,969571362
m6				1,037129013	1,031729724	1	0,989437461	0,943633318					0,999811363
m7					1,030221826	1,010675297	1	0,943199754	0,861360312				0,967088428
m8						1,059733671	1,060220803	1	0,912085772	0,925342441			0,989432664
m9							1,160954349	1,096388115	1	1,089084506	1,137510538		1,095368095
m10								1,080681006	0,918202393	1	1,026698709	0,947102368	0,992876371
m11									0,879112735	0,973995576	1	0,921391606	0,942456799
m12										1,055852075	1,085314858	1	1,046452714
	1,061406138	0,98177115	0,926356438	1,045033904	1,031383598	1,000188672	1,034031606	1,010680198	0,912936784	1,007174739	1,061056593	0,955609852	

3.2.3 Estimated price indices on the basis of averages of Table 3.2.2.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,92497218	0,872763408	0,984574958	0,971714371	0,942324184	0,974209182	0,952208737	0,860120129	0,948906081	0,999670677	0,900323936	0,942146426
m2	1,081113596	1	0,943556386	1,064437373	1,050533618	1,018759487	1,053230792	1,029445811	0,929887565	1,025875265	1,080757561	0,973352448	1,018567311
m3	1,145785893	1,059820075	1	1,128112097	1,113376618	1,079701756	1,116235137	1,091027337	0,985513509	1,087243201	1,145408559	1,031578465	1,079498084
m4	1,015666702	0,939463443	0,886436732	1	0,98693798	0,957087295	0,989471827	0,967126707	0,873595374	0,963772309	1,01533222	0,914429043	0,956906753
m5	1,029108995	0,951897191	0,898168673	1,013234945	1	0,969754294	1,002567432	0,979926576	0,885157361	0,976527783	1,028770086	0,926531461	0,969571362
m6	1,061205918	0,981585951	0,926181698	1,044836772	1,031189041	1	1,033836549	1,010489546	0,91276457	1,006984748	1,060856439	0,955429089	0,999811363
m7	1,026473593	0,949459518	0,895868591	1,010640195	0,997439142	0,967270891	1	0,977417123	0,882890699	0,974027085	1,026135552	0,924158746	0,967088428
m8	1,050189902	0,971396444	0,916567318	1,033990679	1,020484621	0,989619342	1,023104646	1	0,903289474	0,996531584	1,049844051	0,945511107	0,989432664
m9	1,162628297	1,07539883	1,014699434	1,144694706	1,129742624	1,095572761	1,132643162	1,107064821	1	1,103225061	1,162245417	1,046742084	1,095368095
m10	1,053845075	0,974777377	0,919757419	1,03758947	1,024036404	0,9930637	1,026665549	1,003480487	0,906433361	1	1,05349802	0,948801946	0,992876371
m11	1,000329431	0,925276895	0,873050923	0,984899308	0,972034484	0,942634615	0,974530117	0,952522424	0,860403479	0,94921868	1	0,900620531	0,942456799
m12	1,110711334	1,027377084	0,969388208	1,093578565	1,079294165	1,046650151	1,08206518	1,057629086	0,955345175	1,053960739	1,110345551	1	1,046452714
	1,061406138	0,98177115	0,926356438	1,045033904	1,031383598	1,000188672	1,034031606	1,010680198	0,912936784	1,007174739	1,061056593	0,955609852	

3.2.4 Four periods GEKS and the original price index.



Remark It is possible to choose the reference period for a GEKS method in such a way that the last month of this period coincides with the reference month. This has the advantage that the GEKS index can then be computed as soon as the data for the current month become available.

■

Minimal GEKS: numerical results

The GEKS variant that we consider in the present section is called minimal GEKS because the averaging period used (of the typical months, that is away from the boundary) have minimal length, namely three months. Three months is minimum in the nontrivial case: the matrix has to be symmetric. This excludes periods of length 2. For periods of length 1 GEKS has nothing new to offer; it yields the same results.

In Table 3.2.5 an example of a minimal GEKS situation is presented.

3.2.5 Minimal GEKS for a window of 12 months.

	m1	m2	m3	m4	m5	m6	m7	m8	m10	m11	m12
m1	x	x									
m2	x	x	x								
m3		x	x	x							
m4			x	x	x						
m5				x	x	x					
m6					x	x	x				
m7						x	x	x			
m8							x	x	x		
m10								x	x	x	
m11									x	x	x
m12										x	x

Using the scheme of Table 3.2.5 we present some results on real data (derived from Table 3.2.2, which in turn are derived from Table 2.2.4). The data used are presented in Table 3.2.6. The geometric averages per column or row are in the bottom row and right-most column.

In Table 3.2.7 the estimates for minimal GEKS are presented, based on the geometric averages computed in Table 3.2.6.

3.2.6 Example of minimal GEKS. Data in Table 3.2.2 are used. Geometric column and row averages in bottom row and right most column.

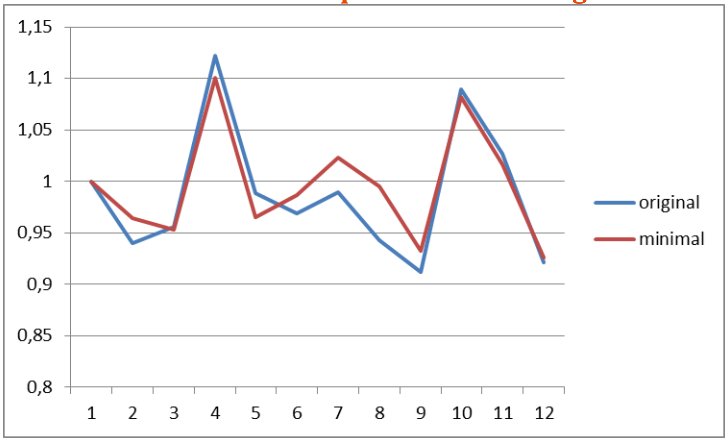
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,940320492											0,969701238
m2	1,063467199	1	0,955660045										1,0054084
m3		1,046397205	1	1,122143269									1,054989679
m4			0,891151806	1	0,988918722								0,958747174
m5				1,011205449	1	0,969246089							0,993324613
m6					1,031729724	1	0,989437461						1,006896344
m7						1,010675297	1	0,943199754					0,984173748
m8							1,060220803	1	0,912085772				0,988880927
m9								1,096388115	1	1,089084506			1,060902077
m10									0,918202393	1	1,026698709		0,980529081
m11										0,973995576	1	0,921391606	0,964569989
m12											1,085314858	1	1,041784458
	1,03124546	0,994620693	0,947876392	1,043027846	1,006720349	0,99315089	1,01608075	1,011244097	0,942594064	1,019857564	1,086731405	0,959891455	

3.2.7 Minimal GEKS results based on Table 3.2.6.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,964484918	0,919156911	1,011425394	0,976217969	0,963059648	0,985294761	0,980604653	0,914034631	0,988957143	1,005319728	0,930807933	0,969701238
m2	1,036822849	1	0,953002887	1,048668959	1,012165096	0,998522248	1,021576121	1,01671331	0,94769199	1,025373362	1,042338464	0,965082933	1,0054084
m3	1,087953523	1,049314764	1	1,100383821	1,062079779	1,047764137	1,071954907	1,066852287	0,994427197	1,075939408	1,09374114	1,01267577	1,054989679
m4	0,988703671	0,953589778	0,908773812	1	0,965190289	0,952180609	0,974164547	0,96952742	0,903709395	0,977785557	0,993963305	0,92029322	0,958747174
m5	1,024361394	0,987981115	0,941548855	1,036065127	1	0,986521124	1,009297916	1,004493549	0,936301789	1,013049518	1,029810718	0,953483712	0,993324613
m6	1,038357284	1,001479939	0,954413274	1,050220925	1,013663039	1	1,023087992	1,018217983	0,949094517	1,026890852	1,043881062	0,966511197	1,006896344
m7	1,01492471	0,978879576	0,932875062	1,026520625	0,990787739	0,977433034	1	0,995239893	0,927676333	1,003717041	1,020323833	0,944699972	0,984173748
m8	1,019778967	0,983561433	0,937336886	1,031430344	0,995526552	0,982107973	1,004782874	1	0,932113292	1,008517694	1,025203914	0,949218353	0,988880927
m9	1,094050451	1,055195159	1,005604033	1,106550409	1,068031709	1,053635842	1,077962178	1,072830962	1	1,081969008	1,099870501	1,018350839	1,060902077
m10	1,011166164	0,975254514	0,929420368	1,022719136	0,987118579	0,97381333	0,996296724	0,991554245	0,924240891	1	1,016545292	0,941201487	0,980529081
m11	0,994708422	0,959381271	0,914293121	1,006073358	0,971052236	0,957963543	0,980080597	0,975415707	0,909197946	0,983723999	1	0,92588249	0,964569989
m12	1,074335493	1,03618038	0,987482894	1,0866102	1,048785613	1,034649162	1,058537133	1,053498383	0,981979846	1,06247176	1,080050666	1	1,041784458
	1,03124546	0,994620693	0,947876392	1,043027846	1,006720349	0,99315089	1,01608075	1,011244097	0,942594064	1,019857564	1,086731405	0,959891455	

For minimal GEKS it is feasible and illuminating to look at the symbolic expressions that the method produces. In principle, these computations can be carried out for any variant of the GEKS method, except that the symbolic expressions can be very long. In case of minimal GEKS the expressions are manageable.

3.2.8 Minimal GEKS compared with the original data.



Minimal GEKS: symbolic results

In Table 3.2.9 an example is shown of a six month period. Some of the indices are shown, symbolically, as well as the row en column (geometric) averages.

3.2.9 Minimal GEKS symbolic calculations

	m1	**m2**	**m3**	**m4**	**m5**	**m6**	
m1	1	$p_{1,2}$	–	–	–	–	$p_{1,2}^{\frac{1}{2}}$
m2	$p_{1,2}^{-1}$	1	$p_{2,3}$	–	–	–	$p_{2,3}^{\frac{1}{3}} p_{1,2}^{-\frac{1}{3}}$
m3	–	$p_{2,3}^{-1}$	1	$p_{3,4}$	–	–	$p_{3,4}^{\frac{1}{3}} p_{2,3}^{-\frac{1}{3}}$
m4	–	–	$p_{3,4}^{-1}$	1	$p_{4,5}$		$p_{4,5}^{\frac{1}{3}} p_{3,4}^{-\frac{1}{3}}$
m5	–	–	–	$p_{4,5}^{-1}$	1	$p_{5,6}$	$p_{5,6}^{\frac{1}{3}} p_{4,5}^{-\frac{1}{3}}$
m6	–	–	–	–	$p_{5,6}^{-1}$	1	$p_{5,6}^{-\frac{1}{2}}$
	$p_{1,2}^{-\frac{1}{2}}$	$p_{1,2}^{\frac{1}{3}} p_{2,3}^{-\frac{1}{3}}$	$p_{2,3}^{\frac{1}{3}} p_{3,4}^{-\frac{1}{3}}$	$p_{3,4}^{\frac{1}{3}} p_{4,5}^{-\frac{1}{3}}$	$p_{4,5}^{\frac{1}{3}} p_{5,6}^{-\frac{1}{3}}$	$p_{5,6}^{\frac{1}{2}}$	

In Table 3.3.1 the minimal GEKS expressions for the indices are shown, based on the symbolic geometric averages in Table 3.2.9.

It is easy to verify directly that the adjusted index in Table 3.3.1 is indeed transitive. Also closer inspection of the expressions in Table 3.3.1 shows how non-intuitive they are in fact. Table 3.3.1 brings to light how complex the method basically is, not so much computationally but in understanding how it actually works.

3.3 Weighted averages

In Section 3.2 an extension of the standard GEKS method was considered in which the averaging period was tied to the month to which it was associated. This allowed to work locally, whereas the GEKS method works globally, that is, over the entire time window. Another aspect of the GEKS method is the fact that each input index is weighted with the same weight. This may be less attractive for price indices with base month and reporting month far apart. In the extreme case there may be no overlap in articles at all. In the less extreme, but still unattractive case, there may be a handful of articles present in both months. In the method of Section 3.2 such price indices can be left out of the weighting, by defining a more localized averaging period. This can be viewed as a situation where price indices have a weight 0 or 1. It is possible to consider a more general case were indices can be weighted with nonzero weights. In the cycle method the same kind of weights can be used, for exactly the same reason.

Example: symbolic results

In this example we illustrate weighted GEKS with symbolic computation. The situation concerns a 5 month period. In Table 3.3.2 the price indices are shown, weighted with weights w_1 or w_2. The choice of the weights depends on the separation of the base and reporting month, which is either 1 (w_1) or 2 (w_2). We assume $1 \geq w_1 \geq w_2$. The exact choice has to be specified by the analyst. This requires some experimentation, probably, to find some suitable values.

In the bottom row and the right most column the geometric averages of the corresponding columns and rows are given.

In Table 3.3.3, symbolic expressions for the weighted GEKS price index are given.

3.3.1 Minimal GEKS price indices, symbolic expressions

	m1	m2	m3	m4	m5	m6	
m1	1	$p_{1,2}^{\frac{5}{6}}p_{2,3}^{-\frac{1}{3}}$	$p_{1,2}^{\frac{1}{2}}p_{2,3}^{\frac{1}{3}}p_{3,4}^{-\frac{1}{3}}$	$p_{1,2}^{\frac{1}{2}}p_{3,4}^{\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}$	$p_{1,2}^{\frac{1}{2}}p_{4,5}^{\frac{1}{3}}p_{5,6}^{-\frac{1}{3}}$	$p_{1,2}^{\frac{1}{2}}p_{5,6}^{\frac{1}{2}}$	$p_{1,2}^{\frac{1}{2}}$
m2	$p_{1,2}^{-\frac{5}{6}}p_{2,3}^{\frac{1}{3}}$	1	$p_{1,2}^{-\frac{1}{3}}p_{2,3}^{\frac{2}{3}}p_{3,4}^{-\frac{1}{3}}$	$p_{1,2}^{-\frac{1}{3}}p_{2,3}^{\frac{1}{3}}p_{3,4}^{\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}$	$p_{1,2}^{-\frac{1}{3}}p_{2,3}^{\frac{1}{3}}p_{4,5}^{\frac{1}{3}}p_{5,6}^{-\frac{1}{3}}$	$p_{1,2}^{-\frac{1}{3}}p_{2,3}^{\frac{1}{3}}p_{5,6}^{\frac{1}{2}}$	$p_{2,3}^{\frac{1}{3}}p_{1,2}^{-\frac{1}{3}}$
m3	$p_{1,2}^{-\frac{1}{2}}p_{2,3}^{-\frac{1}{3}}p_{3,4}^{\frac{1}{3}}$	$p_{1,2}^{\frac{1}{3}}p_{2,3}^{-\frac{2}{3}}p_{3,4}^{\frac{1}{3}}$	1	$p_{2,3}^{-\frac{1}{3}}p_{3,4}^{\frac{2}{3}}p_{4,5}^{-\frac{1}{3}}$	$p_{2,3}^{-\frac{1}{3}}p_{3,4}^{\frac{1}{3}}p_{4,5}^{\frac{1}{3}}p_{5,6}^{-\frac{1}{3}}$	$p_{2,3}^{-\frac{1}{3}}p_{3,4}^{\frac{1}{3}}p_{5,6}^{\frac{1}{2}}$	$p_{2,3}^{-\frac{1}{3}}p_{3,4}^{\frac{1}{3}}$
m4	$p_{1,2}^{-\frac{1}{2}}p_{3,4}^{-\frac{1}{3}}p_{4,5}^{\frac{1}{3}}$	$p_{1,2}^{\frac{1}{3}}p_{2,3}^{-\frac{1}{3}}p_{3,4}^{-\frac{1}{3}}p_{4,5}^{\frac{1}{3}}$	$p_{2,3}^{\frac{1}{3}}p_{3,4}^{-\frac{2}{3}}p_{4,5}^{\frac{1}{3}}$	1	$p_{3,4}^{-\frac{1}{3}}p_{4,5}^{\frac{2}{3}}p_{5,6}^{-\frac{1}{3}}$	$p_{3,4}^{-\frac{1}{3}}p_{4,5}^{\frac{1}{3}}p_{5,6}^{\frac{1}{2}}$	$p_{3,4}^{-\frac{1}{3}}p_{4,5}^{\frac{1}{3}}$
m5	$p_{1,2}^{-\frac{1}{2}}p_{4,5}^{-\frac{1}{3}}p_{5,6}^{\frac{1}{3}}$	$p_{1,2}^{\frac{1}{3}}p_{2,3}^{-\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}p_{5,6}^{\frac{1}{3}}$	$p_{2,3}^{\frac{1}{3}}p_{3,4}^{-\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}p_{5,6}^{\frac{1}{3}}$	$p_{3,4}^{\frac{1}{3}}p_{4,5}^{-\frac{2}{3}}p_{5,6}^{\frac{1}{3}}$	1	$p_{4,5}^{-\frac{1}{3}}p_{5,6}^{\frac{5}{6}}$	$p_{4,5}^{-\frac{1}{3}}p_{5,6}^{\frac{1}{3}}$
m6	$p_{1,2}^{-\frac{1}{2}}p_{5,6}^{-\frac{1}{2}}$	$p_{1,2}^{\frac{1}{3}}p_{2,3}^{-\frac{1}{3}}p_{5,6}^{-\frac{1}{2}}$	$p_{2,3}^{\frac{1}{3}}p_{3,4}^{-\frac{1}{3}}p_{5,6}^{-\frac{1}{2}}$	$p_{3,4}^{\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}p_{5,6}^{-\frac{1}{2}}$	$p_{4,5}^{\frac{1}{3}}p_{5,6}^{-\frac{5}{6}}$	1	$p_{5,6}^{-\frac{1}{2}}$
	$p_{1,2}^{-\frac{1}{2}}$	$p_{1,2}^{\frac{1}{3}}p_{2,3}^{-\frac{1}{3}}$	$p_{2,3}^{\frac{1}{3}}p_{3,4}^{-\frac{1}{3}}$	$p_{3,4}^{\frac{1}{3}}p_{4,5}^{-\frac{1}{3}}$	$p_{4,5}^{\frac{1}{3}}p_{5,6}^{-\frac{1}{3}}$	$p_{5,6}^{\frac{1}{2}}$	

3.3.2 Example of weighted price indices. The weights depend on the difference between the pair of months compared.

	m1	m2	m3	m4	m5	
m1	1	$p_{1,2}^{w_1}$	$p_{1,3}^{w_2}$	–	–	$p_{1,2}^{\frac{w_1}{3}}p_{1,3}^{\frac{w_2}{3}}$
m2	$p_{2,1}^{w_1}$	1	$p_{2,3}^{w_1}$	$p_{2,4}^{w_2}$	–	$p_{2,1}^{\frac{w_1}{4}}p_{2,3}^{\frac{w_1}{4}}p_{2,4}^{\frac{w_2}{4}}$
m3	$p_{3,1}^{w_2}$	$p_{3,2}^{w_1}$	1	$p_{3,4}^{w_1}$	$p_{3,5}^{w_2}$	$p_{3,1}^{\frac{w_2}{5}}p_{3,2}^{\frac{w_1}{5}}p_{3,4}^{\frac{w_1}{5}}p_{3,5}^{\frac{w_2}{5}}$
m4	–	$p_{4,2}^{w_2}$	$p_{4,3}^{w_1}$	1	$p_{4,5}^{w_1}$	$p_{4,2}^{\frac{w_2}{4}}p_{4,3}^{\frac{w_1}{4}}p_{4,5}^{\frac{w_1}{4}}$
m5	–	–	$p_{5,3}^{w_2}$	$p_{5,4}^{w_1}$	1	$p_{5,3}^{\frac{w_2}{3}}p_{5,4}^{\frac{w_1}{3}}$
	$p_{2,1}^{\frac{w_1}{3}}p_{3,1}^{\frac{w_2}{3}}$	$p_{1,2}^{\frac{w_1}{4}}p_{3,2}^{\frac{w_1}{4}}p_{4,2}^{\frac{w_2}{4}}$	$p_{1,3}^{\frac{w_2}{5}}p_{2,3}^{\frac{w_1}{5}}p_{4,3}^{\frac{w_1}{5}}p_{5,3}^{\frac{w_2}{5}}$	$p_{2,4}^{\frac{w_2}{4}}p_{3,3}^{\frac{w_1}{4}}p_{5,4}^{\frac{w_1}{4}}$	$p_{3,5}^{\frac{w_2}{3}}p_{4,5}^{\frac{w_1}{3}}$	

3.3.3 Weighted GEKS based on the marginals of Table 3.3.2.

	m1	m2	m3	m4	m5	
m1	1	$p_{1,2}^{\frac{7}{12}w_1} p_{1,3}^{\frac{w_2}{3}} p_{3,2}^{\frac{w_1}{4}} p_{4,2}^{\frac{w_2}{4}}$	$p_{1,2}^{\frac{w_1}{3}} p_{1,3}^{\frac{8}{15}w_2} p_{2,3}^{\frac{w_1}{5}} p_{4,3}^{\frac{w_1}{5}} p_{5,3}^{\frac{w_2}{5}}$	$p_{1,2}^{\frac{w_1}{3}} p_{1,3}^{\frac{w_2}{3}} p_{2,4}^{\frac{w_2}{4}} p_{3,3}^{\frac{w_1}{4}} p_{5,4}^{\frac{w_1}{4}}$	$p_{1,2}^{\frac{w_1}{3}} p_{1,3}^{\frac{w_2}{3}} p_{3,5}^{\frac{w_2}{3}} p_{4,5}^{\frac{w_1}{3}}$	$p_{1,2}^{\frac{w_1}{3}} p_{1,3}^{\frac{w_2}{3}}$
m2	$p_{2,1}^{\frac{7}{12}w_1} p_{2,3}^{\frac{w_1}{4}} p_{2,4}^{\frac{w_2}{4}} p_{3,1}^{\frac{w_2}{3}}$	1	$p_{1,3}^{\frac{w_2}{5}} p_{2,1}^{\frac{w_1}{4}} p_{2,3}^{\frac{9}{20}w_1} p_{2,4}^{\frac{w_2}{4}} p_{4,3}^{\frac{w_1}{5}} p_{5,3}^{\frac{w_2}{5}}$	$p_{2,1}^{\frac{w_1}{4}} p_{2,3}^{\frac{w_1}{4}} p_{2,4}^{\frac{w_2}{2}} p_{3,4}^{\frac{w_1}{4}} p_{5,4}^{\frac{w_1}{4}}$	$p_{2,1}^{\frac{w_1}{4}} p_{2,3}^{\frac{w_1}{4}} p_{2,4}^{\frac{w_2}{4}} p_{3,5}^{\frac{w_2}{3}} p_{4,5}^{\frac{w_1}{3}}$	$p_{2,1}^{\frac{w_1}{4}} p_{2,3}^{\frac{w_1}{4}} p_{2,4}^{\frac{w_2}{4}}$
m3	$p_{2,1}^{\frac{w_1}{3}} p_{3,1}^{\frac{8}{15}w_2} p_{3,2}^{\frac{w_1}{5}} p_{3,4}^{\frac{w_1}{5}} p_{3,5}^{\frac{w_2}{5}}$	$p_{1,2}^{\frac{w_1}{4}} p_{3,1}^{\frac{w_2}{5}} p_{3,2}^{\frac{9}{20}w_1} p_{3,4}^{\frac{w_1}{5}} p_{3,5}^{\frac{w_2}{5}} p_{4,2}^{\frac{w_2}{4}}$	1	$p_{2,4}^{\frac{w_2}{4}} p_{3,1}^{\frac{w_2}{5}} p_{3,2}^{\frac{w_1}{5}} p_{3,4}^{\frac{9}{20}w_1} p_{3,5}^{\frac{w_2}{5}} p_{5,4}^{\frac{w_1}{4}}$	$p_{3,1}^{\frac{w_2}{5}} p_{3,2}^{\frac{w_1}{5}} p_{3,4}^{\frac{w_1}{5}} p_{3,5}^{\frac{8}{15}w_2} p_{4,5}^{\frac{w_1}{3}}$	$p_{3,1}^{\frac{w_2}{5}} p_{3,2}^{\frac{w_1}{5}} p_{3,4}^{\frac{w_1}{5}} p_{3,5}^{\frac{w_2}{5}}$
m4	$p_{2,1}^{\frac{w_1}{3}} p_{3,1}^{\frac{w_2}{3}} p_{3,3}^{\frac{w_1}{4}} p_{4,2}^{\frac{w_2}{4}} p_{4,5}^{\frac{w_1}{4}}$	$p_{1,2}^{\frac{w_1}{4}} p_{3,2}^{\frac{w_1}{4}} p_{4,2}^{\frac{w_2}{2}} p_{4,3}^{\frac{w_1}{4}} p_{4,5}^{\frac{w_1}{4}}$	$p_{1,3}^{\frac{w_2}{5}} p_{2,3}^{\frac{w_1}{5}} p_{4,2}^{\frac{w_2}{4}} p_{4,3}^{\frac{9}{20}w_1} p_{4,5}^{\frac{w_1}{4}} p_{5,3}^{\frac{w_2}{5}}$	1	$p_{3,3}^{\frac{w_1}{4}} p_{3,5}^{\frac{w_2}{3}} p_{4,2}^{\frac{w_2}{4}} p_{4,5}^{\frac{7}{12}w_1}$	$p_{4,2}^{\frac{w_2}{4}} p_{4,3}^{\frac{w_1}{4}} p_{4,5}^{\frac{w_1}{4}}$
m5	$p_{2,1}^{\frac{w_1}{3}} p_{3,1}^{\frac{w_2}{3}} p_{5,3}^{\frac{w_2}{3}} p_{5,4}^{\frac{w_1}{3}}$	$p_{1,2}^{\frac{w_1}{4}} p_{3,2}^{\frac{w_1}{4}} p_{4,2}^{\frac{w_2}{4}} p_{5,3}^{\frac{w_2}{3}} p_{5,4}^{\frac{w_1}{3}}$	$p_{1,3}^{\frac{w_2}{5}} p_{2,3}^{\frac{w_1}{5}} p_{4,3}^{\frac{w_1}{5}} p_{5,3}^{\frac{8}{15}w_2} p_{5,4}^{\frac{w_1}{3}}$	$p_{2,4}^{\frac{w_2}{4}} p_{3,3}^{\frac{w_1}{4}} p_{5,3}^{\frac{w_2}{3}} p_{5,4}^{\frac{7}{12}w_1}$	1	$p_{5,3}^{\frac{w_2}{3}} p_{5,4}^{\frac{w_1}{3}}$
	$p_{2,1}^{\frac{w_1}{3}} p_{3,1}^{\frac{w_2}{3}}$	$p_{1,2}^{\frac{w_1}{4}} p_{3,2}^{\frac{w_1}{4}} p_{4,2}^{\frac{w_2}{4}}$	$p_{1,3}^{\frac{w_2}{5}} p_{2,3}^{\frac{w_1}{5}} p_{4,3}^{\frac{w_1}{5}} p_{5,3}^{\frac{w_2}{5}}$	$p_{2,4}^{\frac{w_2}{4}} p_{3,3}^{\frac{w_1}{4}} p_{5,4}^{\frac{w_1}{4}}$	$p_{3,5}^{\frac{w_2}{3}} p_{4,5}^{\frac{w_1}{3}}$	

3.4 Alternative ways of averaging

In the previous sections geometric averaging was used as the method for averaging price indices. However, instead of geometric averages one could use other averaging methods, like arithmetic, harmonic, etc. There is no particular reason why geometric averaging should be used. Nor is there a compelling reason to look for alternative for taking geometric means. So it is interesting to consider a few examples where other averaging methods are used in order to see if there are significant differences. As the price indices do not vary wildly (they typically are bounded from below (0) and from above (say, 3, in a normal economic situation) there is no particular reason to consider robust averaging methods, such as the median or trimmed means. But choosing them should be fine.

It should be stressed that the choice of the averaging method used in GEKS is unimportant for obtaining transitive price indices. Any averaging method should be allowed, in principle. What is important to produce transitive indices is that ratios of these averages are used to derive the adjusted indices.

We first look at arithmetical means. We take the data of Table 2.2.4 as input. In Table 3.4.1 the original data are shown, with the arithmetical mean of the columns (bottom row) and their reciprocals (in the right-most column). Note that the reciprocal values of the arithmetical means of the column averages are generally different from the arithmetical means of the correspon-ding row values. In that sense the method of calculation differs from that for the geometric means in standard GEKS and the variants discussed above.

On the basis of the averages in Table 3.4.1 we can calculate a GEKS index. The result is shown in Table 3.4.2.

3.4.1 Original data with arithmetical column averages in the bottom row and the right most column.

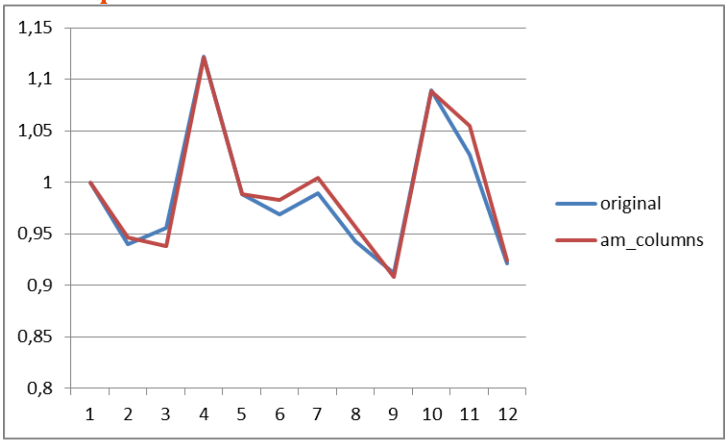
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	am_columns
m1	1	0,940320492	0,889363527	0,992824078	0,96563071	0,954406142	0,961536407	0,929931164	0,847754061	0,938004494	0,974491835	0,913626909	0,940423786
m2	1,063467199	1	0,955660045	1,059086442	1,042946219	1,023345947	1,034205914	0,978198886	0,892849386	0,897777379	1,040803194	0,969707966	0,993208496
m3	1,124399606	1,046397205	1	1,122143269	1,110308886	1,089186907	1,09895587	1,025155067	0,942518711	1,047036886	1,103998423	1,025477886	1,058446056
m4	1,007227788	0,944209991	0,891151806	1	0,988918722	0,964200199	0,970998943	0,920599818	0,843748391	0,934152424	0,980191886	0,911265552	0,944007477
m5	1,035592582	0,958822211	0,900650272	1,011205449	1	0,969246089	0,97066474	0,928774893	0,852098167	0,947910964	0,993439674	0,919789433	0,954746808
m6	1,047771966	0,977186652	0,918116068	1,037129013	1,031729724	1	0,989437461	0,943633318	0,864836454	0,959210038	1,000628471	0,923845172	0,971548907
m7	1,040002222	0,966925432	0,909954647	1,029867239	1,030221826	1,010675297	1	0,943199754	0,861360812	0,958180249	0,987227857	0,907737195	0,967336491
m8	1,075348412	1,022286995	0,975462183	1,086248314	1,076687158	1,059733671	1,060220808	1	0,912085772	0,925342441	1,035881042	0,944715738	1,010914226
m9	1,179587389	1,120009731	1,060986894	1,185187446	1,173573702	1,156287984	1,160954349	1,096388115	1	1,089084506	1,137510538	1,04005003	1,113579692
m10	1,066092974	1,113861881	0,95507619	1,070489113	1,054951401	1,042524536	1,043644973	1,080681006	0,918202393	1	1,026698709	0,947102368	1,023319138
m11	1,026175864	0,960796437	0,905798396	1,020208404	1,006603648	0,999371923	1,012937381	0,965361812	0,879112735	0,973995576	1	0,921391606	0,970888083
m12	1,09453869	1,031238306	0,975155109	1,097374961	1,087205358	1,082432458	1,101640436	1,058519467	0,961492208	1,055852075	1,085314858	1	1,050854902
	1,063350891	1,006837944	0,944781261	1,059313644	1,047398113	1,029284263	1,03376644	0,989203608	0,898004882	0,977212253	1,030515541	0,952059155	

3.4.2 GEKS estimates on the basis of the arithmetical column means in Table 3.4.1.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	am_columns
m1	1	0,946854351	0,88849477	0,996203747	0,984998098	0,967963408	0,972178549	0,930270602	0,844505151	0,918993646	0,969121326	0,895339074	0,940423786
m2	1,056128642	1	0,938364775	1,05211931	1,040284704	1,022293874	1,02674561	0,982485428	0,891906078	0,970575511	1,02351679	0,945593241	0,993208496
m3	1,125499028	1,065683651	1	1,121226349	1,108614402	1,089441869	1,094186011	1,047018658	0,950489726	1,034326455	1,09074511	1,007708257	1,058446056
m4	1,003810719	0,950462547	0,891880675	1	0,988751649	0,97165204	0,975883248	0,933815602	0,847723323	0,922495673	0,972814375	0,89875096	0,944007477
m5	1,015230887	0,961275309	0,902026889	1,011376315	1	0,98270586	0,986985204	0,944438983	0,857367291	0,932990274	0,983881418	0,908975435	0,954746808
m6	1,03309691	0,978192304	0,917901201	1,029175012	1,017598491	1	1,004354654	0,961059684	0,872455662	0,949409495	1,001196247	0,924972031	0,971548907
m7	1,028617636	0,973951084	0,91392139	1,024712743	1,013186415	0,995664227	1	0,956892747	0,868672892	0,945293071	0,996855287	0,920961562	0,967336491
m8	1,074956037	1,017826801	0,955092817	1,070875232	1,058829652	1,040518104	1,0450492	1	0,907805911	0,987877768	1,04176282	0,962450143	1,010914226
m9	1,1841254	1,121194288	1,052089226	1,179630161	1,166361267	1,146190052	1,151181313	1,101557049	1	1,088203719	1,147561178	1,06019374	1,113579692
m10	1,088146805	1,030316537	0,966812746	1,084015925	1,071822534	1,053286285	1,057872982	1,012270984	0,918945582	1	1,054546275	0,974260854	1,023319138
m11	1,031862547	0,977023543	0,916804477	1,027945336	1,016382647	0,998805183	1,003154634	0,959911393	0,871413236	0,948275125	1	0,923866858	0,970888083
m12	1,116895296	1,05753717	0,992355629	1,112655278	1,100139742	1,081113771	1,085821647	1,039014859	0,94322383	1,02641968	1,08240705	1	1,050854902
	1,063350891	1,006837944	0,944781261	1,059313644	1,047398113	1,029284263	1,03376644	0,989203608	0,898004882	0,977212253	1,030515541	0,952059155	

In Figure 3.4.3 the index based on the arithmetical column means of the original indices is shown, in comparison with the original indices.

3.4.3 GEKS based on arithmetical mean of columns compared with the original price indices.



Now we consider arithmetical row averages of the original data. In Table 3.4.1 they are presented in the final column. In the bottom row of this table the corresponding reciprocals of these averages are presented.

In Table 3.4.2 the GEKS estimates based on the row averages are presented. Comparison with Table 3.3.3 reveals a difference with the GEKS estimate based on column arithmetical averages, but these differences are small. For practical purposes these differences are insignificant.

3.4.4 Original data with arithmetical row averages in the right most column and their reciprocals in the bottom row.

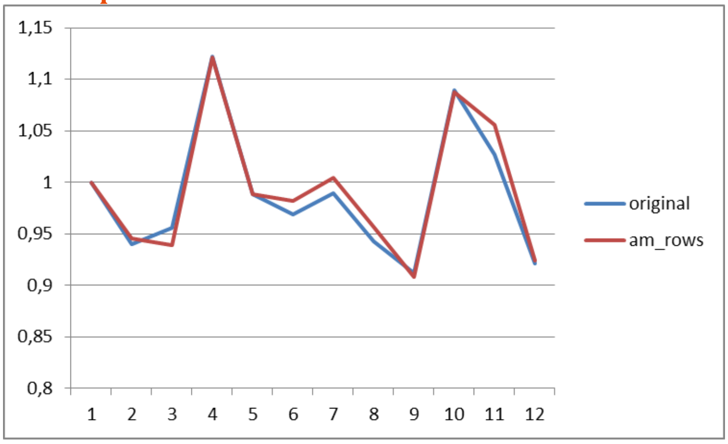
	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	am_rows
m1	1	0,940320492	0,889363527	0,992824078	0,96563071	0,954406142	0,961536407	0,929931164	0,847754061	0,938004494	0,974491835	0,913626909	0,942324152
m2	1,063467199	1	0,955660045	1,059086442	1,042946219	1,023345947	1,034205914	0,978198886	0,892849386	0,897777379	1,040803194	0,969707966	0,996504048
m3	1,124399606	1,046397205	1	1,122143269	1,110308886	1,089186907	1,09895587	1,025155067	0,942518711	1,047036886	1,103998423	1,025477886	1,061298226
m4	1,007227788	0,944209991	0,891151806	1	0,988918722	0,964200199	0,970998943	0,920699818	0,843748391	0,934152424	0,980191886	0,911265552	0,946388793
m5	1,035592582	0,958822211	0,900650272	1,011205449	1	0,969246089	0,97066474	0,928774893	0,852098167	0,947910964	0,993439674	0,919789433	0,95734954
m6	1,047771966	0,977186652	0,918116068	1,037129013	1,031729724	1	0,989437461	0,943633318	0,864836454	0,959210038	1,000628471	0,923845172	0,974460861
m7	1,040002222	0,966925432	0,909954647	1,029867239	1,030221826	1,010675297	1	0,943199754	0,861360312	0,958180249	0,987227857	0,907737195	0,970446002
m8	1,075348412	1,022286995	0,975462183	1,086248314	1,076687158	1,059733671	1,060220808	1	0,912085772	0,925342441	1,035881042	0,944715738	1,014501044
m9	1,179587389	1,120009731	1,060986894	1,185187446	1,173573702	1,156287984	1,160954349	1,096388115	1	1,089084506	1,137510538	1,04005008	1,116635057
m10	1,066092974	1,113861881	0,95507619	1,070489113	1,054951401	1,042524536	1,043644973	1,080681006	0,918202393	1	1,026698709	0,947102368	1,026610462
m11	1,026175864	0,960796437	0,905798396	1,020208404	1,006603648	0,999371923	1,012937381	0,965361812	0,879112735	0,973995576	1	0,921391606	0,972646148
m12	1,09453869	1,031238306	0,975155109	1,097374961	1,087205358	1,082432458	1,101640436	1,058519467	0,961492208	1,055852075	1,085314858	1	1,05256366
	1,061205954	1,003508216	0,942242223	1,056648184	1,044550562	1,026209007	1,030454036	0,985706231	0,895547738	0,9740793	1,028123127	0,9500613	

3.4.5 GEKS estimates on the basis of the arithmetical row means in Table 3.3.2.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	am_rows
m1	1	0,945630029	0,887897603	0,995705104	0,984305222	0,967021532	0,971021725	0,928854787	0,843896263	0,91789845	0,968825254	0,895265709	0,942324152
m2	1,057496029	1	0,938948189	1,052954193	1,040898864	1,02262143	1,026851618	0,982260249	0,892416947	0,970673965	1,024528858	0,946739932	0,996504048
m3	1,126255996	1,06502149	1	1,121418844	1,108579659	1,089113799	1,093619041	1,046128274	0,950443226	1,033788633	1,091145252	1,008298373	1,061298226
m4	1,004313422	0,94970893	0,89172748	1	0,988550946	0,971192704	0,975210152	0,93286133	0,847536344	0,921857733	0,973004206	0,899127368	0,946388793
m5	1,015945031	0,960708129	0,902055158	1,011581653	1	0,982440721	0,986504697	0,943665406	0,857352215	0,932534369	0,984273203	0,909540749	0,95734954
m6	1,034103137	0,977878979	0,918177697	1,029661771	1,017873119	1	1,004136612	0,96053165	0,872675773	0,949201666	1,001865234	0,925797078	0,974460861
m7	1,029843075	0,973850537	0,914395199	1,025420006	1,013679918	0,995880429	1	0,956574671	0,869080723	0,945291362	0,997737979	0,921983191	0,970446002
m8	1,076594548	1,018060133	0,955905719	1,071970686	1,059697636	1,041090109	1,045396695	1	0,908534116	0,988204467	1,043031986	0,963838181	1,014501044
m9	1,184979771	1,120552454	1,052140688	1,179890405	1,166381777	1,145900953	1,150641101	1,100674133	1	1,087691094	1,148038327	1,060871754	1,116635057
m10	1,089445134	1,030212034	0,967315724	1,084766081	1,072346535	1,053516903	1,057874894	1,011936329	0,919378678	1	1,055481959	0,975342871	1,026610462
m11	1,032177884	0,976058402	0,916468269	1,027744787	1,015978081	0,998138238	1,002267149	0,958743369	0,871051059	0,947434479	1	0,924073465	0,972646148
m12	1,116986823	1,056256282	0,991769923	1,112189481	1,099455963	1,080150309	1,084618472	1,037518558	0,942621006	1,025280473	1,082165042	1	1,05256366
	1,061205954	1,003508216	0,942242223	1,056648184	1,044550562	1,026209007	1,030454036	0,985706231	0,895547738	0,9740793	1,028123127	0,9500613	

For this reason the plot of these indices compared with the original indices in Figure 3.4.6 looks indistinguishable from the plot in Figure 3.4.3, based on column averages.

3.4.6 GEKS based on the arithmetical mean of rows compared with the original price indices.



Example: HM-GEKS

In the present example we have used the harmonic mean to compute column and row averages, that have been used, together with their reciprocal to compute the GEKS estimates. As input data we have also used those of Table 2.2.4. The results for the column averages are presented in Table 3.4.7, and those for the row averages in Table 3.4.8. As one can see, the results are very close.

Example: median-GEKS

In the present example we have used the median to compute column and row averages, that have been used, together with their reciprocal to compute the GEKS estimates. As input data we have used those of Table 2.2.4, as before. The results for the column averages are presented in Table 3.4.9, and those for the row averages in Table 3.4.10. As one can see the results are (again) very close.

3.4.7 Harmonic means of the columns used to compute the GEKS estimate of the data in Table 2.2.4.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	rec_hm_columns
m1	1	0,945630029	0,887897603	0,995705104	0,984305222	0,967021532	0,971021725	0,928854787	0,843896263	0,91789845	0,968825254	0,895265709	0,942324152
m2	1,057496	1	0,938948189	1,052954193	1,040898864	1,02262143	1,026851618	0,982260249	0,892416947	0,970673965	1,024528858	0,946739932	0,996504048
m3	1,126256	1,06502149	1	1,121418844	1,108579659	1,089113799	1,093619041	1,046128274	0,950443226	1,033788633	1,091145252	1,008298373	1,061298226
m4	1,004313	0,94970893	0,89172748	1	0,988550946	0,971192704	0,975210152	0,93286133	0,847536344	0,921857733	0,973004206	0,899127368	0,946388793
m5	1,015945	0,960708129	0,902055158	1,011581653	1	0,982440721	0,986504697	0,943665406	0,857352215	0,932534369	0,984273208	0,909540749	0,95734954
m6	1,034103	0,977878979	0,918177697	1,029661771	1,017873119	1	1,004136612	0,96053165	0,872675773	0,949201666	1,001865234	0,925797078	0,974460861
m7	1,029843	0,973850537	0,914395199	1,025420006	1,013679918	0,995880429	1	0,956574671	0,869080723	0,945291362	0,997737979	0,921983191	0,970446002
m8	1,076595	1,018060133	0,955905719	1,071970686	1,059697636	1,041090109	1,045396695	1	0,908534116	0,988204467	1,043031986	0,963838181	1,014501044
m9	1,18498	1,120552454	1,052140698	1,179890405	1,166381777	1,145900953	1,150641101	1,100674133	1	1,087691094	1,148038327	1,060871754	1,116635057
m10	1,089445	1,030212084	0,967315724	1,084766081	1,072346535	1,053516908	1,057874894	1,011936329	0,919378678	1	1,055481959	0,975342871	1,026610462
m11	1,032178	0,976058402	0,916468269	1,027744787	1,015978081	0,998138238	1,002267149	0,958743369	0,871051059	0,947434479	1	0,924073465	0,972646148
m12	1,116987	1,056256282	0,991769923	1,112189481	1,099455963	1,080150309	1,084618472	1,037518558	0,942621006	1,025280473	1,082165042	1	1,05256366
	1,061206	1,003508216	0,942242223	1,056648184	1,044550662	1,026209007	1,030454036	0,985706231	0,895547738	0,9740793	1,028123127	0,9500613	

3.4.8 Harmonic means of the rows used to compute the GEKS estimate of the data in Table 2.2.4.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	hm_rows
m1	1	0,946854351	0,88849477	0,996203747	0,984998098	0,967963403	0,972178549	0,930270602	0,844505151	0,918993646	0,969121326	0,895339074	0,940423786
m2	1,056129	1	0,938364775	1,05211931	1,040284704	1,022293874	1,02674561	0,982485428	0,891906078	0,970575511	1,02351679	0,945593241	0,993208496
m3	1,125499	1,065683651	1	1,121226349	1,108614402	1,089441869	1,094186011	1,047018658	0,950489726	1,034326455	1,09074511	1,007703257	1,058446056
m4	1,003811	0,950462547	0,891880575	1	0,988751649	0,97165204	0,975883248	0,933815602	0,847723323	0,922495673	0,972814375	0,89875096	0,944007477
m5	1,01523	0,961275309	0,902026889	1,011376315	1	0,98270586	0,986985204	0,944438983	0,857367291	0,932990274	0,983881418	0,908975435	0,954746808
m6	1,033097	0,978192304	0,917901201	1,029175012	1,017598491	1	1,004354654	0,961059684	0,872455662	0,949409495	1,001196247	0,924972031	0,971548907
m7	1,028618	0,973951084	0,91392139	1,024712743	1,013186415	0,995664227	1	0,956892747	0,868672892	0,945293071	0,996855287	0,920961562	0,967336491
m8	1,074956	1,017826801	0,955092817	1,070875232	1,058829652	1,040518104	1,0450492	1	0,907805911	0,987877768	1,04176282	0,962450143	1,010914226
m9	1,184125	1,121194288	1,052089226	1,179630161	1,166361267	1,146190052	1,151181313	1,101557049	1	1,088203719	1,147561178	1,06019374	1,113579692
m10	1,088147	1,030316537	0,966812746	1,084015925	1,071822534	1,053286285	1,057872982	1,01227084	0,918945582	1	1,054546275	0,974260854	1,023319138
m11	1,031863	0,977023543	0,916804477	1,027945336	1,016382647	0,998805183	1,003154634	0,959911393	0,871413236	0,948275125	1	0,923866858	0,970388083
m12	1,116895	1,05753717	0,992355629	1,112655278	1,100139742	1,081113771	1,085821647	1,039014859	0,94322383	1,02641968	1,08240705	1	1,050354902
	1,06335	1,006837944	0,944781261	1,059313644	1,047398113	1,029284263	1,03376644	0,989203608	0,898004882	0,977212253	1,030515541	0,952059155	

3.4.9 Median of the columns used to compute the GEKS estimate of the data in Table 2.2.4.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	col_median_
m1	1	0,936505292	0,887247778	0,992883937	0,982681629	0,963425309	0,969640639	0,920578175	0,839299568	0,908182416	0,960254628	0,885054115	0,94731096
m2	1,067799625	1	0,947402844	1,060201096	1,049307075	1,028745183	1,035381911	0,98299303	0,896203764	0,969756844	1,025359532	0,945060452	1,011538288
m3	1,127080873	1,055517208	1	1,119060495	1,107561669	1,085858238	1,092863419	1,037566053	0,94595849	1,023595031	1,082284625	0,997527564	1,067696064
m4	1,007167064	0,943217286	0,89360674	1	0,989724572	0,97033024	0,976590116	0,927176018	0,845314882	0,914691418	0,967136835	0,891397354	0,954100898
m5	1,017623582	0,953009871	0,902884262	1,010382109	1	0,980404314	0,986729181	0,93680206	0,854091034	0,924187844	0,977177755	0,900651939	0,964006972
m6	1,037963183	0,972058014	0,920930527	1,030576972	1,019987352	1	1,006451285	0,955526253	0,871162052	0,942659912	0,99670895	0,918653586	0,983273899
m7	1,031309909	0,965827188	0,915027425	1,023971043	1,013449302	0,993590068	1	0,949401394	0,865577962	0,936617526	0,990320114	0,912765079	0,97697118
m8	1,086273852	1,017301211	0,963794061	1,078543859	1,067461359	1,046543721	1,053295272	1	0,911709175	0,986534812	1,043099494	0,961411142	1,029039125
m9	1,191469694	1,115817674	1,057128838	1,182991121	1,17083538	1,147892058	1,155297436	1,096840996	1	1,082071826	1,144114288	1,054515155	1,128692299
m10	1,101100376	1,031186329	0,976948861	1,093264876	1,082031111	1,060827969	1,067671673	1,013648974	0,92415307	1	1,057336732	0,974533418	1,043084454
m11	1,041390451	0,975267668	0,923971363	1,033979851	1,023355265	1,003301916	1,009774503	0,95868132	0,874038556	0,945772496	1	0,921686903	0,986520587
m12	1,129874415	1,058133369	1,002478564	1,121834158	1,110306831	1,088549607	1,09557215	1,040137727	0,948303109	1,026132077	1,084967136	1	1,070342416
	1,055619583	0,988593326	0,936596129	1,048107727	1,037337972	1,017010622	1,023571647	0,971780349	0,88598106	0,958695143	1,01366359	0,934280455	

3.4.10 Median of the rows used to compute the GEKS estimate of the data in Table 2.2.4.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	row_median_
m1	1	0,936432367	0,886951377	0,992829867	0,982707217	0,963441169	0,969589563	0,920588892	0,839295514	0,908232349	0,9601489	0,884992611	0,947363317
m2	1,067882781	1	0,947160103	1,060225919	1,049416116	1,028842235	1,035407999	0,983081026	0,896269227	0,969885687	1,025326478	0,945068371	1,011672974
m3	1,127457521	1,055787714	1	1,1193735	1,107960643	1,086238992	1,093171045	1,03792487	0,946270039	1,023993392	1,082527098	0,997791575	1,068111897
m4	1,007221916	0,943195202	0,893356865	1	0,989804246	0,97039906	0,976591857	0,927237307	0,845356835	0,914791526	0,967083014	0,891383953	0,954205095
m5	1,017597085	0,952910847	0,902559136	1,010300778	1	0,980394925	0,986651513	0,936788574	0,854064669	0,924214591	0,977044722	0,900565902	0,96408415
m6	1,037946096	0,971966319	0,920607719	1,030503884	1,01999712	1	1,006381702	0,955521647	0,871143502	0,942696221	0,996582802	0,918574626	0,983312056
m7	1,031364237	0,965802853	0,91476993	1,023969218	1,013529079	0,993658766	1	0,94946246	0,865619377	0,936718364	0,990263238	0,912749729	0,977076645
m8	1,086261206	1,017210152	0,963460872	1,078472568	1,067476727	1,046548766	1,053227528	1	0,911694157	0,986577567	1,042972502	0,961333141	1,029084019
m9	1,19147545	1,115736175	1,056780791	1,182932412	1,170871524	1,1479165	1,155242161	1,096859065	1	1,082136547	1,143993843	1,054446969	1,128760134
m10	1,10103984	1,031049343	0,976568801	1,093145237	1,081999797	1,06078711	1,067556737	1,013606046	0,924097798	1	1,057162191	0,974412122	1,043084755
m11	1,041505125	0,975299109	0,923764404	1,034037394	1,023494603	1,003428915	1,009832499	0,958798049	0,874130579	0,945928646	1	0,92172434	0,98668375
m12	1,129952937	1,058124503	1,002213313	1,121851024	1,110412906	1,088643178	1,095590674	1,040222123	0,948364431	1,02625981	1,08492307	1	1,070475962
	1,055561242	0,988461712	0,936231497	1,047992728	1,037307651	1,016971157	1,023461164	0,971737955	0,885927815	0,958694867	1,013495966	0,9341639	

Example: abstract GEKS

Now we consider the case an otherwise specified averaging method is introduced for averaging rows or columns. It is also possible that the averaging method for the rows is different for that of the columns.

We start with Table 3.4.7, where a PIM is shown together with row and columns averages, computed in one way are another (that's, is using some averaging method, not necessarily a geometric average).

3.4.11 PIM with row and column averages

P_{11}	$\cdots$	P_{1n}	ρ_1
$\vdots$	$\ddots$	$\vdots$	$\vdots$
P_{n1}	$\cdots$	P_{nn}	ρ_n
κ_1	…	κ_n	

In Table 3.4.8 is the GEKS-like price index shown, derived from Table 3.4.7, based on the row and column averages. These averages are determined in one way or another, and not necessarily using geometric averages as in standard GEKS. In order for this to be a price index we require that

$$\rho_i \kappa_i = 1, \qquad (3.1)$$

for $i = 1, \dots, n$.

3.4.12 GEKS-like estimator for PIM of Table 3.4.11.

$\rho_1 \kappa_1 = 1$	$\cdots$	$\rho_1 \kappa_n$	ρ_1
$\vdots$	$\ddots$	$\vdots$	$\vdots$
$\rho_1 \kappa_n$	…	$\rho_n \kappa_n = 1$	ρ_n
κ_1	…	κ_n	

We can write the interior of Table 3.4.12 more succinctly as follows:

$$\begin{pmatrix} \rho_1\kappa_1 & \cdots & \rho_1\kappa_n \\ \vdots & \ddots & \vdots \\ \rho_1\kappa_n & \cdots & \rho_n\kappa_n \end{pmatrix} = \begin{pmatrix} \kappa_1 \\ \vdots \\ \kappa_n \end{pmatrix} (\rho_1, \dots, \rho_n) = \begin{pmatrix} \kappa_1 \\ \vdots \\ \kappa_n \end{pmatrix} \left(\frac{1}{\kappa_1}, \dots, \frac{1}{\kappa_n}\right) = \begin{pmatrix} \frac{1}{\rho_1} \\ \vdots \\ \frac{1}{\rho_n} \end{pmatrix} (\rho_1, \dots, \rho_n). \qquad (3.2)$$

This shows clearly that the method depends on n row or column averages.

We can define two other GEKS-like price indices, using either row or column matrices. These are shown in Tables 3.4.13 and 3.4.14.

3.4.13 Result of the GEKS-like method applied to the row averages.

1	$\frac{\rho_1}{\rho_2}$	$\cdots$	$\frac{\rho_1}{\rho_{n-1}}$	$\frac{\rho_1}{\rho_n}$	ρ_1
$\frac{\rho_2}{\rho_1}$	1	$\cdots$	$\cdots$	$\frac{\rho_2}{\rho_n}$	ρ_2
$\vdots$	$\cdots$	$\ddots$	$\cdots$	$\vdots$	$\vdots$
$\frac{\rho_{n-1}}{\rho_1}$	$\cdots$	$\cdots$	1	$\frac{\rho_{n-1}}{\rho_n}$	ρ_{n-1}
$\frac{\rho_n}{\rho_1}$	$\frac{\rho_n}{\rho_2}$	$\cdots$	$\frac{\rho_n}{\rho_{n-1}}$	1	ρ_n
$\frac{1}{\rho_1}$	$\frac{1}{\rho_2}$	$\cdots$	$\frac{1}{\rho_{n-1}}$	$\frac{1}{\rho_n}$	

3.4.14 Result of the GEKS-like method applied to the column averages.

1	$\frac{\kappa_2}{\kappa_1}$	$\cdots$	$\frac{\kappa_{n-1}}{\kappa_1}$	$\frac{\kappa_n}{\kappa_1}$	$\frac{1}{\kappa_1}$
$\frac{\kappa_1}{\kappa_2}$	1	$\cdots$	$\cdots$	$\frac{\kappa_n}{\kappa_2}$	$\frac{1}{\kappa_2}$
$\vdots$	$\cdots$	$\ddots$	$\cdots$	$\vdots$	$\vdots$
$\frac{\kappa_1}{\kappa_{n-1}}$	$\frac{\kappa_2}{\kappa_{n-1}}$	$\cdots$	1	$\frac{\kappa_n}{\kappa_{n-1}}$	$\frac{1}{\kappa_{n-1}}$
$\frac{\kappa_1}{\kappa_n}$	$\frac{\kappa_2}{\kappa_n}$	$\cdots$	$\frac{\kappa_{n-1}}{\kappa_n}$	1	$\frac{1}{\kappa_n}$
κ_1	κ_2	$\cdots$	κ_{n-1}	κ_n	

As a rule, the results shown in Tables 3.4.13 and 3.4.14 will be different. The question is only how much.

One can use the matrices in Tables 3.4.13 or 3.4.14 as input for another round of 'GEKSification'. We use geometric averaging in our example. What happens can be illustrated by taking one case of these two, say the one represented in Table 3.4.13. We then find the results as presented in Table 3.4.15. Here $\rho = \left(\prod_{i=1}^{n} \rho_i\right)^{\frac{1}{n}}$, the geometric average of the column averages.

3.4.15 Table 1. Marginals for the elements of Table 3.4.13.

1	$\frac{\rho_1}{\rho_2}$	$\cdots$	$\frac{\rho_1}{\rho_{n-1}}$	$\frac{\rho_1}{\rho_n}$	$\frac{\rho_1}{\rho}$
$\frac{\rho_2}{\rho_1}$	1	$\cdots$	$\cdots$	$\frac{\rho_2}{\rho_n}$	$\frac{\rho_2}{\rho}$
$\vdots$	$\cdots$	$\ddots$	$\cdots$	$\vdots$	$\vdots$
$\frac{\rho_{n-1}}{\rho_1}$	$\cdots$	$\cdots$	1	$\frac{\rho_{n-1}}{\rho_n}$	$\frac{\rho_{n-1}}{\rho}$
$\frac{\rho_n}{\rho_1}$	$\frac{\rho_n}{\rho_2}$	$\cdots$	$\frac{\rho_n}{\rho_{n-1}}$	1	$\frac{\rho_n}{\rho}$
$\frac{\rho}{\rho_1}$	$\frac{\rho}{\rho_2}$	$\cdots$	$\frac{\rho}{\rho_{n-1}}$	$\frac{\rho}{\rho_n}$	

If we use the marginals in Table 3.4.15 to calculate the corresponding GEKS coefficients, we find, after simplification, exactly same the results. This shows that the GEKS method applied twice to an index yields the same index after applying the method once.

If instead of a geometric average we would have taken an arithmetic or an harmonic average, and used either the row or column averages, we would have reached a similar conclusion. In Table 3.4.16 the results is shown, were $\bar{\rho} = \sum_{i=1}^{n} \frac{1}{\rho_i}$ in case arithmetic averages were used to calculate the row averages, or $\bar{\rho} = \sum_{i=1}^{n} \frac{1}{\rho_i}$ in case the harmonic mean was used. In fact, the details of the averaging method are not important here. Important is that each row average is the product of the original row average (the ρ_i's) and a positive constant (that this should be an average is not important for the result). The reciprocals of these values should be put in the bottom row. Using these values will also yield the same matrix (of row average ratios).

3.4.16 Matrix with arithmetical column averages (right-most column) and their reciprocals (bottom row).

1	$\frac{\rho_1}{\rho_2}$	$\cdots$	$\frac{\rho_1}{\rho_{n-1}}$	$\frac{\rho_1}{\rho_n}$	$\rho_1\bar{\rho}$
$\frac{\rho_2}{\rho_1}$	1	$\cdots$	$\cdots$	$\frac{\rho_2}{\rho_n}$	$\rho_2\bar{\rho}$
$\vdots$	$\cdots$	$\ddots$	$\cdots$	$\vdots$	$\vdots$
$\frac{\rho_{n-1}}{\rho_1}$	$\cdots$	$\cdots$	1	$\frac{\rho_{n-1}}{\rho_n}$	$\rho_1\bar{\rho}$
$\frac{\rho_n}{\rho_1}$	$\frac{\rho_n}{\rho_2}$	$\cdots$	$\frac{\rho_n}{\rho_{n-1}}$	1	$\rho_1\bar{\rho}$
$\frac{1}{\rho_1\bar{\rho}}$	$\frac{1}{\rho_2\bar{\rho}}$	$\cdots$	$\frac{1}{\rho_{n-1}\bar{\rho}}$	$\frac{1}{\rho_n\bar{\rho}}$	

3.5 Dealing with missing index values

As several examples above have already illustrated, the GEKS method can be extended in such a way that it is not required that all price indices are present in the input PIM. If one has (nontrivial) averaging periods for each month one can calculate averages and compute with these indices for the entire period, for every pair of months, even those for which there was no input. In fact, one can represent a transitive index with a minimum amount of information. A PIDG corresponding to a spanning tree is enough.

We illustrate the use of GEKS here in case the PIDG is incomplete, meaning that for a subset of the pairs of months the price indices are known. We assume that the pattern of missings is symmetric with respect to the main diagonal. This means that if for the pair of months (i, j) there is no index, there is also no index for the pair (j, i) either. The idea is that if p_{ij} is known, so is $p_{ji} = \frac{1}{p_{ij}}$.

In Table 3.5.1 an incomplete PIDG is shown, derived from Table 2.2.4. The 'holes' is the matrix (corresponding to missing price indices) are symmetric, with respect to the main diagonal.

3.5.1 Incomplete PIDG with geometric row and column averages.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,940320492	0,889363527	0,992824078	0,95563071		0,961536407	0,929931164		0,938004494	0,974491835		0,954122746
m2	1,063467199	1	0,955660045	1,059086442	1,042946219	1,023345947		0,978198886	0,892849386		1,040803194		1,00482588
m3	1,124399606	1,046397205	1	1,122143269	1,110306886		1,09895587	1,025155067		1,047036886	1,103998423	1,025477886	1,069482279
m4	1,007227788	0,944209991	0,891151806	1	0,988918722	0,964200199		0,920599818	0,843748391		0,980191886		0,947465524
m5	1,035592582	0,958822211	0,900650272	1,011205449	1	0,969246089	0,97066474	0,928774893	0,852098167	0,947910964		0,919789433	0,952732738
m6		0,977186652		1,037129013	1,031729724	1	0,989437461			0,959210038		0,923845172	0,987668204
m7	1,040002222		0,909954647		1,030221826	1,010675297	1	0,943199754	0,861360312		0,987227857	0,907737195	0,963770456
m8	1,075348412	1,022286995	0,975462183	1,086248314	1,076687158		1,060220803	1	0,912085772	0,925342441	1,035881042		1,015171809
m9		1,120009731		1,185187446	1,173573702		1,160954349	1,096388115	1	1,089084506	1,137510538	1,04005003	1,109847538
m10	1,066092974		0,95507619		1,054951401	1,042524536		1,080681006	0,918202393	1	1,026698709	0,947102368	1,008640524
m11	1,026175864	0,960796437	0,905798396	1,020208404			1,012937381	0,965361812	0,879112735	0,973995576	1	0,921391606	0,96535981
m12			0,975155109		1,087205358	1,082432458	1,101640436		0,961492208	1,055852075	1,085314858	1	1,042296696
	1,048063178	0,995197297	0,935031855	1,055447375	1,0496123	1,012485768	1,037591466	0,985054985	0,901024669	0,991433495	1,03588319	0,959419808	

3.5.2 Estimates on the basis of the geometric averages in Table 3.2.9.

	m1	m2	m3	m4	m5	m6	m7	m8	m9	m10	m11	m12	
m1	1	0,949540378	0,892135161	1,007026348	1,00145897	0,966035701	0,989989619	0,93986332	0,859688122	0,945949249	0,988359714	0,915404262	0,954122746
m2	1,053141102	1	0,939544207	1,060540838	1,054677603	1,017371903	1,042598758	0,989808692	0,905372896	0,996218034	1,040882238	0,964049853	1,00482588
m3	1,120906386	1,064345874	1	1,128782264	1,122541755	1,082835587	1,109685686	1,053498797	0,963629906	1,060320554	1,107858715	1,026082483	1,069482279
m4	0,993022678	0,942915128	0,885910447	1	0,994471468	0,959295358	0,983082142	0,93330559	0,853689801	0,939349056	0,981463609	0,909017191	0,947465524
m5	0,998543156	0,948157045	0,890835459	1,005559267	1	0,964628337	0,988547358	0,938494085	0,85843569	0,944571148	0,986919828	0,91407066	0,952732738
m6	1,03515843	0,982924727	0,923501233	1,042431813	1,036668695	1	1,0247961	0,972907439	0,889913407	0,97920734	1,02310889	0,947588439	0,987668204
m7	1,010111602	0,959141753	0,901156077	1,017208998	1,011585325	0,97580387	1	0,949366844	0,868380946	0,955514311	0,998353614	0,924660466	0,963770456
m8	1,063984496	1,01029624	0,94921798	1,07146042	1,065536817	1,027847008	1,053333605	1	0,914694833	1,006475334	1,051599412	0,973975942	1,015171809
m9	1,163212536	1,10451727	1,037742803	1,171385671	1,164909628	1,123704837	1,151568334	1,093260795	1	1,100340024	1,149672409	1,064809712	1,109847538
m10	1,057139166	1,003796323	0,943111021	1,064566993	1,0586815	1,021234175	1,0465568	0,993566326	0,908809984	1	1,044833764	0,967709698	1,008640524
m11	1,011777378	0,960723474	0,902642175	1,018886478	1,013253531	0,977413069	1,001649101	0,950932445	0,869812994	0,957090051	1	0,926185324	0,96535981
m12	1,092413529	1,037290755	0,97458052	1,100089206	1,094007327	1,055310469	1,081478053	1,026719406	0,939134985	1,033367757	1,079697523	1	1,042296696
	1,048063178	0,995197297	0,935031855	1,055447375	1,0496123	1,012485768	1,037591466	0,985054985	0,901024669	0,991433495	1,03588319	0,959419808	

Geometric row and column averages have also been given, in the right-most column and the bottom line.[7]

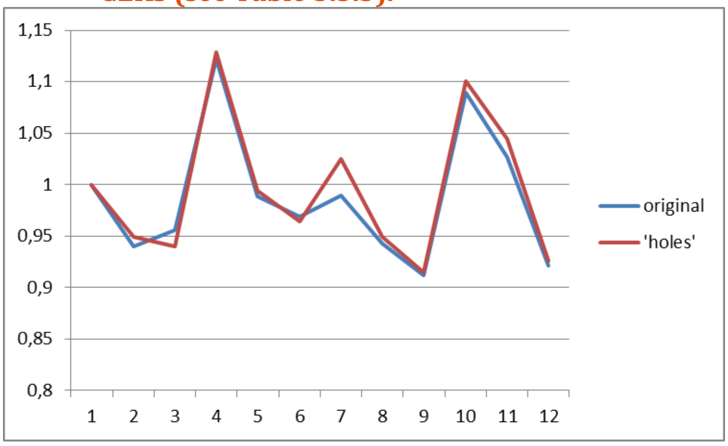
Table 3.5.2 contains the estimates based on the averages in Table 3.5.1.

The method used here is based on the same ideas of the GEKS method, except that averages are now taken for the indices that are present in a month. Because of the symmetry of the holes in the PIDG the row averages are reciprocals of the column averages, as is required to produce a transitive price index as in the GEKS method.

The method above can be seen as a kind of imputation method. It is not exactly the same as it also changes original price indices. It produces a price index that is transitive. The transitivity makes sure that the missing values get a value as well.

A drawback of this 'imputation method' is that in each month in principle the entire window is used, only using those indices that are present in a row or column. One would like to work more locally to achieve better results, more geared towards the local situation. In some cases this can be achieved by partitioning the PIDG, and applying the GEKS approach to the appropriate parts of the PIDG.

3.5.3 Original data compared with the GEKS estimate based on the imputed GEKS (see Table 3.3.3).



3.6 Spatial and spatio-temporal setting

The GEKS method and variants above have been illustrated with examples from applications involving temporal indices. However, formally there is no difference between a temporal setting (concerning development of prices of articles in time) and a spatial setting (concerning price comparisons of a spatial regions (a province, a country) at a given point or period in time) or a spatio-temporal setting (involving the development in time of prices of articles in various regions or countries). In a more abstract formulation, one is in fact comparing 'states'. A state can be the price of an article or a group of articles at a point or period in time, or the average price of prices of articles in a region, etc. Comparing pairs of states produces price ratios or

[7] The example is artificial. The holes are created for the sake of the example, to illustrate how the method works. With real data one could have real holes in long periods for months that are far apart.

price indices, etc. This yields a digraph where the states form the vertices and the price comparison relation yields the arcs.

4. Updating with GEKS

So far we have assumed that the data for a time window are all present and GEKS (whatever variant) can be applied in one go. In practice, however, information appears in piecemeal fashion, and price indices are usually updated as soon as the results for the last month have been processed. Typically, the publication policy requires that results that have been published earlier cannot be revised (except for blunders, which are exceptional to occur). So if we want to update GEKS estimates, we assume that earlier GEKS estimates are also fixed. Only the new results, concerning the latest month, can be adjusted.

Updating is an issue it its own right. We only want to sketch in the present section how this can be done with the GEKS method. A full treatment is reserved for another report. The aim is to give the reader a feel of the approach by considering an example. We do not fill in every detail. But we trust that the attentive reader will understand the approach and is able to describe how it works in a similar situation.

To illustrate updating with (a generalization of) GEKS we consider an example, with symbolic data. This is chosen to see better what it going on. In this example we consider a time period of 6 months to which the information of a new month is added, month 7. The GEKS indices for the first 6 months have been calculated and are assumed to be fixed. The GEKS price indices for the first 6 months are represented by a linear PIDG, of MoM-price indices.

The situation is presented in Table 4.1.1. The known information for the first 6 months of the 6 month period are in black. They are the GEKS-indices for these months, represented by a linear PIDG. As the GEKS index is transitive the price indices for the remaining month pairs can be derived by transitive closure. The price indices related to the newest month (m7) in red. We assume that they are time reversible, that is $p_{i7} = p_{7i}$, for all $i = 1, \dots, 7$. Because the GEKS price indices for the first 6 months are fixed, we cannot expect that calculating updated GEKS indices using the new month 7 as a base or reporting month will produce the same GEKS indices. Therefore 6 correction factors f_i, $i = 1, \dots, 6$ have been introduced to correct for the differences with earlier GEKS estimates. The updated price indices involving month 7 are $f_i p_{i,7}$ and $f_i^{-1} p_{7,i}$ for its time reversal, $i = 1, \dots, 6$.

4.1.1 Update for the GEKS method with correction factors (CF).

	m1	m2	m3	m4	m5	m6	m7	CF
m1	1	$p^G_{1,2}$	–	–	–	–	$p_{1,7}$	f_1
m2	$p^G_{2,1}$	1	$p^G_{2,3}$	–	–	–	$p_{2,7}$	f_2
m3	–	$p^G_{3,2}$	1	$p^G_{3,4}$	–	–	$p_{3,7}$	f_3
m4	–	–	$p^G_{4,3}$	1	$p^G_{4,5}$	–	$p_{4,7}$	f_4
m5	–	–	–	$p^G_{5,4}$	1	$p^G_{5,6}$	$p_{5,7}$	f_5
m6	–	–	–	–	$p^G_{6,5}$	1	$p_{6,7}$	f_6
m7	$p_{7,1}$	$p_{7,2}$	$p_{7,3}$	$p_{7,4}$	$p_{7,5}$	$p_{7,6}$	1	
CF	f_1^{-1}	f_2^{-1}	f_3^{-1}	f_4^{-1}	f_5^{-1}	f_6^{-1}		

The idea of the correction factors is that they are used in the GEKS averages, and that the GEKS estimates for pairs of months (m1,m2), (m2,m3), (m3,m4), (m4,m5) and (m5,m6) yield the same estimates as before, namely $p^G_{1,2}$, $p^G_{2,3}$, $p^G_{3,4}$, $p^G_{4,5}$ and $p^G_{5,6}$, respectively (and their time reversals). This leads to a system of equations that the correction factors must satisfy:

$$
\begin{aligned}
p^G_{1,2} &= \frac{\sqrt[3]{p^G_{1,2}p_{1,7}f_1}}{\sqrt[4]{p^G_{1,2}p^G_{3,2}p_{7,2}f_2^{-1}}} = (p^G_{1,2})^{\frac{1}{12}}(p^G_{2,3})^{\frac{1}{4}}(p_{1,7})^{\frac{1}{3}}(p_{2,7})^{\frac{1}{3}}(f_1)^{\frac{1}{3}}(f_2)^{\frac{1}{4}} \\
p^G_{2,3} &= \frac{\sqrt[4]{p^G_{2,1}p^G_{2,3}p_{2,7}f_2}}{\sqrt[4]{p^G_{2,3}p^G_{4,3}p_{7,3}f_3^{-1}}} = (p^G_{2,1})^{\frac{1}{4}}(p^G_{3,4})^{\frac{1}{4}}(p_{2,7})^{\frac{1}{4}}(p_{3,7})^{\frac{1}{4}}(f_2)^{\frac{1}{4}}(f_3)^{\frac{1}{4}} \\
p^G_{3,4} &= \frac{\sqrt[4]{p^G_{3,2}p^G_{3,4}p_{3,7}f_3}}{\sqrt[4]{p^G_{3,4}p^G_{5,4}p_{7,4}f_4^{-1}}} = (p^G_{3,2})^{\frac{1}{4}}(p^G_{4,5})^{\frac{1}{4}}(p_{3,7})^{\frac{1}{4}}(p_{4,7})^{\frac{1}{4}}(f_3)^{\frac{1}{4}}(f_4)^{\frac{1}{4}} \\
p^G_{4,5} &= \frac{\sqrt[4]{p^G_{4,3}p^G_{4,5}p_{4,7}f_4}}{\sqrt[4]{p^G_{4,5}p^G_{6,5}p_{7,5}f_5^{-1}}} = (p^G_{4,3})^{\frac{1}{4}}(p^G_{5,6})^{\frac{1}{4}}(p_{4,7})^{\frac{1}{4}}(p_{5,7})^{\frac{1}{4}}(f_4)^{\frac{1}{4}}(f_5)^{\frac{1}{4}} \\
p^G_{5,6} &= \frac{\sqrt[4]{p^G_{5,4}p^G_{5,6}p_{5,7}f_5}}{\sqrt[3]{p^G_{5,6}p_{7,6}f_6^{-1}}} = (p^G_{5,4})^{\frac{1}{4}}(p^G_{5,6})^{-\frac{1}{12}}(p_{5,7})^{\frac{1}{4}}(p_{6,7})^{\frac{1}{3}}(f_5)^{\frac{1}{4}}(f_6)^{\frac{1}{3}}
\end{aligned}
\tag{4.1}
$$

4.1.2 Table with GEKS coefficients (GC) based on Table 4.1.1.

	m1	m2	m3	m4	m5	m6	m7	CF	GC
m1	1	$p^G_{1,2}$	–	–	–	–	$p_{1,7}$	f_1	$\sqrt[3]{p^G_{1,2}p_{1,7}f_1}$
m2	$p^G_{2,1}$	1	$p^G_{2,3}$	–	–	–	$p_{2,7}$	f_2	$\sqrt[4]{p^G_{2,1}p^G_{2,3}p_{2,7}f_2}$
m3	–	$p^G_{3,2}$	1	$p^G_{3,4}$	–	–	$p_{3,7}$	f_3	$\sqrt[4]{p^G_{3,2}p^G_{3,4}p_{3,7}f_3}$
m4	–	–	$p^G_{4,3}$	1	$p^G_{4,5}$	–	$p_{4,7}$	f_4	$\sqrt[4]{p^G_{4,3}p^G_{4,5}p_{4,7}f_4}$
m5	–	–	–	$p^G_{5,4}$	1	$p^G_{5,6}$	$p_{5,7}$	f_5	$\sqrt[4]{p^G_{5,4}p^G_{5,6}p_{5,7}f_5}$
m6	–	–	–	–	$p^G_{6,5}$	1	$p_{6,7}$	f_6	$\sqrt[3]{p^G_{6,5}p_{6,7}f_6}$
m7	$p_{7,1}$	$p_{7,2}$	$p_{7,3}$	$p_{7,4}$	$p_{7,5}$	$p_{7,6}$	1		
CF	f_1^{-1}	f_2^{-1}	f_3^{-1}	f_4^{-1}	f_5^{-1}	f_6^{-1}			
GC	$\sqrt[3]{p^G_{2,1}p_{7,1}f_1^{-1}}$	$\sqrt[4]{p^G_{1,2}p^G_{3,2}p_{7,2}f_2^{-1}}$	$\sqrt[4]{p^G_{2,3}p^G_{4,3}p_{7,3}f_3^{-1}}$	$\sqrt[4]{p^G_{3,4}p^G_{5,4}p_{7,4}f_4^{-1}}$	$\sqrt[4]{p^G_{4,5}p^G_{6,5}p_{7,5}f_5^{-1}}$	$\sqrt[3]{p^G_{5,6}p_{7,6}f_6^{-1}}$			

From the system (4.1) we derive a linear system by taking (natural) logarithms:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & \frac{1}{4} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{4} & \frac{1}{3} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}, \tag{4.2}$$

where the a_i , for $i = 1, \ldots ,5$ are expressions in terms of the GEKS indices and $x_j = \log f_j$, for $j = 1, \ldots ,6$. So we have 5 linear equations with 6 unknowns. We can write the 'solution' of (4.2) – the x's - in terms of a single parameter x_5. This is easy to see if we start with the final equation and work backwards to the first one. To simplify matters we first multiply both sides of (4.2) by 12. We than can write the final equation as

$$12a_5 = 3x_5 + 4x_6, \tag{4.3}$$

from which we obtain

$$x_6 = 3a_5 - \tfrac{3}{4}x_5. \tag{4.4}$$

Taking the one but last equation

$$12a_4 = 3x_4 + 3x_5, \tag{4.5}$$

We obtain

$$x_4 = 4a_5 - x_5. \tag{4.6}$$

If we continue in this way, and using previously obtained results, we get expressions for x_3, x_2 and x_1 in terms of x_5. So if we want a single solution we should choose a value for x_5. How should we do this? We should realize that a choice of this parameter determines the CF's. So we should control the perturbation of the price indices $p_{1,7}, \ldots, p_{6,7}$, keeping in mind that the first one of these could be modified the most and the last one, the least. One can use a linear object function

$$\textstyle\sum_{j=1}^{6} w_j x_j, \tag{4.7}$$

with $w_j > 0$ for $j = 1, \ldots ,6$, to formally express the loss of the perturbations to each of the price indices $p_{1,7}, \ldots, p_{6,7}$. It is reasonable to assume that a price index $p_{i,j}$ is more reliable, and hence should be less perturbed, the smaller $|j - i|$ is, that is the closer the base and reporting month. This would translate into weights that obey $w_1 < \cdots < w_6$. How these weights should actually be chosen requires some 'playing with the data' to get some feeling for the consequences of various choices. Minimizing function (4.7), which is actually a function of x_5, yields an optimal value x_5^*. From this we obtain in particular the optimal f_6^*value for f_6 . This yields the corrected value $f_6^* p_{6,7}$ that will be taken instead of the direct estimate $p_{6,7}$. Note that the factor f_6^* contains information from the 'observed' price indices $p_{1,7}, \ldots, p_{6,7}$.

Remark With a similar reasoning one would be able to arrange the updating of the RYGEKS (Rolling year GEKS), which uses a sliding window of one year. We shall not elaborate this case here. It should be clear how it roughly works. ■

Remark Updating with the cycle method uses similar ideas: update only the indices involving the newest arcs. However, we will also not elaborate this subject here. The interested reader is referred to Willenborg (2017) instead. ■

5. On to the cycle method

The GEKS method is a bit of a trick to transitivize price indices. At best one can understand the heuristic reasoning behind (2.3). But personally I find the motivation unsatisfactory. There is also no control over how close the transitivized price index should be to the original price index.

It would be more attractive to derive it from some model, such as an optimization model. In the present section we formulate an optimization model, using certain ideas from the GEKS method. The optimization model we present yields a variant of the GEKS method. In fact, it is formally similar to an optimization model leading to the time product dummy index.

5.1 Loglinear models

As a starting point for the model we take the equation (2.6) and (2.7) and forget about the meaning of the parameters they have there. We have

$$P_{ij}^G = \frac{\kappa_j}{\kappa_i} = \frac{\rho_i}{\rho_j}. \tag{5.1}$$

Taking logarithms we obtain from (5.1):

$$\log P_{ij}^G = \log \kappa_j - \log \kappa_i = \log \rho_i - \log \rho_j, \tag{5.2}$$

which is an expression linear in the logarithms of the κ's or the ρ's. Which formulation we use of the two is irrelevant, of course, as they are equivalent.

Inspired by (5.2) we now consider the following simple loglinear model to describe the price indices as functions of certain parameters z_i to be determined

$$\log P_{ij} = z_i - z_j + \varepsilon_{ij}, \tag{5.3}$$

where the ε_{ij}'s are error terms. An optimization model to estimate the z's could be the following one, minimizing the sum of the squared error terms:

$$\min \sum_{i,j}^{n} \varepsilon_{ij}^2 = \min \sum_{i,j=1}^{n} \left(y_{ij} - z_i + z_j\right)^2, \tag{5.4}$$

where we have written $y_{ij} = \log P_{ij}$. In case the errors terms are weighted, we could obtain the following optimization problem:

$$\min \sum_{i,j}^{n} w_{ij}\varepsilon_{ij}^2 = \min \sum_{i,j=1}^{n} w_{ij}\left(y_{ij} - z_i + z_j\right)^2. \tag{5.5}$$

Of course, there is no particular reason to choose (weighted) quadratic sums of the error terms to quantify the ‘costs’, except perhaps for computational reasons (‘ease of computation’). But other choices are also possible, and none of them is privileged. By choosing different metrics one gets different solutions, and hence different GEKS-like indices.

In the models (5.4) and (5.5) it is tacitly assumed that for each pair of months (i, j) an index $P_{ij}$is given, as in case of GEKS. But in fact there is no compelling reason to do so. If the indices are given for a subset of all possible combinations, a variant of the optimization model (5.5) can still be defined. In fact, this leads to the cycle model, as the next section tries to elucidate.

5.2 Cycle method

At first sight the GEKS method and the cycle method seem completely different. However, the previous section points at the link between the two methods. The cycle model generalizes the optimization models that are inspired by GEKS and discussed in the previous section, in particular model (5.5), which includes model (5.4) as a special case, of course.

We assume that the ordered pairs of months for which the indices are known are defined by a PIDG. Let A be the set of arcs in the PIDG, and V be the set of points (months, say, in the temporal setting). Now instead of (5.5) we assume the optimization model

$$\min \sum_{(i,j)\in A} w_{ij}\varepsilon_{ij}^2 = \min \sum_{(i,j)\in A} w_{ij}\left(y_{ij} - z_i + z_j\right)^2. \tag{5.6}$$

Note that in this formulation there is no talk of cycles. It seems as they are not needed. But in fact they are present, but implicitly. In fact, the sum in (5.6) is more complicated than it looks at first sight. The sum is over a subset of ordered pairs of indices, namely those pairs that define the arcs in the PIDG that is being used.

The formulation at the basis of optimization model (5.6) is called the indirect formulation of the cycle method in Willenborg (2010), because it does not directly deal with the price indices associated with the arcs A of the PIDG but with its vertices. However, one can reformulate (5.6) in terms of the price indices to be adjusted. First, let $x_{ij} = z_i - z_j$, so the sum in (5.6) can be rewritten as follows

$$\sum_{(i,j)\in A} w_{ij}\left(y_{ij} - z_i + z_j\right)^2 = \sum_{(i,j)\in A} w_{ij}\left(y_{ij} - x_{ij}\right)^2. \tag{5.7}$$

The x_{ij}’s can be viewed as the adjusted y_{ij}’s. They have a nice property. Suppose that we start in vertex i on the PIDG. Associated with it is the as yet unknown z_i. Let Γ be a cycle which contains vertex i. Starting from this vertex and walking around in one direction implies that a sequence of vertices in the PIDG is being traversed, in the order specified: $(j_0, j_1, \ldots, j_{n-1}, j_l)$, with $j_0 = j_l = i$ and such that (j_{i-1}, j_i) is an arc in the PIDG for $j = 1, \ldots, l$. We then have

$$\begin{aligned} z_i &= z_{j_0} = z_{j_0} - z_{j_1} + z_{j_1} + \cdots + z_{j_{l-1}} - z_l + z_l \\ &= x_{j_0 j_1} + \cdots + x_{j_{l-1}, j_l} + z_l = x_{i j_1} + \cdots + x_{j_{l-1} i} + z_i, \end{aligned} \tag{5.8}$$

from which it follows that

$$x_{i j_1} + \cdots + x_{j_{l-1} i} = 0. \tag{5.9}$$

That is, the added x-values associated with the arcs on Γ equal 0. This is a condition on the x's that holds for any cycle in the PIDG. The question now arises whether we have to require this for all cycles in the PIDG? The answer is: No, it is sufficient to require it for certain subsets of cycles. These subsets are complete sets of so-called elementary cycles. They can be considered as bases of a finite dimensional vector space. They are usually not unique, as is the case with bases of vector spaces. Any cycle in the PIDG can be written as the 'sum' of a finite number of elementary cycles. The 'sum' is a special addition operator called the ring sum.

At this point we refer the interested reader to Willenborg (2016) for further information on the cycle method, in order not to duplicate the discussion. In this reference the cycle method is explained in detail, but not starting from the GEKS estimator, as in the present document. Among other thing, it is shown there how an complete set of elementary cycles can be found (using spanning trees).

6. Discussion

The motivation for writing the present report was a research effort at CBS where a number of price index methods have been applied to a handful of scanner data sets, for direct comparison. GEKS and the cycle method were among these methods. But they were somewhat difficult to apply, as no software was readily available to do this. I set myself the task to do this, using Excel and R. For the cycle method an R script has been developed, but for a specific problem.

When I started writing this report, GEKS was a method that I knew but was not familiar with. I knew how to apply it but I did not quite understand why it works. Why does it produce transitive indices? Why is it a useful method for this purpose? Earlier, I had developed the cycle method for the same purpose. This method I did understand (obviously), but the GEKS method (of which I had read in Balk, 2008), was a bit of a mystery to me. I also had a faint idea that it was a special case of the cycle method, but I did not understand how exactly.

Writing the present report was instrumental for this and led to my understanding of both of the GEKS method and its relation to the cycle method. Once I understood the GEKS method, it turned out to be easy to compute it in applications using a spreadsheet program (like Excel), and to come up with all sorts of generalizations. One of these generalizations led to the link to the cycle method, as I had suspected to exist.

For the cycle method it was clear that it can be applied in the spatial, temporal and spatio-temporal case. This is also the case with GEKS, although I myself am inclined to see it mainly as a method best suited for temporal applications. This does not mean that GEKS should be applied in all cases, as it has some restrictions, some of them pointed out below.

I hope that the present report will be useful for others who know the GEKS method, but do not know the cycle method. As such it should be a gentle introduction to the cycle method, as gentle as I can hope to produce. And even if they are not interested in the cycle method, I hope that there are some results about GEKS and some of its variants that will spark their interest and increase their understanding of the method. And that reading the present report will increase their understanding of GEKS, as it did in my case writing it.

I also hope that the present report conveys that the GEKS method is a nice method, but not a perfect one. For me, the nicest thing about the method is that it seems to be a first step in the right direction, but that further steps should be taken to get a fully satisfactory method. To wit, the GEKS method is unnecessarily restrictive in requiring that all price indices in a time period (or between all pairs of countries in a spatial setting) should be available. The present report shows that it is possible to produce variants of GEKS that work well when only a subset of the full set of price indices is known. This may result in a substantial reduction of effort to produce the input price index numbers: instead of being quadratic in the number of states (months, countries) it is linear number in the number of states. In case of comparing price parity of countries this may be a significant time saver. Also the fact that each index has equal weight in the GEKS method is not very attractive. One would like to differentiate between indices in terms of the distance of the states, say the number of months separating the base period from the reporting period. Those with a small separation should weigh more in the averaging method than those more distant. If they are too distant one would preferably not use them at all. The paper also shows that the choice in GEKS to use geometric averaging is not the secret behind GEKS. Also other averaging methods can be used. What makes that GEKS (and generalizations) work is that ratios of averages are used.

In the recent years that I have been working and occasionally teaching in the area of price index numbers, I have become convinced that a good price index should be transitive. Many of the well-known classical price indices (Laspeyres, Paasche, Fisher, Törnqvist) are not. In my view these indices can be taken as a starting point and they can then be transitivized by a GEKS-like method or the cycle method. By applyng a transitivizing method to a nontransitive index one can compare the original values to the adjusted ones and find out whether they do not deviate too much. If they do, the original index is perhaps better dismissed as untrustworthy. If an index is transitive it can be very compactly represented, for instance only by the MoM price indices. By transitive closure all other indices in the period can be computed.

In the report updating for GEKS is also considered. The conclusion is that it is possible, but not worth the trouble, as the simplicity of GEKS is lost. Updating is better left to the cycle method.

I consider the present paper as an effort to understand transitivizing methods for price indices. The GEKS approach is an idea for such an approach, or rather, in my view, the germ of an idea that needs further development and maturation. The present paper investigates a few aspects to this end. But there is still more to do. And there is also more to compute. For instance, to answer the questions how the methods fare for longer periods of time, and with price index series that show more extreme variation than the relatively short and tame one considered in the present report? Only computation and experimentation - preferably with real data - can tell. This is also true for the choice of suitable weights. This requires extensive experimentation with real data, in order the get a feel for the right choice of these weights. What is ‘right’ is a matter of judgement.

References

Balk, B. (2008). Price and quantity index numbers, Cambridge University Press.

Chessa, A., J. de Haan, J. Verburg & L. Willenborg (2016). Comparison of price index methods. Report, CBS, The Hague.

De Haan, J., L. Willenborg, A. Chessa & J. Verburg (2016). An overview of price index methods for scanner data, Paper for the UNECE-ILO meeting of the group of experts on consumer price indexes, 2-4 May 2016, Geneva, Switzerland.

Willenborg, L. (1993). An adjustment method based on graph homology, Report, CBS, Voorburg.

Willenborg, L. (2010). Chain indexes and path independence. Report, CBS, The Hague.

Willenborg, L. (2015). Transitivizing elementary price indexes for internet data. Report, CBS, The Hague.

Willenborg, L. (2017). Cycle method – Theory, computation and applications. Report, CBS, The Hague. (in preparation).

Explanation of symbols

Empty cell	Figure not applicable
.	Figure is unknown, insufficiently reliable or confidential
*	Provisional figure
**	Revised provisional figure
2014–2015	2014 to 2015 inclusive
2014/2015	Average for 2014 to 2015 inclusive
2014/’15	Crop year, financial year, school year, etc., beginning in 2014 and ending in 2015
2012/’13–2014/’15	Crop year, financial year, etc., 2012/’13 to 2014/’15 inclusive

Due to rounding, some totals may not correspond to the sum of the separate figures.

Colofon

Publisher
Statistics Netherlands
Henri Faasdreef 312, 2492 JP The Hague
www.cbs.nl

Prepress
Statistics Netherlands, Studio BCO

Design
Edenspiekermann

Information
Telephone +31 88 570 70 70, fax +31 70 337 59 94
Via contactform: www.cbsl.nl/information