

Muradov, Kirill

**Article**

## Structural decomposition analysis with disaggregate factors within the Leontief inverse

Journal of Economic Structures

**Provided in Cooperation with:**

Pan-Pacific Association of Input-Output Studies (PAPAIOS)

*Suggested Citation:* Muradov, Kirill (2021) : Structural decomposition analysis with disaggregate factors within the Leontief inverse, Journal of Economic Structures, ISSN 2193-2409, Springer, Heidelberg, Vol. 10, pp. 1-17,  
<https://doi.org/10.1186/s40008-021-00245-5>

This Version is available at:

<https://hdl.handle.net/10419/261617>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by/4.0/>

RESEARCH

Open Access



# Structural decomposition analysis with disaggregate factors within the Leontief inverse

Kirill Muradov\* 

\*Correspondence:  
kmuradov@mail.ru  
Centre for Information  
and Analysis of Foreign Trade,  
No.9, Leninsky Prospekt,  
119049 Moscow, Russia

## Abstract

A trivial case in input–output structural decomposition analysis is a decomposition of a product of variables, or factors, where one factor is an inverse—typically, Leontief inverse—of a sum of other factors. There may be dozens and hundreds of such factors that describe the changes in subsets of technical coefficients. The existing literature offers ambiguous guidance in this case. The solution that is consistent with the index number theory may be virtually infeasible. The simplified ad hoc solutions require the researcher to make arbitrary choices, lead to biased estimates and do not ensure the consistency-in-aggregation of factors. This paper reviews the ad hoc solutions to the said problem and describes a numerical test to identify the best-performing solution. It is found that calculating the average of the two polar decomposition forms for each factor is superior to other approximations in terms of minimising the errors.

**Keywords:** Structural decomposition analysis, Input–output tables, Leontief inverse

## 1 Introduction

Structural decomposition analysis (SDA) of input–output data explains the change in a product of matrices and vectors—referred to as determinants or factors—in terms of the changes in each of those factors. In a single country input–output framework with  $N$  industries, the result may record the decomposed changes for each industry  $i \in \{1, \dots, N\}$ , as in the classic contribution of Dietzenbacher and Los (1998). The underlying changes in factors usually describe economy-wide changes, covering all  $N$  industries, and the result implicitly aggregates the effects of each industry  $j \in \{1, \dots, N\}$  or groups thereof.

Disaggregation of factors into components that are industry-specific significantly increases the number of factors in the decomposition and, therefore, the complexity of computations. This also means that one of the factors is an inverse—typically, Leontief inverse—of a sum of other, industry-specific factors.

The complexity of structural decomposition where the target variable is a product of a large number of factors was addressed by Dietzenbacher and Los (1998) who reviewed various simplified ad hoc solutions. Since then, many authors have adopted the average

of the two ‘polar decompositions’ as their default SDA strategy, though de Haan (2001) indicated that these ‘polar decompositions’ were not unique, and there was no reason why a particular pair of ‘polar decompositions’ should be preferred. This is even more obvious if an inverse matrix is a sum of factors where their order can be freely altered.

The problem of additive factors nested within the Leontief inverse is the biggest complication in SDA. Because of matrix inversion this problem is unknown in index number analysis, and decompositions that are consistent with the index number theory (surveyed by Wang et al 2017; de Boer and Rodrigues 2020) may be infeasible. A ‘textbook’ solution, discussed by Rose and Casler (1996), uses a factorisation of the change in Leontief inverse. Whether this solution biases the result and whether there is another ad hoc solution that minimises the bias—has not been systematically explored.

Bridging this gap in the literature, this paper reviews the ad hoc solutions that simplify the SDA calculations and approximate the results of complete decompositions. There are four solutions, called here ‘shortcuts’, for the additive SDA and two for the multiplicative SDA. The author runs numerical tests of all ‘shortcuts’ using the time series of input–output data from Statistics Denmark and Statistics Netherlands. The number of factors within the Leontief inverse varies from 5 to 8. The tests reveal that one of the ‘shortcuts’ performs consistently better than others. This ‘shortcut’—the average of the two ‘polar decomposition forms’—is not new, but remained largely unnoticed, though reported to be efficient by Dietzenbacher and Los (1998). Indeed, it makes the SDA ‘short’: instead of computing  $2^{N-1}$  decomposition forms and taking their weighted average for each of  $N$  factors to obtain the true result, it only requires computing 2 decomposition forms and taking their simple average to obtain an approximate result. Furthermore, a pair of ‘polar decomposition forms’ is unique to each factor and does not require the researcher to guess which of the forms should be preferred.

The contribution of this paper is the evidence of the best-performing approximate solution in a special case of SDA where the exact and true solution becomes prohibitively complex. The design of the numerical experiment that yields that evidence stands out for a larger data input and number of factors than in many previous surveys of SDA techniques.

The special case addressed here may become a typical case in SDA of inter-country input–output tables where a factor may describe, e.g., a change in trade relationship between two countries, rather than the change in global trade pattern.

In the language of econometric modelling, SDA is designed to construct a ‘counterfactual’ to answer the question “how the target variable would have changed if only one factor changed and other remained the same?” With country- and industry-specific factors, this ‘counterfactual’ can be constructed at a very disaggregate level. Then the direct calculation of a disaggregate effect by SDA may be an appealing alternative to the indirect estimation of an aggregate effect by means of statistical inference followed by a possible disaggregation. This is true in cases where both SDA and econometric solutions apply, with due account of their conceptual differences, for example, in cases where the target and explanatory variables are connected via global value chains.

The paper is laid out as follows. Section 2 reminds the reader of the classic solutions to the additive and multiplicative SDA that are consistent with the index number theory, briefly reviews generic approaches to reduce the intensity of SDA calculations and to

obtain reasonable approximations. Following is an exposition of a particular case where one of the factors in the decomposition is an inverse of a sum of other factors, or, more precisely, the Leontief inverse. Then the ad hoc solutions, or ‘shortcuts’ to the complete decomposition, are formulated. Section 3 explains how these ‘shortcuts’ are tested and evaluated and how the necessary data are selected and prepared. This section concludes by reporting and discussing the results of the numerical experiment. Final remarks are set out in Sect. 4.

## 2 Methods

### 2.1 The general case of structural decomposition

For the general case of SDA, we will consider a variable  $Y$  that is a product of  $N$  variables:

$$Y = Z_1 Z_2 \dots Z_N = \prod_{n=1}^N Z_n. \tag{1}$$

Each variable  $Z_n$  is called *factor* where the subscript  $n$  identifies the  $n$ th factor in the decomposition. In the general case,  $Y$  and  $Z_n$  are matrices, and in special cases,  $Z_1$  and/or  $Z_N$  may be vectors and  $Y$  may be a vector or a scalar.

Using superscripts (0) and (1) for the initial period 0 and the terminal period 1, define the change in  $Y$ . This can be done in two ways. First, define the absolute change in  $Y$ :

$$Y^{(1)} - Y^{(0)} = Z_1^{(1)} Z_2^{(1)} \dots Z_N^{(1)} - Z_1^{(0)} Z_2^{(0)} \dots Z_N^{(0)} = \prod_{n=1}^N Z_n^{(1)} - \prod_{n=1}^N Z_n^{(0)} = \Delta Y, \tag{2}$$

where  $\Delta$  is the difference operator.  $\Delta Y$  may be understood as an increment of  $Y$  and has the same units of measurement as  $Y$  (usually, monetary units).

Second, define the relative change in  $Y$ :

$$Y^{(1)} \oslash Y^{(0)} = \left( Z_1^{(1)} Z_2^{(1)} \dots Z_N^{(1)} \right) \oslash \left( Z_1^{(0)} Z_2^{(0)} \dots Z_N^{(0)} \right) = \left( \prod_{n=1}^N Z_n^{(1)} \right) \oslash \left( \prod_{n=1}^N Z_n^{(0)} \right) = PY, \tag{3}$$

where  $P$  is the ratio operator and  $\oslash$  signifies the element-by-element division.  $PY$  measures the growth of each element in  $Y$ , and the units of measurement are dimensionless.

The problem of structural decomposition is to decompose  $\Delta Y$  and  $PY$  into  $N$  terms that would attribute the change in  $Y$  to the changes in each  $n$ th factor. For example, the absolute change in  $Y$  that is induced by a change in  $Z_1$  may be written as:

$$\Delta Y(Z_1, t_2, \dots, t_N) = Z_1^{(1)} Z_2^{(t_2)} \dots Z_N^{(t_N)} - Z_1^{(0)} Z_2^{(t_2)} \dots Z_N^{(t_N)}, \tag{4}$$

where the superscripts  $(t_2) \dots (t_N)$  can take values 1 or 0, and factors other than  $Z_1$  can therefore be defined at either initial or terminal periods. For brevity, we will denote  $\Delta Y(Z_1, t_2, \dots, t_N)$  by  $\Delta Y(Z_1)$ . One combination of  $\{t_2, \dots, t_N\}$  corresponds to one *decomposition form* of  $\Delta Y(Z_1)$ . From the literature on the index number theory (Siegel 1945) and structural decomposition analysis (Seibel 2003), it is known that the number of possible unique decomposition forms of  $\Delta Y(Z_1)$  is  $2^{N-1}$ .

It is customary to refer to the two forms of  $\Delta Y(\mathbf{Z}_n)$  where all factors other than  $\mathbf{Z}_n$  are defined at period 1 or period 0 as *polar decomposition forms* (Dietzenbacher and Los 1998; de Haan 2001). For example, the polar forms of  $\Delta Y(\mathbf{Z}_1)$  are:

$$\Delta Y(\mathbf{Z}_1 | t_2 = 1, \dots, t_N = 1) = \mathbf{Z}_1^{(1)} \mathbf{Z}_2^{(1)} \dots \mathbf{Z}_N^{(1)} - \mathbf{Z}_1^{(0)} \mathbf{Z}_2^{(1)} \dots \mathbf{Z}_N^{(1)}, \tag{5a}$$

$$\Delta Y(\mathbf{Z}_1 | t_2 = 0, \dots, t_N = 0) = \mathbf{Z}_1^{(1)} \mathbf{Z}_2^{(0)} \dots \mathbf{Z}_N^{(0)} - \mathbf{Z}_1^{(0)} \mathbf{Z}_2^{(0)} \dots \mathbf{Z}_N^{(0)}. \tag{5b}$$

All  $2^{N-1}$  forms of  $\Delta Y(\mathbf{Z}_n)$  can be classified according to the *distance from the polar form*, denoted by  $k \in \{0, \dots, N - 1\}$ . One of the polar forms needs to be chosen as the starting point in the decomposition: for convenience, let it be (5a), and the corresponding value of  $k$  is 0. This may be understood as follows: none of the remaining factors is defined at period 0, and all of those are at period 1. At  $k = 1$ , one of the remaining factors is now defined at period 0, while the rest are still at period 1. Obviously, there are  $N - 1$  such forms of  $\Delta Y(\mathbf{Z}_n)$ . At  $k = 2$ , two of the remaining factors are defined at period 0, while the rest are at period 1, which continues until the other polar form is reached at  $k = N - 1$ . The number of unique forms of  $\Delta Y(\mathbf{Z}_n)$  at each  $k$  is equal to the number of  $k$  combinations of  $N - 1$ :  $\frac{(N-1)!}{(N-1-k)!k!} = \binom{N-1}{k}$ . We will denote each  $m$ th unique form of  $\Delta Y(\mathbf{Z}_n)$  at  $k$  steps from the polar form by  $\Delta Y(\mathbf{Z}_n)_{k,m}$  with the subscript  $m$  running from 1 to  $\binom{N-1}{k}$ . Calculating the weighted average of all unique forms at each  $k$  with the respective weights  $c_k$  and across all  $k$  yields the aggregate form of  $\Delta Y(\mathbf{Z}_n)$ :

$$\overline{\Delta Y(\mathbf{Z}_n)} = \sum_{k=0}^{N-1} c_k \sum_{m=1}^{\binom{N-1}{k}} \Delta Y(\mathbf{Z}_n)_{k,m} \quad \text{where} \quad c_k = \frac{(N-1-k)!k!}{N!}. \tag{6}$$

Expression (6) may be recognised as the Bennet indicator for the  $n$ th factor (de Boer and Rodrigues 2020) that is the additive counterpart to the well-known Fisher index. Finally, collecting the changes induced by all  $N$  factors produces a full additive decomposition of  $\Delta Y$ :

$$\Delta Y = \sum_{n=1}^N \overline{\Delta Y(\mathbf{Z}_n)} = \sum_{n=1}^N \sum_{k=0}^{N-1} c_k \sum_{m=1}^{\binom{N-1}{k}} \Delta Y(\mathbf{Z}_n)_{k,m}. \tag{7}$$

In case of multiplicative decomposition, the relative change in  $Y$  that is induced by a change in  $\mathbf{Z}_1$  should be written as:

$$PY(\mathbf{Z}_1, t_2, \dots, t_N) = \left( \mathbf{Z}_1^{(1)} \mathbf{Z}_2^{(t_2)} \dots \mathbf{Z}_N^{(t_N)} \right) \oslash \left( \mathbf{Z}_1^{(0)} \mathbf{Z}_2^{(t_2)} \dots \mathbf{Z}_N^{(t_N)} \right). \tag{8}$$

The relative change  $PY(\mathbf{Z}_n)$  aggregated within and between all  $k$  is as follows:

$$\overline{PY(\mathbf{Z}_n)} = \bigodot_{k=0}^{N-1} \left( \binom{N-1}{k} \bigodot_{m=1}^{N-1} PY(\mathbf{Z}_n)_{k,m} \right)^{\circ c_k}, \tag{9}$$

where  $\bigodot$  denotes the Hadamard product of a sequence, the power  $\circ c_k$  applies to the elements of the respective vectors (Hadamard power), and the weights  $c_k$  are defined as in Eq. (6). Expression (9) is nothing but the Fisher index for the  $n$ th factor. And the full multiplicative decomposition of  $PY$  is given by:

$$PY = \bigodot_{n=1}^N \overline{PY(\mathbf{Z}_n)} = \bigodot_{n=1}^N \bigodot_{k=0}^{N-1} \left( \binom{N-1}{k} \bigodot_{m=1}^{N-1} PY(\mathbf{Z}_n)_{k,m} \right)^{\circ c_k}. \tag{10}$$

Decompositions (7) and (10) are *exact* in the sense that they do not have any residual terms. These two formulae may be derived by computing the simple average (respectively, arithmetic or geometric) of all *elementary decompositions* of the change in  $\mathbf{Y}$ . An elementary decomposition is made of a unique sequence of  $N$  decomposition forms denoting the change in  $\mathbf{Y}$  because of the changes in each  $n$ th factor, that is  $\Delta\mathbf{Y}(\mathbf{Z}_n)$  or  $PY(\mathbf{Z}_n)$ , where one  $n$  is consecutively chosen at each  $k$ . The total number of sequences is equal to the number of permutations of  $N$  factors, or  $N!$  (Dietzenbacher and Los 1998).

Computing the simple average from the elementary decompositions requires  $N!$  forms of  $\Delta\mathbf{Y}(\mathbf{Z}_n)$  or  $PY(\mathbf{Z}_n)$  some of which are duplicates, while computing the weighted average only involves  $2^{N-1}$  unique forms thereof. The sum of all weights, with due account of the number of times they apply to all  $m$ th forms at step  $k$ , is unity:

$$\sum_{k=0}^{N-1} c_k \frac{(N-1)!}{(N-1-k)!k!} = \sum_{k=0}^{N-1} \frac{(N-1-k)!k!}{N!} \frac{(N-1)!}{(N-1-k)!k!} = N \frac{(N-1)!}{N!} = 1.$$

Tables 1 and 2 provide an exemplary calculation of the coefficients (weights) and the number of unique decomposition forms for each factor required to implement decompositions (7) and (10) where the number of factors is up to 10. Thanks to the underlying formula, Table 2 contains the Pascal's triangle less the first row. The number of unique decomposition forms for each factor ( $2^{N-1}$ ) and their sum for all factors ( $N2^{N-1}$ ) are a good indication of the complexity of the decomposition exercise: in case of 5 factors, 16 decomposition forms need to be computed for each factor and 80 for all factors, and in case of 10 factors, the respective numbers are 512 and 5120.

Hence, the well-known problem of SDA is an exponential growth of the array of terms to be computed for an exact and full decomposition of  $\mathbf{Y}$  as the number of factors increases. Several approaches have been proposed in the literature on SDA to address this problem.

**Table 1** Denominator of the coefficients (weights) in the decomposition

N	k									
	0	1	2	3	4	5	6	7	8	9
2	2	2								
3	3	6	3							
4	4	12	12	4						
5	5	20	30	20	5					
6	6	30	60	30	20	6				
7	7	42	105	140	105	42	7			
8	8	56	168	280	280	168	56	8		
9	9	72	252	504	630	504	252	72	9	
10	10	90	360	840	1260	1260	840	360	90	10

The numerator in all coefficients is 1. N is the number of factors in the decomposition and k is the distance from one of the polar decomposition forms

**Table 2** Number of unique decomposition forms that correspond to one factor at each step from the polar form

N	k										Sum
	0	1	2	3	4	5	6	7	8	9	
2	1	1									2
3	1	2	1								4
4	1	3	3	1							8
5	1	4	6	4	1						16
6	1	5	10	10	5	1					32
7	1	6	15	20	15	6	1				64
8	1	7	21	35	35	21	7	1			128
9	1	8	28	56	70	56	28	8	1		256
10	1	9	36	84	126	126	84	36	9	1	512

N is the number of factors in the decomposition and k is the distance from one of the polar decomposition forms

It may be reasonable to handle factors in groups. For example, the factors that affect final demand may be delimited from those that affect intermediate demand. This enables *hierarchical* or *nested* SDA (see e.g., Koller and Stehrer 2010). In the most typical case of two groups, the array of N factors is divided into two subsets with R and S factors, where  $R + S = N$ . The decomposition will involve two aggregate factors at the first tier and R + S factors at the second tier. The total number of decomposition forms for all N factors will now be  $R2^R + S2^S$  which is necessarily less than  $N2^{N-1}$  if at least R or S is more than 1.<sup>1</sup>

Dietzenbacher and Los (1998) and Dietzenbacher et al (2000) test the average of the two polar decompositions against the average of all elementary decompositions. One

<sup>1</sup> Reduction in the total number of decomposition forms may be described by the ratio  $\frac{R+S}{\frac{R}{2^{S-1}} + \frac{S}{2^{R-1}}}$ . Note that, in the case of additive SDA with two subsets of factors, two decomposition forms for each aggregate factor may be conveniently merged into one expression, e.g.,  $\Delta Y(Z_1) = (\Delta Z_1) \frac{1}{2} (Z_1 + Z_2)$ . Then  $R2^R + S2^S$  decomposition forms may be computed as  $R2^{R-1} + S2^{S-1}$  expressions, but these expressions must not be confused with the decomposition forms.

polar decomposition is the elementary decomposition where at each  $k$  starting from 0,  $n = k + 1$ , and another polar decomposition is its ‘mirror image’ with  $k$  running in reverse order, starting from  $N - 1$ . There are two polar decompositions that contain the polar decomposition forms for the first and the last factors. This only requires computing two decomposition forms for each factor and  $2N$  forms for all factors. They show that the result is a good approximation of the full decomposition. However, de Haan (2001) stresses that the selection of the two polar decompositions is arbitrary. There exist  $N!/2$  such decomposition pairs among the elementary decompositions, and any factor can be first or last.

Dietzenbacher and Los (1998) also discuss two ad hoc solutions to simplify the SDA calculations that, however, do not provide exact decompositions. One solution is to take the averages of pairs of the polar decomposition forms for each factor.<sup>2</sup> These averages may be treated as special cases of Eqs. (6) and (9) where the coefficients (weights)  $c_k$  are set to  $1/2$  and  $k$  are set to 0 and  $N - 1$  ( $m$  is irrelevant because there is only one polar decomposition form at  $k = 0$  and  $k = N - 1$ ). Then the full decompositions may be written as:

$$\Delta Y = \sum_{n=1}^N \overline{\Delta Y(\mathbf{Z}_n)_{pol}} = \sum_{n=1}^N \sum_{k=0, N-1} \frac{1}{2} \Delta Y(\mathbf{Z}_n)_k + \mathbf{e}_\Delta, \tag{11}$$

$$PY = \bigodot_{n=1}^N \overline{PY(\mathbf{Z}_n)_{pol}} = \left[ \bigodot_{n=1}^N \bigodot_{k=0, N-1} (PY(\mathbf{Z}_n)_k)^{\circ \frac{1}{2}} \right] \circ \mathbf{e}_p, \tag{12}$$

where  $\mathbf{e}_\Delta$  and  $\mathbf{e}_p$  are residual terms, respectively, in the additive and multiplicative cases, the lower index *pol* denotes a polar decomposition form and  $\circ$  denotes the Hadamard product or Hadamard power in the superscript. This ad hoc solution requires computing  $2N$  decomposition forms for all factors.

Another opportunity is to apply the so-called *mid-point weights* to each factor  $\mathbf{Z}_n$ . In the additive case, this signifies that, for example,  $\Delta Y(\mathbf{Z}_1)$  is defined as  $\mathbf{Z}_1^{(1)} \mathbf{Z}_2^{(0/1)} \dots \mathbf{Z}_N^{(0/1)} - \mathbf{Z}_1^{(0)} \mathbf{Z}_2^{(0/1)} \dots \mathbf{Z}_N^{(0/1)}$  plus an error term, where for  $n \neq 1$ ,  $\mathbf{Z}_n^{(0/1)} = \frac{1}{2} \mathbf{Z}_n^{(0)} + \frac{1}{2} \mathbf{Z}_n^{(1)}$ . This solution, also known as the Marshall–Edgeworth decomposition, therefore involves only one decomposition form for each factor and  $N$  forms for all factors, but also has a residual term. It is less clear how the mid-term weights apply in the multiplicative case.

Dietzenbacher and Los (1998) report that both solutions perform rather well, and their results are very close to those from the full and exact additive decomposition of  $\mathbf{Y}$ .

In addition to Bennet and Fisher decompositions that are combinatorial in nature and build on, respectively, the arithmetic and geometric averages, another family of decompositions employs the logarithmic mean. These include Montgomery (additive) decomposition, Montgomery–Vartia (multiplicative) decomposition and Sato–Vartia (additive and multiplicative) decomposition that are exhaustively reviewed by de Boer (2008, 2009); de Boer and Rodrigues (2020). These decompositions, particularly favoured for

<sup>2</sup> Polar decomposition forms here and elsewhere must not be confused with polar decompositions.



SDA of energy and emissions, are also known as ‘Divisia-based’, ‘Divisia-linked’ decomposition approaches or ‘Logarithmic Mean Divisia Index’ (LMDI) methods (reviewed by Su and Ang (2012), Wang et al (2017)).

Montgomery, Montgomery–Vartia and Sato–Vartia decompositions are exact, and the related indicators and indices for each factor  $Z_n$  are shown to be ideal—as are the Bennet indicator and the Fisher index—because they satisfy a number of tests from the index number theory (e.g., time reversal test, product test, etc.). Only one decomposition needs to be computed to obtain indicators or indices for all factors, so the methods based on the logarithmic mean are believed to be easier to implement than the combinatorial ones. However, these methods are not reviewed here because they do not apply to the special case of structural decomposition that motivated this paper.

**2.2 A special case of structural decomposition: factors nested within an inverse matrix**

Further complication arises if one of the factors  $Z_n$  is an inverse of a sum of other factors. This is a typical case in input–output SDA. To elaborate, we now turn to a simple Leontief model<sup>3</sup> where industry output  $\mathbf{x}$  is a product of the Leontief inverse  $\mathbf{L}$  and final demand  $\mathbf{f}$ :

$$\mathbf{x} = \mathbf{L}\mathbf{f} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{f}.$$

$\mathbf{A}$  is a matrix of technical coefficients with the  $j$ th column describing the expenditures on intermediate inputs per one unit of output of industry  $j \in \{1, \dots, J\}$ . For our illustrative example, we will treat columns of  $\mathbf{A}$  as factors. Each  $j$ th factor is then matrix  $\mathbf{A}_j$  where all columns other than the  $j$ th column from  $\mathbf{A}$  are set to zero. Under constant prices, the changes in  $j$ th factor may be understood as the changes in the ‘production recipe’ or technology. The matrix of technical coefficients is now the sum of all  $J$  factors,  $\mathbf{A} = \mathbf{A}_1 + \mathbf{A}_2 + \dots + \mathbf{A}_J$ , and the change in output in the additive case is:

$$\begin{aligned} \Delta\mathbf{x} &= \mathbf{L}^{(1)}\mathbf{f}^{(1)} - \mathbf{L}^{(0)}\mathbf{f}^{(0)} = \\ &= \left(\mathbf{I} - \mathbf{A}_1^{(1)} - \mathbf{A}_2^{(1)} - \dots - \mathbf{A}_J^{(1)}\right)^{-1}\mathbf{f}^{(1)} - \left(\mathbf{I} - \mathbf{A}_1^{(0)} - \mathbf{A}_2^{(0)} - \dots - \mathbf{A}_J^{(0)}\right)^{-1}\mathbf{f}^{(0)}, \end{aligned} \tag{13}$$

and in the multiplicative case:

$$\begin{aligned} \text{Px} &= \left(\mathbf{L}^{(1)}\mathbf{f}^{(1)}\right) \oslash \left(\mathbf{L}^{(0)}\mathbf{f}^{(0)}\right) = \\ &= \left(\left(\mathbf{I} - \mathbf{A}_1^{(1)} - \mathbf{A}_2^{(1)} - \dots - \mathbf{A}_J^{(1)}\right)^{-1}\mathbf{f}^{(1)}\right) \oslash \left(\left(\mathbf{I} - \mathbf{A}_1^{(0)} - \mathbf{A}_2^{(0)} - \dots - \mathbf{A}_J^{(0)}\right)^{-1}\mathbf{f}^{(0)}\right). \end{aligned} \tag{14}$$

Invoking the hierarchical approach, we first decompose the change in output into the changes induced by the change in the Leontief inverse and the change in final demand:

$$\Delta\mathbf{x} = \frac{1}{2}\left(\mathbf{L}^{(1)} - \mathbf{L}^{(0)}\right)\left(\mathbf{f}^{(1)} + \mathbf{f}^{(0)}\right) + \frac{1}{2}\left(\mathbf{L}^{(1)} + \mathbf{L}^{(0)}\right)\left(\mathbf{f}^{(1)} - \mathbf{f}^{(0)}\right), \tag{15}$$

<sup>3</sup> The following description is also applicable to the Ghosh model provided that the researcher is able to correctly interpret the factors therein.

$$P\mathbf{x} = \left( \frac{\mathbf{L}^{(1)}\mathbf{f}^{(1)}}{\mathbf{L}^{(0)}\mathbf{f}^{(1)}} \circ \frac{\mathbf{L}^{(1)}\mathbf{f}^{(0)}}{\mathbf{L}^{(0)}\mathbf{f}^{(0)}} \right)^{\circ \frac{1}{2}} \circ \left( \frac{\mathbf{L}^{(1)}\mathbf{f}^{(1)}}{\mathbf{L}^{(1)}\mathbf{f}^{(0)}} \circ \frac{\mathbf{L}^{(0)}\mathbf{f}^{(1)}}{\mathbf{L}^{(0)}\mathbf{f}^{(0)}} \right)^{\circ \frac{1}{2}}. \tag{16}$$

For easier exposition, in the multiplicative case (16) the element-by-element division symbol  $\circ$  is replaced by the fraction sign, and the power applies to each element of the vector in brackets. The first term on the right-hand side of Eqs. (15) and (16) describes the changes in output related to the changes in the Leontief inverse. Note that this term is the average of the two underlying decomposition forms of  $\Delta\mathbf{x}(\mathbf{L})$  or  $P\mathbf{x}(\mathbf{L})$ .

There are two basic options (Rose and Casler 1996) to further decompose the changes related to the Leontief inverse. First, similar to Eq. (4), the difference of  $\mathbf{x}$  that is attributed, for example, to the changes in the outlays of the first industry ( $j = 1$ ) may be written as follows:

$$\begin{aligned} \Delta\mathbf{x}(\mathbf{A}_1, t_2, \dots, t_j) &= \frac{1}{2} \left( \mathbf{I} - \mathbf{A}_1^{(1)} - \mathbf{A}_2^{(t_2)} - \dots - \mathbf{A}_j^{(t_j)} \right)^{-1} \left( \mathbf{f}^{(1)} + \mathbf{f}^{(0)} \right) - \\ &\quad - \frac{1}{2} \left( \mathbf{I} - \mathbf{A}_1^{(0)} - \mathbf{A}_2^{(t_2)} - \dots - \mathbf{A}_j^{(t_j)} \right)^{-1} \left( \mathbf{f}^{(1)} + \mathbf{f}^{(0)} \right). \end{aligned} \tag{17}$$

Then the complete decomposition of  $\Delta\mathbf{x}(\mathbf{A})$  will be identical to that of  $\Delta\mathbf{Y}$  in Sect. 2.1 involving a total of  $2^J$  decomposition forms<sup>4</sup> of  $\Delta\mathbf{x}(\mathbf{A}_j)$  for all  $J$  factors and their weighted averages, leading to  $J$  Bennet indicators. Similarly, the decomposition of the ratio  $P\mathbf{x}(\mathbf{A})$  involves  $J$  Fisher indices.

The second option utilises the known property of the Leontief inverse:  $\Delta\mathbf{L} = \mathbf{L}^{(1)}(\Delta\mathbf{A})\mathbf{L}^{(0)} = \mathbf{L}^{(0)}(\Delta\mathbf{A})\mathbf{L}^{(1)}$ . Replacing  $\Delta\mathbf{A}$  with the sum of the changes in  $J$  factors yields:

$$\Delta\mathbf{L} = \mathbf{L}^{(1)}(\Delta\mathbf{A})\mathbf{L}^{(0)} = \mathbf{L}^{(1)}(\Delta\mathbf{A}_1)\mathbf{L}^{(0)} + \mathbf{L}^{(1)}(\Delta\mathbf{A}_2)\mathbf{L}^{(0)} + \dots + \mathbf{L}^{(1)}(\Delta\mathbf{A}_J)\mathbf{L}^{(0)}, \tag{18a}$$

$$\Delta\mathbf{L} = \mathbf{L}^{(0)}(\Delta\mathbf{A})\mathbf{L}^{(1)} = \mathbf{L}^{(0)}(\Delta\mathbf{A}_1)\mathbf{L}^{(1)} + \mathbf{L}^{(0)}(\Delta\mathbf{A}_2)\mathbf{L}^{(1)} + \dots + \mathbf{L}^{(0)}(\Delta\mathbf{A}_J)\mathbf{L}^{(1)}. \tag{18b}$$

For any  $j$ ,  $\mathbf{L}^{(1)}(\Delta\mathbf{A}_j)\mathbf{L}^{(0)}$  is no longer equal to  $\mathbf{L}^{(0)}(\Delta\mathbf{A}_j)\mathbf{L}^{(1)}$ , and an average of Eqs. (18a) and (18b) can be taken to avoid an arbitrary choice. Then there will be 2 decomposition forms for each  $j$ th factor and  $2J$  forms for all  $J$  factors.

There are two issues with this second option. Although mathematically correct, the terms related to each  $j$ th factor in Eqs. (18a) and (18b) or their average do not exactly capture changes attributable to that factor:  $\mathbf{L}^{(1)}(\Delta\mathbf{A}_j)\mathbf{L}^{(0)} \neq \mathbf{L}^{(1)} - (\mathbf{I} - \mathbf{A}^{(1)} + \Delta\mathbf{A}_j)^{-1}$  and  $\mathbf{L}^{(0)}(\Delta\mathbf{A}_j)\mathbf{L}^{(1)} \neq \mathbf{L}^{(0)}(\Delta\mathbf{A}_j)(\mathbf{I} - \mathbf{A}^{(0)} - \Delta\mathbf{A}_j)^{-1}$ . The correct term for each  $j$ th factor would be  $\mathbf{L}^{(1)}(\Delta\mathbf{A}_j)(\mathbf{I} - \mathbf{A}^{(1)} + \Delta\mathbf{A}_j)^{-1}$  and  $\mathbf{L}^{(0)}(\Delta\mathbf{A}_j)(\mathbf{I} - \mathbf{A}^{(0)} - \Delta\mathbf{A}_j)^{-1}$ , but these will not add up to  $\Delta\mathbf{L}$ . Furthermore, this simplifying option is not available in the case of multiplicative decomposition.

Note that the decompositions based on the logarithmic mean require that the dependent variable be expressed as a product of factors as in Eq. (1). As an inverse of a sum of

<sup>4</sup> The number of decomposition forms doubles because the decomposition is hierarchical, and, for each  $j$ , final demand can be defined at either period 0 or period 1.

factors cannot be expressed as a product of those factors, the Montgomery, Montgomery–Vartia and Sato–Vartia decompositions are irrelevant in this special case.

The problem is now to find a reasonable ‘shortcut’ to the decomposition of  $\Delta\mathbf{x}(\mathbf{A})$  and  $\mathbf{Px}(\mathbf{A})$  and to avoid computing all  $2^J$  decomposition forms where each form includes an inverse of a unique combination of factors therein.

### 2.3 ‘Shortcuts’ to the complete decomposition with factors nested in the Leontief inverse

There is no specific order of factors within the Leontief inverse: any factor can appear first or last in the sum without affecting the result. It is even more obvious in this case that the choice of two polar decompositions is arbitrary.<sup>5</sup> Therefore, we will not consider the average of two polar decompositions as a viable shortcut.

The shortcuts to be considered are as follows:

*Shortcut 1* The averages of pairs of the polar decomposition forms for each factor

This shortcut applies Eqs. (11) and (12) to the factors within the Leontief inverse. For each  $j$ , compute two polar decomposition forms of  $\Delta\mathbf{x}(\mathbf{A}_j)$  according to Eq. (17) where all time periods  $t$  other than  $t_j$  are set to 1 and 0 (in other words, two decomposition forms at distance  $k \in \{0, J - 1\}$  from the polar form), then take the average of the two forms. The change of  $\mathbf{x}$  because of change in the  $j$ th factor then is:

$$\Delta\mathbf{x}(\mathbf{A}_j)_{S1} = \overline{\Delta\mathbf{x}(\mathbf{A}_j)_{pol}} = \sum_{k=0, J-1} \frac{1}{2} \Delta\mathbf{x}(\mathbf{A}_j)_k, \tag{19}$$

$$\mathbf{Px}(\mathbf{A}_j)_{S1} = \overline{\mathbf{Px}(\mathbf{A}_j)_{pol}} = \bigodot_{k=0, N-1} (\mathbf{Px}(\mathbf{A}_j)_k)^{\circ \frac{1}{2}}. \tag{20}$$

The above requires computing only 2 forms for each factor,  $2J$  forms for all factors and does not provide an exact decomposition.

*Shortcut 2* The normalised averages of pairs of the polar decomposition forms for each factor

The information from Eqs. (19) and (20) can be utilised to modify the averaged polar forms  $\overline{\Delta\mathbf{x}(\mathbf{A}_j)_{pol}}$  and  $\overline{\mathbf{Px}(\mathbf{A}_j)_{pol}}$ , so that the decompositions of  $\Delta\mathbf{x}(\mathbf{A})$  and  $\mathbf{Px}(\mathbf{A})$  are exact. Therefore, the residual term is distributed across the indicators or indices for all  $J$  factors. In the additive case, this can be achieved by multiplying each element in the averaged  $j$ th polar form by a respective coefficient:

$$\Delta\mathbf{x}(\mathbf{A}_j)_{S2} = \mathbf{c}_\Delta \circ \overline{\Delta\mathbf{x}(\mathbf{A}_j)_{pol}} \quad \text{where} \quad \mathbf{c}_\Delta = (\Delta\mathbf{x}(\mathbf{A})) \oslash \left( \sum_{j=1}^J \overline{\Delta\mathbf{x}(\mathbf{A}_j)_{pol}} \right). \tag{21}$$

In the multiplicative case, the coefficients need to be defined as powers that apply to the base on an element-by-element basis:

<sup>5</sup> In the generic case of the decomposition of a product of  $N$  variables as in Eq. (1), the choice of the two polar decompositions may be justified by the order of those variables that cannot be arbitrarily altered. This is the standard approach in the related literature.

$$\begin{aligned}
 \mathbf{Px}(\mathbf{A}_j)_{S2} &= \left( \overline{\mathbf{Px}(\mathbf{A}_j)_{pol}} \right)^{\circ \mathbf{c}_p}, \\
 \text{where } (\mathbf{c}_p)_i &= \log_{(\mathbf{b})_i} (\mathbf{Px}(\mathbf{A}))_i, \quad (\mathbf{b})_i = \prod_{j=1}^J \left( \overline{\mathbf{Px}(\mathbf{A}_j)_{pol}} \right)_i,
 \end{aligned}
 \tag{22}$$

and  $i$  denotes the  $i$ th element in the respective  $J \times 1$  vector.

The above modification of shortcut 1 requires the computation of  $2J$  polar decomposition forms for each factor and for all factors and provides an exact decomposition.

*Shortcut 3* Decomposition with mid-point weights (Marshall–Edgeworth decomposition)—only for additive decomposition

For the calculation of the change attributed to the  $j$ th factor, other factors are defined as the arithmetic mean of their values at period 0 and period 1. We modify Eq. (17) to formally describe shortcut 3:

$$\begin{aligned}
 \Delta \mathbf{x}(\mathbf{A}_j)_{S3} &= \frac{1}{2} \left( \mathbf{I} - \mathbf{A}_j^{(1)} - \sum_{h \neq j} \frac{1}{2} (\mathbf{A}_h^{(1)} + \mathbf{A}_h^{(0)}) \right)^{-1} (\mathbf{f}^{(1)} + \mathbf{f}^{(0)}) - \\
 &\quad - \frac{1}{2} \left( \mathbf{I} - \mathbf{A}_j^{(0)} - \sum_{h \neq j} \frac{1}{2} (\mathbf{A}_h^{(1)} + \mathbf{A}_h^{(0)}) \right)^{-1} (\mathbf{f}^{(1)} + \mathbf{f}^{(0)}).
 \end{aligned}
 \tag{23}$$

There are two decomposition forms for each factor and  $2J$  forms for all factors. An aggregation across all  $J$  factors does not provide an exact decomposition.

*Shortcut 4* Factorisation of the change in the Leontief inverse—only for additive decomposition

We will finally compute the difference of  $\mathbf{x}$  attributed to the change in each factor that builds on the factorisation of  $\Delta \mathbf{L}$  as it appears in Eqs. (18a)–(18b):

$$\Delta \mathbf{x}(\mathbf{A}_j)_{S4} = \frac{1}{2} \left( \mathbf{L}^{(1)} (\Delta \mathbf{A}_j) \mathbf{L}^{(0)} + \mathbf{L}^{(0)} (\Delta \mathbf{A}_j) \mathbf{L}^{(1)} \right) \frac{1}{2} (\mathbf{f}^{(1)} + \mathbf{f}^{(0)}).
 \tag{24}$$

The above computes two decomposition forms for each factor, merged into one expression, and  $2J$  forms for all factors. The decomposition is exact. Although shortcut 4 cannot be interpreted as a correct measure of the effect of the change in  $\mathbf{A}_j$ , it can still provide an approximation of the true result, and we will see whether that approximation is good.

### 3 Results and discussion

#### 3.1 Testing strategy

We will evaluate the performance of the shortcuts in Sect. 2.3 according to the following logic:

1. Compute the true vectors of changes  $\Delta \mathbf{x}(\mathbf{A}_j)$  in additive decomposition and  $\mathbf{Px}(\mathbf{A}_j)$  in multiplicative decomposition as, respectively, Bennet indicators and Fisher indices for each factor  $j$ . An aggregation across all  $J$  factors must yield a vector of changes in industry output induced by the changes in technical coefficients and, hence, the changes in the Leontief inverse:  $\sum_{j=1}^J (\Delta \mathbf{x}(\mathbf{A}_j)) = \Delta \mathbf{x}(\mathbf{A})$  and  $\odot_{j=1}^J (\mathbf{Px}(\mathbf{A}_j)) = \mathbf{Px}(\mathbf{A})$ , which are equal to the first terms on the right-hand side of,

respectively, Eqs. (15) and (16). This is the consistency-in-aggregation property that shortcuts 1 and 3 do not have.

2. Compute the estimates of the vectors of changes using shortcuts 1–4 for additive decomposition and shortcuts 1–2 for multiplicative decomposition for each  $j$ th factor and aggregates for all  $J$  factors.
3. Use standard matrix comparison methods to evaluate the performance of the tested shortcuts. In related literature, it is customary to use an array of measures of distance that meet researcher's criteria. A measure of distance is a scalar that weighs, scales and aggregates the differences between each element of an estimated matrix (vector) and the respective element of the true matrix (vector). The best estimate should be the closest, i.e., should have the smallest distance, to the true values. A survey of literature led to a selection of 12 measures of distance listed in the Additional file 1. To avoid misleading conclusions, measures of similarity rather than distance were disregarded.<sup>6</sup> Calculate all selected measures of distance for each shortcut and each  $j$ th factor. Identify the best-performing shortcut by each measure of distance and for each  $j$ th factor. Then aggregate the information across all 12 measures and select the best-performing shortcut for each  $j$ th factor. Finally, summarise the results for all  $J$  factors, selecting one best-performing shortcut (or, if not possible, two or more shortcuts). In the case of additive decomposition, the measures of distance will also be calculated for the aggregate change representing all  $J$  factors to find out whether shortcut 1 or 3 provides estimates that are closer to the exact decomposition.
4. Calculate an additional summary statistic for each shortcut: the share of entries in the estimated vectors of changes for each  $j$ th factor where this shortcut provides the best estimate, i.e., the closest to the true value. This statistic ignores the average distance to the true results and, rather, indicates a success rate in producing the best estimates. Use this percentage as the score to identify the best-performing shortcut.

### 3.2 Data management strategy

Input–output data for the numerical experiment can be generated from random numbers or can be drawn from recognised sources of statistical information. As the structure of real input–output tables is not entirely random, the latter option is preferred here, which is also in line with previous studies of SDA (e.g., de Boer and Rodrigues, 2020). Data selection and processing steps are as follows:

- Look for time series of input–output tables at current prices and previous year's prices, preferably, released by national statistical offices. A survey of official releases of national input–output data available online revealed two sources: Statistics Denmark and Statistics Netherlands. Statistics Denmark maintains perhaps the longest annual series of input–output tables starting from 1966 in 117-industry and 69-industry classifications of which the latter is sufficient for the purpose of this paper. Statistics Netherlands offers annual input–output tables starting from 1995 in

<sup>6</sup> Such measures as R-squared (coefficient of determination) and various information-based statistics are invariant to scaling: multiplying an estimate by a scalar—and, therefore, magnifying the differences with the true value—does not change the result of the comparison (see e.g., Steen-Olsen et al. 2016).

76-industry classification. Both statistical offices provide the tables at current prices and previous year's prices which ensures the consistency of structural decomposition. For the proposed numerical test, the time coverage of tables was limited to 2006–2015.<sup>7</sup>

- Adapt and aggregate input–output tables. Input–output tables for both countries were aggregated to four alternate classifications with 5 to 8 industries to ensure that the structural decomposition is computationally manageable.<sup>8</sup> The aggregation schemes (available with the supporting dataset, see the data availability statement) are not random and are designed to delineate larger classes of economic activities.
- Run the SDA on the aggregated input–output tables for Denmark and the Netherlands according to the testing strategy with 5 to 8 factors that correspond to the changes in the 'production recipe' or technology of the aggregate industries. Ignore other factors that are not related to the changes in technical coefficients.

In terms of the number of factors and time coverage, the data input in this numerical experiment is larger than in many previous studies of SDA techniques. For interested readers, executable codes that can be used to replicate the computations are in Additional file 2.

### 3.3 Results of the numerical experiment

The results of the numerical experiment, including the true and estimated vectors of changes by factor and indicators of their performance, are deposited at the figshare repository and are available at the link that can be found in the data availability statement. Tables 3, 4, 5, 6 and 7 summarise the results, indicating which shortcut provides the best estimates for each set of factors and each of the 9 yearly intervals. The last row identifies the single best-performing shortcut for all intervals. Two or three shortcuts appear in one cell in the cases where these shortcuts perform equally well.

Tables 3 and 4 reveal that, by and large, shortcut 1 outperforms other shortcuts in additive SDA. The results are not uniform and vary across time and the number of factors. For example, in the additive SDA of industry output of the Netherlands in 2010–2011, shortcut 4 provides the best estimates with five factors corresponding to the five aggregate industries, shortcut 1 ranks the best with six factors, shortcut 3 with seven factors, and both shortcuts 1 and 3 provide the best estimates with the number of factors increasing to eight, as measured by the distance to the true results. For any set of factors, except five-factor SDA for the Netherlands, shortcut 1 provides the estimates at minimal overall distance from the true values in most yearly intervals (see Table 3).

<sup>7</sup> For Denmark, retrieved from [https://www.dst.dk/ext/4372765735/0/inout/Excel-files-with-IO-data-for-the-period-2006-2015-\(69-industries\)--zip](https://www.dst.dk/ext/4372765735/0/inout/Excel-files-with-IO-data-for-the-period-2006-2015-(69-industries)--zip), accessed 6 February 2020. For the Netherlands, retrieved from [https://www.cbs.nl/-/media/\\_excel/2016/35/input-output-tables-1995-2015.xls](https://www.cbs.nl/-/media/_excel/2016/35/input-output-tables-1995-2015.xls), accessed 10 December 2019.

<sup>8</sup> In the Dutch tables, trade and transport margins are exogenous to the intermediate consumption. There is a column vector of trade and transport margins that are recorded as products supplied by domestic industries or as imported products. There is also a row vector that allocates those margins to the intermediate demand by industries and to final demand. For consistency with the usual input–output table layout and with the Danish tables, the supply of trade and transport margins in the Netherlands was distributed between the intermediate and final demand and across the purchasing industries. The matrix of imputed flows of trade and transport margins was constructed with the RAS method using the matrix of intermediate transactions and the aggregated vector of final demand as prior estimates. This did not involve a change of valuation from basic to purchasers' prices.

**Table 3** Best-performing shortcuts to the complete additive SDA, by measured distance to the true results

Period	5 factors		6 factors		7 factors		8 factors	
	DNK	NLD	DNK	NLD	DNK	NLD	DNK	NLD
2006–2007	1	4	1	1,4	1	1	1	1
2007–2008	3	2	3	1	3	1,4	3	2
2008–2009	1	1	3	1	3	1	1,3	1
2009–2010	1	2	1	2,3	1	1,2,4	1	2
2010–2011	1	4	1	1	1,2,4	3	3	1,3
2011–2012	3	2	2,3	1	3	3	3	3
2012–2013	1,2	2	1	2	1	1,4	1	4
2013–2014	1	1	1	1	1	1	2,3	1
2014–2015	1	1	1	1	1	1	1	1
All	1	2	1	1	1	1	1	1

NLD stands for the Netherlands and DNK for Denmark. The values in the table are numbers of the shortcuts as listed in Sect. 2.3

Source: author's calculations

**Table 4** Best-performing shortcuts to the complete additive SDA, by share of the best estimates

Period	5 factors		6 factors		7 factors		8 factors	
	DNK	NLD	DNK	NLD	DNK	NLD	DNK	NLD
2006–2007	1	2	2	2	1	1	1,3	1
2007–2008	2	2	3	1	3	2	3	2
2008–2009	4	1	2	1,2	3	1	3	1
2009–2010	1	2	1	4	1,2	2	1	2
2010–2011	1,3	4	1	1,4	1	3	3	3
2011–2012	3	1,2	3	1	3	3	3	3
2012–2013	1	2	3	2	1	2	2	2
2013–2014	1	2	1	1	1,2	1	2	1
2014–2015	1	2	2	1	2	1	2	1
All	1	2	1,2,3	1	1	1	3	1

NLD stands for the Netherlands and DNK for Denmark. The values in the table are numbers of the shortcuts as listed in Sect. 2.3

Source: author's calculations

The prevalence of shortcut 1 is somewhat weaker if judged by the share of the best estimates that this shortcut yields. Still shortcut 1 ensures the largest share of best estimates in most yearly intervals for all set of factors except five-factor SDA for the Netherlands, six-factor and eight-factor SDA for Denmark (see Table 4).

Table 5 shows that shortcut 1 outperforms shortcut 3 in terms of the distance of the aggregate estimates to the aggregate true results.

A reasonable indicator of the accuracy of estimates by shortcut 1 in additive SDA is the mean absolute percentage error (MAPE, see Additional file 1) that is calculated for each  $j$ th factor and each yearly interval. MAPE is rather small: maximum 0.12% for Denmark and 0.21% for the Netherlands.

In multiplicative SDA, shortcut 1 is the best-performing approach in the vast majority of cases, as Tables 6 and 7 show. A possible reason is that shortcuts 3 and

**Table 5** Best-performing shortcuts to the complete additive SDA, by measured distance to the aggregate true results

Period	5 factors		6 factors		7 factors		8 factors	
	DNK	NLD	DNK	NLD	DNK	NLD	DNK	NLD
2006–2007	1	1	1	3	1	1	1	1
2007–2008	3	1	1	3	3	1	3	3
2008–2009	3	1	1	3	3	1	3	1
2009–2010	1	1	1	1	1	1	1	1
2010–2011	1	3	3	1	1	3	3	3
2011–2012	3	3	3	1	3	3	3	1
2012–2013	1	1	1	1	1	1	1	1
2013–2014	1	1	1	3	1	1	1	1
2014–2015	1	1	1	1	1	1	1	1
All	1	1	1	1	1	1	1	1

NLD stands for the Netherlands and DNK for Denmark. The values in the table are numbers of the shortcuts as listed in Sect. 2.3

Source: author's calculations

**Table 6** Best-performing shortcuts to the complete multiplicative SDA, by measured distance to the true results

Period	5 factors		6 factors		7 factors		8 factors	
	DNK	NLD	DNK	NLD	DNK	NLD	DNK	NLD
2006–2007	1	1	1	1	1	1	1, 2	1
2007–2008	1	2	1	1	1	1	1	1
2008–2009	1	1	1	1	1	1	1	1
2009–2010	1	2	1	1	1	1	1	1
2010–2011	1	1	1	2	1	1	1	1
2011–2012	1	2	1	1	1	2	1	2
2012–2013	1	1	1	1	1	1	1	1
2013–2014	1	2	1	1	1	1	1, 2	1
2014–2015	1	1	1	1, 2	1	2	1	2
All	1	1	1	1	1	1	1	1

NLD stands for the Netherlands and DNK for Denmark. The values in the table are numbers of the shortcuts as listed in Sect. 2.3

Source: author's calculations

4 do not apply to multiplicative SDA. As the results of multiplicative decomposition are ratios, we can judge the accuracy of estimates by considering the mean absolute difference (MAD, see Additional file 1) transformed into percentage form, i.e., multiplied by 100. The maximum MAD for Denmark is 0.000066 percentage points and for the Netherlands 0.000021 percentage points.

Another important observation from Tables 3 and 4 is that shortcut 4—that builds on the factorisation of the change in the Leontief inverse and is largely accepted in the literature (e.g., Oosterhaven and Hoen 1998; Miller and Blair 2009, chap.13)—produces a bias in the results of additive SDA, and this bias tends to be larger than that of other shortcuts.



**Table 7** Best-performing shortcuts to the complete multiplicative SDA, by share of the best estimates

Period	5 factors		6 factors		7 factors		8 factors	
	DNK	NLD	DNK	NLD	DNK	NLD	DNK	NLD
2006–2007	1	1	1	1	1	1	1	1
2007–2008	1	1	1	1	1	1	1	1
2008–2009	1	1	1	1	1	1	1	1
2009–2010	1	2	1	1	1	1	1	1
2010–2011	1	1	1	1	1	1	1	1
2011–2012	1	2	1	1	1	2	1	1
2012–2013	1	1	1	1	1	1	1	1
2013–2014	1	1	1	1	1	1	1	1
2014–2015	1	1	1	1	1	1	1	1
All	1	1	1	1	1	1	1	1

NLD stands for the Netherlands and DNK for Denmark. The values in the table are numbers of the shortcuts as listed in Sect. 2.3

Source: author's calculations

#### 4 Conclusion

When the number of factors that are nested within the Leontief inverse increases to hundreds or thousands, the structural decomposition that is consistent with the index number theory becomes incredibly cumbersome or even infeasible. Researchers need a reasonable approximation with minimum computations and maximum accuracy of results. A solution that is largely acceptable in the literature is to take the average of the two polar decompositions. However, the selection of such polar decompositions is entirely arbitrary: any decomposition with its ‘mirror image’ can be treated as polar.

This paper tested three ad hoc solutions known from the literature and one new variant to estimate the changes in the dependent variable induced by the changes of multiple factors within the Leontief inverse. All four solutions avoid arbitrary choice of decompositions, or the order of changes in factors, and require computing only two decomposition forms for each factor. The numerical experiment with time series of real input–output tables produced evidence that supports employing one of the four ad hoc solutions: taking the average of the two polar decomposition forms for each factor. It is more likely to obtain estimates with minimum errors with this solution, rather than with the other solutions.

The above finding is thought to be in particular useful for SDA that involves inter-country input–output tables. With a reliable ‘shortcut’ to complete decomposition, investigators should be able to identify the country and industry sources of changes that affect the performance of a country and/or industry of interest. While the bias this ‘shortcut’ minimises may be small, it is important that the selection of computational approach is not arbitrary and is based on sound evidence.

#### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40008-021-00245-5>.

**Additional file 1.** Measures of distance used for the comparison of estimated and true vectors of changes.

**Additional file 2.** Executable codes that were used to run the numerical experiment (including a “read me” file).

**Acknowledgements**

Not applicable.

**Authors' contributions**

Not applicable.

**Funding**

The author has not received funding from any sources for this research.

**Availability of data and materials**

The dataset supporting the conclusions of this article is available in the figshare repository, <https://doi.org/10.6084/m9.figshare.12478868.v1>.

**Declarations****Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Competing interests**

The author declares that he has no competing interests.

Received: 25 September 2020 Revised: 8 August 2021 Accepted: 9 August 2021

Published online: 30 August 2021

**References**

- de Boer P (2008) Additive structural decomposition analysis and index number theory: an empirical application of the Montgomery decomposition. *Econ Syst Res* 20(1):97–109. <https://doi.org/10.1080/09535310801892066>
- de Boer P (2009) Multiplicative decomposition and index number theory: an empirical application of the Sato-Vartia decomposition. *Econ Syst Res* 21(2):163–174. <https://doi.org/10.1080/09535310902937638>
- de Boer P, Rodrigues JFD (2020) Decomposition analysis: when to use which method? *Econ Syst Res* 32(1):1–28. <https://doi.org/10.1080/09535314.2019.1652571>
- Dietzenbacher E, Los B (1998) Structural decomposition techniques: sense and sensitivity. *Econ Syst Res* 10(4):307–324. <https://doi.org/10.1080/09535319800000023>
- Dietzenbacher E, Hoen AR, Los B (2000) Labor productivity in Western Europe 1975–1985: an intercountry, interindustry analysis. *J Region Sci* 40(3):425–452. <https://doi.org/10.1111/0022-4146.00182>
- de Haan M (2001) A structural decomposition analysis of pollution in the Netherlands. *Econ Syst Res* 13(2):181–196. <https://doi.org/10.1080/09537320120052452>
- Koller W, Stehrer R (2010) Trade integration, outsourcing and employment in Austria: a decomposition approach. *Econ Syst Res* 22(3):237–261. <https://doi.org/10.1080/09535314.2010.493504>
- Miller RE, Blair PD (2009) *Input-output analysis: foundations and extensions*, 2nd edn. Cambridge University Press, New York
- Oosterhaven J, Hoen A (1998) Preferences, technology, trade and real income changes in the European Union: an intercountry decomposition analysis for 1975?1985. *Ann Reg Sci* 32:505–524. <https://doi.org/10.1007/s001680050086>
- Rose A, Casler S (1996) Input-output structural decomposition analysis: a critical appraisal. *Econ Syst Res* 8(1):33–62. <https://doi.org/10.1080/09535319600000003>
- Seibel S (2003) Decomposition analysis of carbon dioxide emission changes in Germany—conceptual framework and empirical results. Luxembourg: Office for Official Publications of the European Communities. <https://ec.europa.eu/eurostat/documents/3888793/5817109/KS-DB-03-002-EN.PDF/f4f5f9a0-5336-412d-93f2-7e0053b744b7>. Accessed 30 June 2020
- Siegel IH (1945) The generalized “ideal” index-number formula. *J Am Stat Assoc* 40(232):520–523
- Steen-Olsen K, Owen A, Barrett J, Guan D, Hertwich EG, Lenzen M, Wiedmann T (2016) Accounting for value added embodied in trade and consumption: an intercomparison of global multiregional input-output databases. *Econ Syst Res* 28(1):78–94. <https://doi.org/10.1080/09535314.2016.1141751>
- Su B, Ang BW (2012) Structural decomposition analysis applied to energy and emissions: some methodological developments. *Energy Econ* 34(1):177–6188. <https://doi.org/10.1016/j.eneco.2011.10.009>
- Wang H, Ang BW, Su B (2017) Assessing drivers of economy-wide energy use and emissions: IDA versus SDA. *Energy Policy* 107:585–599. <https://doi.org/10.1016/j.enpol.2017.05.034>

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.