Energy decomposition analysis: the generalized Fisher index revisited

Paul de Boer

Econometric Institute, Erasmus University Rotterdam,
P.O. Box 1738, NL 3000 DR Rotterdam, the Netherlands
e-mail: pmdeboer@few.eur.nl

Abstract

In recent papers I have proposed to use the Montgomery decomposition in case of an additive decomposition (De Boer, 2008) and the Sato-Vartia decomposition for the multiplicative case (De Boer, 2007), rather than the commonly used methods of taking the average of the two polar decompositions or the average of all decompositions (Dietzenbacher and Los, 1998).The disadvantage of the average of the two polar decompositions is that it does not satisfy factor reversal which means that the order of appearance of the factors in the decomposition matters. The average of all elementary decompositions meets this requirement, but needs the computation of $r!$ decompositions, r being the number of factors. Both the Montgomery and the Sato-Vartia decomposition, borrowed from index number theory, are ideal and require the computation of only one decomposition. The disadvantage is that they are not “change-in-sign robust”, i.e. they cannot handle values that in one period are positive and in the other negative. Ang, Liu and Chung (2004) have proposed to use a generalized Fisher index approach which is ideal and negative value robust and give the formulae for the cases of three and four factors. I show that this approach is equivalent to the method of taking the average of all elementary decompositions. The example of Chung and Rhee (2001) which deals with energy-related CO_2 emissions for seven intermediate demand sectors in the Korean economy serves as empirical application. Moreover, I give two tables from which the formulae for the cases of five and six factors are easily implemented.

Rotterdam, 2 May 2008

1. Introduction

In recent years a large number of empirical articles on structural decomposition analysis, which aims at disentangling an aggregate change in a variable into its r factors, have been published in Economic Systems Research. For an additive decomposition commonly used methods are the arithmetic average of the two polar decompositions and the arithmetic average of all elementary decompositions ($r!$), whereas for a multiplicative decomposition the corresponding geometric averages are commonly used. The advantage of the average of all elementary decompositions is that it is satisfies factor reversal, so that it is "ideal", but it requires $r!$ decompositions, whereas for the average of the two polar decompositions only two decompositions are required, but it is not ideal. In recent papers I have proposed to use the Montgomery decomposition in case of an additive decomposition (De Boer, 2008) and the Sato-Vartia decomposition for the multiplicative case (De Boer, 2007). Both decompositions, borrowed from index number theory, are ideal and require the computation of only one decomposition. The disadvantage is that they are not "change-in-sign robust", i.e. they cannot handle values that in one period are positive and in the other negative.

In the framework of a multiplicative decomposition[1] Ang, Liu and Chung (2004) have proposed to use a generalized Fisher index approach which is ideal and change-in-sign robust, but, according to Ang c.s., the formula, either based on Siegel (1945) or on Shapley (1953), is relatively complex, especially for a large number of factors r. We apply the generalized Fisher index approach to the same example they use, i.c. the example of Chung and Rhee (2001) which deals with energy-related CO_2 emissions for seven intermediate demand sectors in the Korean economy. We show that the complicated formula for the cases of $r = 3$ and $r = 4$ given by Ang c.s. are nothing but the geometric average of the $3! = 6$ and $4! = 24$ elementary decompositions, respectively.

The organization of this paper is as follows: in section two I apply the reasoning of SDA to the well-known decomposition of a change in value into changes in price and quantity. It is easily shown that the SDA approach is equivalent to the use of the Fisher indices for two factors in IDA. The formula is summarized in the form of a table which will be generalized to a higher number of factors. In section 3 I use the example of Chung and Rhee (2001) of the decomposition of energy-related CO_2 emissions for seven intermediate demand sectors in the Korean economy in order to deal with the case of three factors. Again, I show that the generalized Fisher approach is equivalent to SDA and that commonly used methods of SDA yield empirical results that are very close to each other. A summarizing table is presented, as well. Section 4 is devoted to the treatment of the four-factor case by Ang,c.s. (2004) in the framework of the same example. The difference between their approach and mine is that they did not realize that the decomposition, reading in four factors, can be reduced to a decomposition reading in three factors. I show that SDA and generalized Fisher yield the very same formulae and provide the summarizing table. In section 5 I give the summarizing tables from which the formulae can easily be derived for the cases of five (120 elementary decompositions) and six factors (720 elementary decompositions) . Section 6, finally, contains some final remarks. Last, but not least, in order to give proper credit to the contributions of Siegel and Shapley, I propose to replace the name of "generalized Fisher" or "Structural Decomposition Analysis" by "Siegel-Shapley decomposition".

2. Decomposition and index number theory: the case of two factors (price and quantity)

2.1. The Fisher index

Let $p_i(1)$ and $p_i(0)$ denote the prices of commodity i (= 1,…,n) in comparison and base period, and let $q_i(1)$ and $q_i(0)$ be the corresponding quantities. Then, the ratio of total expenditure in comparison and in base period is defined as:

$$DV[1,0] = \frac{V(1)}{V(0)} = \frac{\sum_{i=1}^{n} p_i(1)q_i(1)}{\sum_{i=1}^{n} p_i(0)q_i(0)} \tag{1}$$

In the terminology of decomposition analysis we have to decompose (1) into its factors “price” and “quantity”. One possible solution, the so-called first *polar decomposition*, is:

$$DV[1,0] = \frac{\sum_{i=1}^{n} p_i(1)q_i(1)}{\sum_{i=1}^{n} p_i(0)q_i(1)} \times \frac{\sum_{i=1}^{n} p_i(0)q_i(1)}{\sum_{i=1}^{n} p_i(0)q_i(0)} \tag{2}$$

In index number theory (omitting the commonly used factor 100) the price term is the named the **price index of Paasche** and the quantity term the **quantity index of Laspeyres,** so that:

$$DV[1,0] = P^P \times Q^L \tag{3}$$

It is easily seen that if we reverse ***base and comparison period*** (0 to 1 and 1 to 0) that for the first polar decomposition (2) generally $DV[1,0] \times DV[0,1] \neq 1$ holds true. In the terminology of index number theory, the first polar decomposition does not meet the requirement of ***time reversal***:

$$DV[1,0] \times DV[0,1] = 1$$

However, this is not the only possibility. By reversing the time periods in the ***weights*** (0 to 1, and 1 to 0) we obtain the *second polar decomposition*:

$$DV[1,0] = \frac{\sum_{i=1}^{n} p_i(1)q_i(0)}{\sum_{i=1}^{n} p_i(0)q_i(0)} \times \frac{\sum_{i=1}^{n} p_i(1)q_i(1)}{\sum_{i=1}^{n} p_i(1)q_i(0)} \tag{4}$$

In index number theory the price term is the named the **price index of Laspeyres** and the quantity term the **quantity index of Paasche** so that:

$$DV[1,0] = P^L \times Q^P \tag{5}$$

The second polar decomposition does not meet the requirement of time reversal, either.

The solution that is commonly adopted in decomposition analysis is to take the *geometric mean of the two polar decompositions* (2) and (4) which meets the requirement of time reversal. In terms of index number theory we take the geometric mean of (3) and (5), which can be written as:

$$\mathrm{DV}[1,0] = \left(\mathrm{P}^{\mathrm{P}} \times \mathrm{P}^{\mathrm{L}}\right)^{1/2} \left(\mathrm{Q}^{\mathrm{L}} \times \mathrm{Q}^{\mathrm{P}}\right)^{1/2} \tag{6}$$

The first term is the definition of the **Fisher price index** (P^{F}) and the second one the **Fisher quantity index** (Fisher, 1922). Consequently, the geometric mean of the two polar decompositions yields:

$$\mathrm{DV}[1,0] = \mathrm{P}^{\mathrm{F}} \times \mathrm{Q}^{\mathrm{F}} \tag{7}$$

It can easily shown that if in the formula of the Fisher price index, we reverse the factors (p to q and q to p) that we obtain the Fisher quantity index. Indices that exhibit this property of ***factor reversal*** are called ***"ideal".***

2.2. Summary

In view of the generalization to more than 2 factors I give the following summary.
In case of $r = 2$ factors, there are $r! = 2! = 2$ *permutations*, which are called the elementary decompositions, i.c. the polar ones: (2) and (4). Consider the first factor: "price". In (2) the quantity term in numerator and denominator is the one in the comparison period {1}; the number of duplicates is 1 (which means that in case of $r = 2$ *"permutation"* and *"combination"* are synonyms), whereas the exponent in the geometric average (6) is equal to $1/2$. In (4) the quantity term in numerator and denominator is the one in the base period {0}; the number of duplicates is 1 again, whereas the exponent in the geometric average is also equal to $1/2$. This can be summarized as follows:

Table 1. Case of two factors

Number of ones	Combinations	Number of duplicates	Exponent
1	{1}	1	½
0	{0}	1	½

If we look at the second factor, "quantity", we observe that the combinations are either {1} or {0} again, that the number of duplicates is 1, as well, whereas the exponent is also equal to $1/2$. Consequently, the table applies to both (=all) factors.

3. Decomposition and index number theory: the case of three factors (Chung and Rhee)

3.1. Decomposition analysis

Chung and Rhee (2001) made a decomposition of the sources of carbon dioxide emissions for $n = 7$ Korean industries. They gently supplied the data in their article, so that other researchers, like me, can profitably make use of their example. It reads: the emissions of CO_2 from the intermediate demand sectors, C_p, can be estimated using the input-output relation:

$$C_p = f^{'}Duy$$

where:

f : vector with typical element f_i, the amount of CO_2 emitted per unit of production in industry i;
D : Leontief inverse matrix with typical element d_{ij};
u : vector with typical element u_j, the share of industry j in final demand, and
y : gross domestic product (GDP).

The task is to apply a multiplicative decomposition of the change in the emissions from the intermediate demand sectors, C_p, into the changes in emission coefficient, $f_i(1)/f_i(0)$, in production technology, $d_{ij}(1)/d_{ij}(0)$, in the structure of the final demand, $u_j(1)/u_j(0)$, and in the size of the economy, $y(1)/y(0)$, i.e.:

$$D_{C_p} = \frac{C_{p_1}}{C_{p_0}} = \frac{f_1^{'}D_1u_1y_1}{f_0^{'}D_0u_0y_0} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)y(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)y(0)} = \frac{y(1)}{y(0)} \times \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)} \qquad (8)$$

In (8) we have already separated the contribution of the size of the economy, (y), from the remaining $r = 3$ factors, viz. emission coefficients (f_i), production technology (d_{ij}) and structure of the final demand (u_j).

There are 3! = 6 permutations (= elementary decompositions) of the second term in (8) which are given in Table 2.

Table 2. Elementary decompositions of the Chung-Lee example

e	DF_e	DD_e	DU_e
1	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$
2	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(0)}$
3	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$
4	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(1)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(0)}$
5	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(1)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(0)}$
6	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(0)d_{ij}(0)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(0)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(0)u_j(0)}$	$\dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(1)}{\sum_{i=1}^{n}\sum_{j=1}^{n} f_i(1)d_{ij}(1)u_j(0)}$

Decompositions 1 and 6 are called the *polar decompositions*; see for instance Dietzenbacher and Los (1998). In practice, quite often researchers use the geometric average of these two polar decompositions as “generalization” of the Fisher index (7) to three (or more) factors. But De Haan (2001) has argued that this is but one mirror pair (changing zeros into ones and ones into zeros). In this case there are two other mirror pairs, viz. 2 and 4, and 3 and 5. Each of them constitutes another “generalization” of the Fisher index. These two mirror pairs satisfy *time reversal*, as well. Dietzenbacher and Los (1998), finally, propose to use the average of all elementary decompositions, which also constitutes a generalization of the Fisher index. In De Boer (2007) I argued that the geometric average of all elementary decompositions is to be preferred to each one of the mirror pairs, because it does not only satisfy time reversal, but also factor reversal, i.e. it is, like the Fisher index in case of two factors, *ideal* .
The results for the six elementary decompositions, the average of the mirror pairs and the geometric average of the six elementary decompositions are given in Table 3.

Table 3. Numerical results of the decompositions of the Chung-Rhee example (the change in the size of the economy is equal to 2.107494 in all decompositions)

Decomposition	CO_2 per unit (DF_e)	Leontief inverse (DD_e)	Industry share (DU_e)
Elementary: e_1	0.751053	0.966940	1.027354
Elementary: e_2	0.751053	0.964832	1.029580
Elementary: e_3	0.757830	0.958292	1.027353
Elementary: e_4	0.754164	0.958292	1.032348
Elementary: e_5	0.747087	0.964832	1.035064
Elementary: e_6	0.754164	0.955778	1.035064
Polar: (e_1 and e_6)	0.752607	0.961343	1.031202
Mirror pair 1: (e_2 and e_4)	0.752607	0.961556	1.030972
Mirror pair 2: (e_3 and e_5)	0.752439	0.961556	1.031202
Generalized Fisher	0.752551	0.961485	1.031125

From an empirical point of view the results of the geometric average of the polar decompositions, the geometric average of mirror pairs 1 and 2, and the geometric average of all elementary decompositions (named "generalized Fisher", see below) are extremely close to each other.

3.2. Index number theory

Consider the second column of Table 2 in which I give the elementary decompositions of the factor f (amount of CO_2 emitted per unit of production in the industries). We remark that the terms 1 and 2 are equal to each other, as well as the terms 5 and 6. Collecting the duplicates we can rewrite the geometric average of the six elementary decompositions, D_{x_1}, alternatively as:

$$\left[\frac{\sum x_1(1)x_2(1)x_3(1)}{\sum x_1(0)x_2(1)x_3(1)}\left(\frac{\sum x_1(1)x_2(1)x_3(0)}{\sum x_1(0)x_2(1)x_3(0)}\frac{\sum x_1(1)x_2(0)x_3(1)}{\sum x_1(0)x_2(0)x_3(1)}\right)^{\frac{1}{2}}\frac{\sum x_1(1)x_2(0)x_3(0)}{\sum x_1(0)x_2(0)x_3(0)}\right]^{\frac{1}{3}} \quad (9)$$

where we replaced f_i by x_1, d_{ij} by x_2, and u_j by x_3.

Expression (9) is the generalization of Gini (1937) of the Fisher index to three factors. Siegel (1945) has generalized the Fisher index to an arbitrary number of factors r. His formula, however, is hardly readable and is given without proof. He supplies the results for the special cases of $r=2$ (Fisher), $r=3$ (Gini) and $r=4$. The latter will be presented in the next section. Expression (9) is also equivalent to the formula (7) given in the article by Ang, Liu and Chung (2004) who make use of the very complicated formula of the n-factor Shapley value (Shapley,1953). Ang c.s give the name of “generalized Fisher” to this decomposition.

3.3. Summary

In (9) we first have a term, where the weight is given by the combination {1,1}, which occurs two times in Table 3, and where the exponent is equal to 1/3; in the middle we have two terms with the combinations {1,0} and {0,1} as weights, and exponent 1/6, whereas the final term has weight {0,0}, occurs two times in Table 3, whereas the exponent is 1/3. This is summarized in the following table.

Table 4. Summary for the case of three factors

Number of ones	Combinations		Number of duplicates	Exponent
2	{1,1}		2	1/3
1	{1,0}	{0,1}	1	1/6
0	{0,0}		2	1/3

As before this table is valid for each of the three factors.

4. The case of four factors (Ang, Liu and Chung)

Ang, Liu and Chung (2004) have used the same example as the one I used in section 3. The difference between their approach and mine is that they did not realize that when taking the ratio of D_{C_p} in equation (8) the scalar $y(1)/y(0)$ was independent from the indices and, consequently, could be factorized out, reducing the decomposition reading in four factors to a decomposition reading in three factors. If one uses $r=4$ factors, then we have $4!=24$ permutations (elementary decompositions) that are given in Table 5. For ease of exposition, we have replaced the name of the factors by x_1 through x_4.

Consider the first term of decomposition 1. Mathematically, it reads: $\frac{\sum x_1(1)x_2(1)x_3(1)x_4(1)}{\sum x_1(0)x_2(1)x_3(1)x_4(1)}$. But the first term of the decompositions 2,3,4,5 and 6 is exactly the same, so that this expression occurs six times.

Consider the first term of decomposition 10. Like the first term of the decompositions 12,16,18, 22 and 24 it reads: $\frac{\sum x_1(1)x_2(0)x_3(0)x_4(0)}{\sum x_1(0)x_2(0)x_3(0)x_4(0)}$. This expression occurs six times, as well.

Next, consider the first term of decomposition 7: $\frac{\sum x_1(1)x_2(0)x_3(1)x_4(1)}{\sum x_1(0)x_2(0)x_3(1)x_4(1)}$, which equals the first term of the decomposition 8; this expression occurs twice.

The first terms of the decompositions 17 & 23 are equal to each; the same applies to the first terms of decompositions 11 & 21; 9 &15; 19 & 20; 13 & 14.

Table 5. Elementary decompositions* in case n=4

#	$d(x_1)$	x_2	x_3	x_4	$\times$	x_1	$d(x_2)$	x_3	x_4	$\times$	x_1	x_2	$d(x_3)$	x_4	$\times$	x_1	x_2	x_3	$d(x_4)$
1		1	1	1		0		1	1		0	0		1		0	0	0	
2		1	1	1		0		1	1		0	0		0		0	0	1	
3		1	1	1		0		0	1		0	1		1		0	0	0	
4		1	1	1		0		0	0		0	1		1		0	1	0	
5		1	1	1		0		1	0		0	0		0		0	1	1	
6		1	1	1		0		0	0		0	1		0		0	1	1	
7		0	1	1		1		1	1		0	0		1		0	0	0	
8		0	1	1		1		1	1		0	0		0		0	0	1	
9		0	0	1		1		1	1		1	0		1		0	0	0	
10		0	0	0		1		1	1		1	0		1		1	0	0	
11		0	1	0		1		1	1		0	0		0		1	0	1	
12		0	0	0		1		1	1		1	0		0		1	0	1	
13		1	0	1		0		0	1		1	1		1		0	0	0	
14		1	0	1		0		0	0		1	1		1		0	1	0	
15		0	0	1		1		0	1		1	1		1		0	0	0	
16		0	0	0		1		0	1		1	1		1		1	0	0	
17		1	0	0		0		0	0		1	1		1		1	1	0	
18		0	0	0		1		0	0		1	1		1		1	1	0	
19		1	1	0		0		1	0		0	0		0		1	1	1	
20		1	1	0		0		0	0		0	1		0		1	1	1	
21		0	1	0		1		1	0		0	0		0		1	1	1	
22		0	0	0		1		1	0		1	0		0		1	1	1	
23		1	0	0		0		0	0		1	1		0		1	1	1	
24		0	0	0		1		0	0		1	1		0		1	1	1	

* $d(x_i) = x_i(1)/x_i(0)$; defining $d(x_i) = x_i(1) - x_i(0)$ and replacing "$\times$" by "$+$" we have the 24 elementary decompositions for the additive case.

If we take the geometric mean of all 24 elementary decompositions and we collect the duplicates, we find:

$$D_{x_1} = \left[\frac{\sum x_1(1)x_2(1)x_3(1)x_4(1)}{\sum x_1(0)x_2(1)x_3(1)x_4(1)} \right]^{\frac{1}{4}} \left[\frac{\sum x_1(1)x_2(0)x_3(1)x_4(1)}{\sum x_1(0)x_2(0)x_3(1)x_4(1)} \right]^{\frac{1}{12}} \left[\frac{\sum x_1(1)x_2(1)x_3(0)x_4(1)}{\sum x_1(0)x_2(1)x_3(0)x_4(1)} \right]^{\frac{1}{12}}$$

$$\left[\frac{\sum x_1(1)x_2(1)x_3(1)x_4(0)}{\sum x_1(0)x_2(1)x_3(1)x_4(0)} \right]^{\frac{1}{12}} \left[\frac{\sum x_1(1)x_2(0)x_3(0)x_4(1)}{\sum x_1(0)x_2(0)x_3(0)x_4(1)} \right]^{\frac{1}{12}} \left[\frac{\sum x_1(1)x_2(0)x_3(1)x_4(0)}{\sum x_1(0)x_2(0)x_3(1)x_4(0)} \right]^{\frac{1}{12}}$$

$$\left[\frac{\sum x_1(1)x_2(1)x_3(0)x_4(0)}{\sum x_1(0)x_2(1)x_3(0)x_4(0)} \right]^{\frac{1}{12}} \left[\frac{\sum x_1(1)x_2(0)x_3(0)x_4(0)}{\sum x_1(0)x_2(0)x_3(0)x_4(0)} \right]^{\frac{1}{4}}$$

This expression is given in Siegel (1945) for the case $r = 4$.

It is equivalent to the formula that Ang, Liu and Chung (2004, p. 763) derived from the very complicated formula for the Shapley value in case $r = 4$. This formula can be summarized as follows:

Table 6. Summary for the case of four factors[2]

Number of ones	Combinations			Number of duplicates	Exponent
3	{1,1,1}			6	¼
2	{1,1,0}	{1,0,1}	{0,1,1}	2	1/12
1	{0,0,1}	{0,1,0}	{1,0,0}	2	1/12
0	{0,0,0}			6	¼

Again, this table applies to all four factors.

5. Results of Siegel for the case of five and six factors

As said before, Siegel (1945) gave, without proof, a complicated and rather inaccessible formula for generating the combinations and their exponents. He only supplied the results for the cases $r = 2,3$ and 4. In the previous sections I gave these results, as well as a summary in the form of a table. It can be shown that it follows from Siegel's formula that for the cases $r = 5$ and $r = 6$ these summary tables read:

Table 7. Summary for the case of five factors

Number of ones	Combinations				Number of duplicates	Exponent
4	{1,1,1,1}				24	1/5
3	{1,1,1,0}	{1,1,0,1}	{1,0,1,1}	{0,1,1,1}	6	1/20
2	{1,1,0,0}	{1,0,1,0}	{0,1,1,0}		4	1/30
	{0,0,1,1}	{0,1,0,1}	{1,0,0,1}		4	1/30
1	{0,0,0,1}	{0,0,1,0}	{0,1,0,0}	{1,0,0,0}	6	1/20
0	{0,0,0,0}				24	1/5

and:

Table 8. Summary for the case of six factors

Number of ones	Combinations					Number of duplicates	Exponent
5	{1,1,1,1,1}					120	1/6
4	{1,1,1,1,0}	{1,1,1,0,1}	{1,1,0,1,1}	{1,0,1,1,1}	{0,1,1,1,1}	24	1/30
3	{1,1,1,0,0}	{1,1,0,1,0}	{1,0,1,1,0}	{0,1,1,1,0}		12	1/60
	{1,1,0,0,1}	{1,0,1,0,1}	{0,1,1,0,1}			12	1/60
	{1,0,0,1,1}	{0,1,0,1,1}				12	1/60
	{0,0,1,1,1}					12	1/60
2	{0,0,0,1,1}	{0,0,1,0,1}	{0,1,0,0,1}	**{1,0,0,0,1}**		12	1/60
	{0,0,1,1,0}	{0,1,0,1,0}	{1,0,0,1,0}			12	1/60
	{0,1,1,0,0}	{1,0,1,0,0}				12	1/60
	{1,1,0,0,0}					12	1/60
1	{0,0,0,0,1}	{0,0,0,1,0}	{0,0,1,0,0}	{0,1,0,0,0}	{1,0,0,0,0}	24	1/30
0	{0,0,0,0,0}					120	1/6

It can be verified from the tables 1, 4, 6, 7 and 8 that Siegel has reduced the computational burden of calculating $r!$ permutations (and taking their *unweighted* geometric average) to calculating 2^{r-1} combinations (and taking their *weighted* geometric average). The weights (exponents) are given in my tables 1, 4, 6, 7, and 8. In case $r = 6$ (Table 8), for example, the number of decompositions to be calculated is reduced from 720 to 32. How to use these tables? As example I take the combination **{1,0,0,0,1}** ,boldfaced in Table 8, and use it for the contribution of factor x_3; i.e. $x_3(\mathbf{1})/x_3(\mathbf{0})$. In the geometric average it reads:

$$\left[\frac{\sum x_1(\mathbf{1})x_2(\mathbf{0})x_3(\mathbf{1})x_4(\mathbf{0})x_5(\mathbf{0})x_6(\mathbf{1})}{\sum x_1(\mathbf{1})x_2(\mathbf{0})x_3(\mathbf{0})x_4(\mathbf{0})x_5(\mathbf{0})x_6(\mathbf{1})}\right]^{1/60}$$

With the aid of these tables the computer program to calculate the generalized Fisher index, although tedious, is easily implemented.

6. Concluding remarks

For the case of a multiplicative decomposition Siegel (1945)[3] reduced, by collecting duplicates, the calculation of $r!$ permutations to the calculation of 2^{r-1} combinations. Then, he proposed to calculate the weighted geometric average of the combinations, the number of duplicates being the exponent, which is equivalent to the calculation of the (unweighted) geometric average of all permutations, of course. Independently from Siegel, Shapley (1953) followed the same route for the additive decomposition: he reduced permutations to combinations and proposed to take the weighted arithmetic average, the number of duplicates being the divisor of each combination in the arithmetic average, see Albrecht et al. (2003)[4].

Last, but not least, in order to give credit to both Siegel's and Shapley's contributions I propose to use "Siegel-Shapley decomposition" rather than "generalized Fisher index" or "Structural Decomposition Analysis".

References

Albrecht, J., François, D., Schoors, K., 2002. A Shapley decomposition of carbon emissions without residuals. Energy Policy 30, 727-736.

Ang, B.W., Liu, F.L., Chew, E.P., 2003. Perfect decomposition techniques in energy and environmental analysis. Energy Policy 31, 1561-1566.

Ang, B.W., Liu, F.L., Chung, H.-S., 2004. A generalized Fisher index approach to energy decomposition analysis. Energy Economics 26, 757-763.

Chung, H.S., Rhee, H.C., 2001. A residual-free decomposition of the sources of carbon dioxide emissions: a case for the Korean industries. Energy 26 (1), 15-30.

De Boer, P.M.C., 2008. Additive structural decomposition analysis and index number theory: an empirical application of the Montgomery decomposition. Economic Systems Research 20 (1), 97-109.

De Boer, P.M.C., 2007. Multiplicative structural decomposition analysis and index number theory: an empirical application of the Sato-Vartia decomposition. Report EI

2007-16, Econometric Institute, Erasmus University Rotterdam, submitted to Economic Systems Research.

De Haan, M., 2001. A structural decomposition analysis of pollution in the Netherlands. Economic Systems Research 13 (2), 181- 196.

Dietzenbacher, E. ,Los, B. ,1998. Structural decomposition techniques: sense and sensitivity. Economic Systems Research 10, 307-323.

Fisher,I.,1922. The Making of Index Numbers. Houghton Mifflin, Boston.

Gini, C., 1937. Methods of eliminating the influence of several groups of factors. Econometrica 5 (1), 56-73.

Shapley, L.,1953. A value for *n*-person games. In: Kuhn, H.W., Tucker, A.W. (Eds.), Contributions to the Theory of Games vol.2. Princeton University, Princeton, pp. 307-317.

Siegel, I.H., 1945. The generalized "ideal" index-number formula. Journal of the American Statistical Association 40, 520-523.

Sun, J.W., 1998. Changes in energy consumption and energy intensity: a complete decomposition method. Energy Economics 20 (1), 85-100.

[1] For the additive decomposition, the use of the formula of Shapley has been proposed by Albrecht et al. (2002). Ang et al. (2003) have shown that it is equivalent to the method proposed by Sun (1998).

[2] De Haan (2001) has presented a similar table in case of an additive decomposition with four factors.

[3] I quote Siegel : " … the problem considered here is the development of a general formula … satisfying the relationship $A_n.B_n.C_n... = V_n$ where the n factors on the left are the appropriate indexes of the $a_i, b_i, c_i, ...(i = 1,..., n)$,respectively, for the time period t_1 with respect to the base period t_0 , and $V_n = \sum a_1 b_1 c_1 ... / \sum a_0 b_0 c_0 ...$ is the unique index of the $v_i = a_i b_i c_i ...$ " (page 520) and "The principle underlying the construction of our general formula is essentially simple. …. In fact, there are n! possible sets of aggregative indexes (including duplicates of individual measures) satisfying the relationship $A_n^{'}.B_n^{'}.C_n^{'}... = V_n^{'}$. Now, these raw aggregative indexes do not meet the time-reversal and factor-reversal tests …. These two defects are easily overcome, however, if we take the geometric mean of the n! possible equations of the type $A_n^{'}.B_n^{'}.C_n^{'}... = V_n^{'}$ and define A_n as the geometric mean over all the $A_n^{'}$, including duplications, B_n as the geometric mean over all the $B_n^{'}$, including duplications, etc; …. Each has 2^{n-1} distinct aggregative components…" (page 521).

[4] I quote from Albrecht et al. (2003), page 731: " Indeed, the decomposition problem has formal similarities with a classical problem in cooperative game theory. Shapley (1953) was the first to give a formula for the real power of any given voter in a coalition voting game with transferable utility. This is commonly referred to as the Shapley value" …. "The Shapley decomposition iterates the cumulative approach for every possible order (permutation) of variables. With n variables, we need to make n! calculations, with each calculation based on another order for including new variables. The Shapley value implies that taking the average of the n! estimated contributions of each factor, yields the true contribution for each variable."