`rdecompose`: Decompose Aggregate Values in Stata

Jinjing Li[1] Yohannes Kinfu[2]

[1]NATSEM, Institute of Governance and Policy Analysis
[2]CeRAPH, Health Research Institute
University of Canberra

24 September 2015
Oceania Stata User Group Meeting 2015

Outline

Decomposition

- Micro-data based decomposition
  - Blinder-Oaxaca decomposition approach etc.
- Macro-data based decomposition
  - Rate Decomposition (Kitagaw and Gupta's decomposition models, various Gini decomposition etc.)

Decomposition in Stata

- Well developed commands for Blinder-Oaxaca type decomposition (e.g. oaxaca, nldecompose)
- Some highly specialised rate decomposition (e.g. Gini decomposition with descogini)
- No readily available command for general rates/aggregate data decomposition

Generalised Rate Decomposition

- Assume that the rate r can be expressed by k factors

$$r\left(x_1 \cdots x_k\right) = \prod_{i=1}^{k} x_i$$

2 Factors Decomposition

- In the case of $k = 2$

$$\begin{cases} C(x_1) = & \frac{1}{2}(x_2^a + x_2^b)(x_1^a - x_1^b) \\ C(x_2) = & \frac{1}{2}(x_1^a + x_1^b)(x_2^a - x_2^b) \end{cases}$$

- Intuitively speaking, the contribution of the factor is conditionally on the mean value of the other factors.
- We then standardised the contribution from $C(x_1)$ and $C(x_2)$

3 or more factors

Different specification

- In the case where $k \geq 3$

$$C(x_i) = \sum_{j=1}^{k-1} \frac{R(j-1, i)}{k \begin{pmatrix} k-1 \\ j-1 \end{pmatrix}} (x_i^a - x_i^b)$$

- where $R(j, i)$ is the sum of all possible values of the product of $k - 1$ factors (excluding x_i), out of which j factors from population a and all other factors from population b.

3 or more factors

Generalised specification

- It is also possible that the rate function $r\left(x_1 \cdots x_k\right)$ is more than the simple product function, e.g. $newborn = \sum_{age} fertility \cdot women_{age}$
- sum over a specified group is a common operation in cross-classifed data

3 or more factors

Different specification

- The number of permutations increase much faster than k
  - for six factor decomposition, we need to calculate $r(\cdot)$ 192 times
  - No publicly available software (in any language or statistical package) to handle large k
  - Gupta published some Fortran code for small and medium size k but it requires the end user to tweak the code for each case
- In most cases, we also need to aggregate values over a group (e.g. age groups, location etc.)
- Mostly done in Excel or manual calculations, which are prone to mistakes

rdecompose

- We developed a new Stata command **rdecompose** to assist decompositions using Gupta's method
- rdecompose currently supports decomposition where the aggregate rates r is calculated based on k factors, and aggregated over s, i.e. $r = \sum_s f(x_1 \cdots x_k)$
  - ability to decompose with any arbitrary number of factors
  - ability to automatically aggregate values over a group
  - ability to specify non-standard functional form instead of product only (e.g. $x_1 e^{x_2} \ln(x_3 + x_4)$)
  - ability to interact with other commands for further processing
  - in Stata

rdecompose syntax

rdecompose *variables* [**if** exp], **group**(variable) [**sum**(varlist) **detail reverse function**(string) **transform**(variable) **multi baseline**(#)]

variables: factors that contribute to the rates

group: population identifier (string or numeric)

sum: indicates the rate is the sum of the specified variable (Default: none)

function: specifies the funciton form (Default: $f(\cdot) = \prod_{i=1}^{k} x_i$)

Example 1: Data on total fertility and proximate determinants of fertility

- Example from Gupta(1994)
- Decompose total fertility rate in Korea between 1960 and 1970
- Following Moreno(1991)

$$TFR = C_m C_c C_x \cdots C_{others}$$

Example 1: Data on total fertility and proximate determinants of fertility

- Data as in Stata

year	Marriage	Contraception	Abortion	Lactation	Fecundity
1970	.58	.76	.84	.66	16.573
1960	.72	.97	.97	.56	16.158

Table: Fertility Rate Decomposition in Korea

Example 1: Data on total fertility and proximate determinants of fertility

```
. rdecompose Marriage Contraception Abortion Lactation Fecundity , group(year)

Decomposition between year == 1960 (6.13)
                  and year == 1970 (4.05)
Func Form = Marriage*Contraception*Abortion*Lactation*Fecundity
```

Component	Absolute Difference	Proportion (%)
Marriage	−1.09	52.46
Contraception	−1.23	59.13
Abortion	−.728	35.00
Lactation	.84	−40.38
Fecundity	.129	−6.20
Overall	−2.08	100.00

Number of Obs : 10

Example 2: Data on demand for additional children in Nepal

- Example data from Clogg and Eliason(1998) on population size and percent desiring more children
- Decompose the difference between women with one child and women with 4+ children

age group	Age composition	Rate	Parity
20-24	27	37.037	One child
25-29	152	19.079	One child
	....		
20-24	363	90.083	4+ Children
25-29	208	76.923	4+ Children

Table: National Fertility Survey, Clogg and Eliason(1988)

Example 2: Data on demand for additional children in Nepal

```
rdecompose Age_composition Rate , group( Parity ) transform( Size ) sum(
     age_group )

Decomposition between Parity == 1 (11.49)
                  and Parity == 2 (72.09)
Func Form = \sum(age_group){Size*Rate}
```

Component	Absolute Difference	Proportion (%)
Age_composition(*)	23.1	38.07
Rate	37.5	61.93
Overall	60.6	100.00

(*) indicates transformed variables   Number of Obs : 20

Example 3 : Global Burden of Disease Data

- Latest IHME Data (2015)
- Decompose the mortality rate due to *ageing effect* and the change in *Communicable disease*, *Non-communicable disease* and *Injuries* between developed and developing countries.

Example 3 : Global Burden of Disease Data

- Data as in Stata

age_group	age_structure	CDM	NCD	Injuries	group
		......			
1-4 years	0.464026	284.31	55.8	46.33	1
5-9 years	0.054843	47.17	18.6	19.13	1
		......			
80 years and above	0.006002	1433.16	11589.4	491.64	1
		......			
1-4 years	0.044478	4.39	12.23	9.98	2
5-9 years	0.053709	1.21	6.19	5.64	2
		......			
80 years and above	0.045763	650.37	9489.93	329.66	2

Table: Disease Burden between Developed and Developing countries

Example 3 : Global Burden of Disease Data

```
. rdecompose age_structure CDM NCD Injuries , group(group) sum(age_group_i) func
      (age_structure*( CDM + NCD + Injuries ))

Decomposition between group == 1 (576.46)
                  and group == 2 (993.90)
Func Form = \sum(age_group_i){age_structure*( CDM + NCD + Injuries )}
```

Component	Absolute Difference	Proportion (%)
age_structure	843	201.88
CDM	−200	−47.86
NCD	−193	−46.17
Injuries	−32.8	−7.85
Overall	417	100.00

Number of Obs : 160

Example 4 : China Health Expenditure

- Ongoing project between Health Research Institute at University of Canberra and China National Health Develop Research Centre
- Decompose the increase of total health expenditure between 1993 to 2012 into
  - Prevalence by age and disease group
  - Population Size
  - Demographic Ageing
  - Expenditure per case
  - Excess Health Inflation

Example 4 : China Health Expenditure

```
Decomposition between year == 1993 (124535.22)
                  and year == 2012 (2475451.03)
Func Form = \sum(disease_group)\sum(agegroup){prevalencerate*population*ageing*
      exppercase*healthpriceinflation}
```

Component	Absolute Difference	Proportion (%)
prevalencerate	−77942	−3.32
population	139708	5.94
ageing	191016	8.13
exppercase	1536974	65.38
healthpricein~n	561159	23.87
Overall	2350916	100.00

Saved Results

- rdecompose also returns some results in scalar/macro/matrix format for further processing
  - Scalar e(N) contains the number of observations used in the estimation
  - Scalar e(rate1) contains the rate calculated for the first group
  - Scalar e(rate2) contains the rate calculated for the second group
  - Scalar e(diff) shows the total differences between two groups
  - Macro e(basegroup_value) shows the baseline group value
  - Matrix e(b) contains the total contributions for each factor

Summary

- **rdecompose** decomposes aggregated rates from two populations using Gupta's decomposition model
  - Wide range of applications in the field of demography, health, economics etc.
  - Support flexible a functional form and avoid cumbersome calculations with a large number of factors

- Limitations and Next steps
  - Other common data transformations
  - Other rate decomposition models
  - Improved support for more than two populations (currently with the option multi)

ECONOMIC SYSTEMS RESEARCH
https://doi.org/10.1080/09535314.2019.1652571

Decomposition analysis: when to use which method?

Paul de Boer[a] and João F. D. Rodrigues [b]

[a]Econometric Institute, Erasmus School of Economics, Erasmus University Rotterdam Rotterdam, the Netherlands; [b]Institute of Environmental Sciences CML, Leiden University, Leiden, the Netherlands

ABSTRACT

Structural and index decomposition analyses allow identifying the main drivers of observed changes over time of energy and environmental impacts. These decomposition analyses have become very popular in recent decades and, many alternative methods to implement them have become available. Several of the most popular methods have been developed earlier in index number theory, a context in which each particular method is defined by adhering to a set of properties. The goal of the present paper is to review the main results of index number theory and discuss its connection to decomposition analyses. By doing so, we can present a decision tree that allows users to choose a decomposition method that meets desired properties. We report as hands-on example an empirical case study of the carbon footprint of the Netherlands in the period 2004–2005.

ARTICLE HISTORY
Received 17 December 2018
In final form 2 August 2019

KEYWORDS
Index number theory; index decomposition analysis; structural decomposition analysis; decision tree

1. Introduction

In recent decades, many studies have attempted to identify the drivers of observed changes over time of energy and environmental impacts (Hoekstra and van den Bergh, 2003; Su and Ang, 2012). Such decomposition analyses can fall under two distinct but related categories: index decomposition analysis (IDA), in which the link between impact (energy, environmental, employment or whatever) and production level is explored; and structural decomposition analysis (SDA), in which the link between impact and consumption activities is explored. Hence, SDA is more comprehensive than IDA (since it requires explicitly accounting for the link between production and consumption) but also requires more data. A practitioner who is new in the field of IDA and/or SDA can find much advice in literature, for example Ang (2004), Ang et al. (2009), Su and Ang (2012), Ang (2015) and Wang et al. (2017b). In fact, there are so many possible references that a general overview is tough to disentangle; in other words the new practitioner will have trouble to 'see the forest for the trees'. Moreover, the mathematics employed and notation used is usually unduly complicated. In this paper, we follow a different route: we go back to the 'forest', that is to say to the collective stock of knowledge called index number theory in which the mathematics

CONTACT João F. D. Rodrigues j.rodrigues@leidenuniv.nl Institute of Environmental Sciences CML, Leiden University, Einsteinweg 2, 2333 CC, Leiden, the Netherlands

Supplemental data for this article can be accessed here. https://doi.org/10.1080/09535314.2019.1652571

is usually easily accessed. Then, we turn to IDA and SDA, which – as one of the referees correctly points out- ‘in quite a number of cases proceeds by, knowingly or unknowingly, re-inventing or replicating well-known results from index number theory.’

This paper is structured as follows. In Section 2, we review the relevant properties and formulas of index number theory and, in Section 3, discuss its connection with IDA and SDA. Section 4 deals with the general case of ideal decompositions (index number theory), which leave no residual term (IDA and SDA), of an aggregate change into of n factors. Sections 5 and 6 are devoted to a hands-on example of a decomposition of carbon dioxide emissions of sectors of the Dutch economy into five factors. In the supplementary material, to enable replication we provide in the Appendices D–G our Matlab programs and Excel files used in the hands-on example. Based on theory and empirics we present in Section 7, a decision tree that we hope will help practitioners who seek to select the appropriate decomposition method for their problem at hand. Section 8 concludes.

2. Index number theory

2.1. Historical background

Traditionally index number theory is about the measurement of aggregate price and quantity change. Prominent applications are the consumer price index (CPI), producer price index (PPI), purchasing power parity (PPP) and human development index (HDI.) There is a vast literature on this subject. The oldest price index is attributed to the French economist Dutot (1738) who disposed of data on prices of several commodities in 1515 (base period), under the reign of King Louis XII, and in 1738 (comparison period), under the reign of King Louis XV. He took the arithmetic mean of the prices in the comparison period (1738) and of the prices in the base period (1515) and divided them by each other. According to the price index of Dutot the price level was multiplied by the factor 22. Dutot concluded that Louis XV was worse off, when compared to his ancestor, because his income was only multiplied by the factor 13. The price index of Dutot depends on the units of measurement (the price of, say, salt, can be measured per ounce or per pound) so that the outcome is arbitrary. The Italian economist Carli (1764) constructed the first index free from the units of measurement using price relatives, that is to say he divided the price of, say one ounce of salt, in the comparison period by the corresponding price in the base period, and took the arithmetic mean of the price relatives. The disadvantage of Carli’s approach is that it does not take into account the importance of a commodity in the budget of a consumer who will spend in (say) one year more money on bread than on salt. According to Balk (2008a) the first person who recognized the necessity of introducing weights into a price index was Young (1812). The two most important contributions in the nineteenth century are undoubtedly the ones by Laspeyres (1871) and Paasche (1874). Laspeyres considered a basket of commodities in the base period and computed the total expenditure. He also computed total expenditure of this basket in the comparison period. His price index is the ratio of the total expenditure in the comparison period and the total expenditure in the base period. Laspeyres’ choice of the base period’s basket is arbitrary. Paasche chose the basket of the comparison period. In his seminal book, Fisher (1922) introduced time reversal and the product test as two desirable properties of indices (to be described later in greater detail). Laspeyres and Paasche do not meet time reversal. Fisher proposed as new price (quantity) index the geometric mean of the price (quantity) indices of Laspeyres and Paasche which satisfies time reversal. He proved that from all at that time existing pairs of price and quantity indices only the product of his pair satisfied the product test. That is the reason why he called his pair of indices 'ideal'. After the contribution of Fisher two other pairs of ideal index numbers were discovered. Montgomery (1929, 1937) and, independently, Vartia (1974; 1976), proposed a solution, the so-called Montgomery-Vartia index. Another solution was, independently, proposed by Sato (1976) and Vartia (1974; 1976), the so-called Sato-Vartia index.

Above, we give a short history of the classical index number problem in which one wants to decompose the ratio value change in expenditure into the product of two factors, called price and quantity 'indices' (singular 'index'). The alternative problem, lesser known but equally old (and more relevant in the context of index and structural decomposition analysis), is the decomposition of the difference between the values of two expenditures as the sum of two parts, the price and quantity 'indicators', which measure respectively the changes due to price and quantity differences. The indicators of Bennet (1920) are the additive counterpart of the indices of Fisher. Another pair of ideal indicators was discovered by Montgomery (1929; 1937).

The remainder of this section is organized as follows. Section 2.2 is devoted to the notation, problem formulation and properties used to characterize indices and indicators. The following subsections then report specific indices and indicators. For computational purposes these turn out to cluster around two basic methods: Fisher and Bennet methods, which are combinatorial in nature (addressed in Section 2.3), while Montgomery-Vartia, Sato-Vartia and Montgomery involve logarithmic transformations (addressed in Section 2.4). In Section 2.5 we discuss another important contribution, viz. the concept of consistency-in-aggregation. Section 2.6, finally, summarizes the relevant properties and formulas of index number theory for IDA and SDA.

2.2. Notation, problem formulation and properties

2.2.1. Notation

$p_i^t > 0$: price of commodity i ($i = 1, \ldots, N$)in base ($t = 0$) or in comparison period ($t = 1$)

$q_i^t > 0$: quantity of commodity i ($i = 1, \ldots, N$)in base ($t = 0$) or in comparison period ($t = 1$)

v_i^t: value of commodity i in base ($p_i^0 q_i^0$) or in comparison period ($p_i^1 q_i^1$) (1)

V^t: total value in base $\left(\sum_{i=1}^N v_i^0\right)$ or in comparison period $\left(\sum_{i=1}^N v_i^1\right)$

s_i^t: share of value of commodity i in total value in base or in comparison period (2)

RV : ratio change in value : $V^1/V^0 = \frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^0 q_i^0}$ (3)

RP: price index measuring the price change from the base to the comparison period

RQ: quantity index measuring the quantity change from the base to the comparison period

DV: difference change in value: $V^1 - V^0 = \sum_{i=1}^N (v_i^1 - v_i^0) = \sum_{i=1}^N p_i^1 q_i^1 - \sum_{i=1}^N p_i^0 q_i^0$ (4)

DP: price indicator measuring the change in prices from the base to the comparison period

DQ: quantity indicator measuring the change in quantities from the base to the comparison period.

2.2.2. Problem formulation and properties

We quote from the abstract of Balk (2008b): 'The index number problem is known as that of decomposing aggregate value change, in ratio or in difference form, into two, ideally symmetric, factors.' In our notation:

$$RV = RP \times RQ \tag{5}$$

and

$$DV = DP + DQ. \tag{6}$$

The basic properties (Eichhorn and Voeller, 1976, and Eichhorn, 1978) on the price and quantity indices in Equation 5 comprise:

- global monotonicity: the price (quantity) index is non-decreasing in comparison prices and non-increasing in base prices (quantities);
- linear homogeneity in comparison prices (quantities): if all comparison prices are multiplied by a common factor, the price (quantity) index is multiplied by that common factor, as well;
- identity: if all comparison prices (quantities) in the comparison period are the same as those in the base period, the price (quantity) index is equal to one;
- homogeneity of degree zero in prices (quantities): if all prices (quantities) in comparison and base period are multiplied by a common factor, the price (quantity) index remains the same;
- invariance to changes to the units of measurement of the commodities.

A price (quantity) index that satisfies the requirements of 'linear homogeneity in comparison prices (quantities)' and of 'identity' satisfies a stronger requirement, namely that of

- proportionality with respect to prices (quantities). If all the individual price (quantity) relatives are the same, then the price (quantity) index number must be equal to these relatives.

Other desirable properties (Fisher, 1922) are:

- time reversal (symmetry): the price (quantity) index for the base period relative to the comparison period must be equal to the reciprocal of the price (quantity) index for the comparison period relative to the base period;
- product test: it requires that the ratio change in value can be decomposed as product of a price and a quantity index, like in Equation 5;

- idealness: a pair of price and quantity indices which, in addition to satisfying the product test, also have the same functional form, i.e. by interchanging prices and quantities the price index turns into the quantity index and vice versa, is called 'ideal'.

Indicators return monetary values, which may be negative or zero. Therefore the basic properties must be modified a bit (Diewert, 2005, Balk, 2008a). They comprise: global monotonicity, modified identity (if all comparison prices and quantities in the comparison period are the same as those in the base period, the price (quantity) index is equal to zero); homogeneity of degree 1 in prices (quantities), invariance to changes to the units of measurement of the commodities. Please note that there is ***no*** analogue to the property of 'linear homogeneity in comparison prices (quantities)', so that there is ***no*** analogue of 'proportionality' for indicators, as well. Other desirable properties are:

- time reversal (symmetry): the price (quantity) indicator for the base period relative to the comparison period must be equal to the opposite of the price (quantity) indicator for the comparison period relative to the base period;
- sum test: the sum test requires that the difference change in value can be decomposed as the sum of a price and a quantity index, like in Equation 6;
- idealness: a pair of price and quantity indicators which, in addition to satisfying the sum test, also have the same functional form, i.e. by interchanging prices and quantities the price indicator turns into the quantity indicator and vice versa, is called 'ideal'.

The property of idealness is an important contribution of index number theory to structural and index decomposition analysis since it implies that the decomposition is '*complete*', that is to say that there is *no residual term*. We refer to the review of index number theory of Balk (2016) and to Balk's monograph (Balk 2008a) for the mathematical presentation.

2.3. Combinatorial indices and indicators

2.3.1. Introduction

This class consists of the 'traditional' indices and indicators of Laspeyres (superscript L), Paasche (superscript P), Fisher (superscript F) and its additive counterpart Bennet (superscript B).

2.3.2. Indices

The indices are defined as:

$$RP^L = \frac{\sum_{i=1}^N p_i^1 q_i^0}{\sum_{i=1}^N p_i^0 q_i^0};\ RQ^L = \frac{\sum_{i=1}^N p_i^0 q_i^1}{\sum_{i=1}^N p_i^0 q_i^0}, \tag{7}$$

$$RP^P = \frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^0 q_i^1};\ RQ^P = \frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^1 q_i^0}, \tag{8}$$

$$RP^F = (RP^L RP^P)^{\frac{1}{2}} = \left(\frac{\sum_{i=1}^N p_i^1 q_i^0}{\sum_{i=1}^N p_i^0 q_i^0}\right)^{\frac{1}{2}} \left(\frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^0 q_i^1}\right)^{\frac{1}{2}}$$

$$RQ^F = (RQ^L RQ^P)^{\frac{1}{2}} = \left(\frac{\sum_{i=1}^N p_i^0 q_i^1}{\sum_{i=1}^N p_i^0 q_i^0}\right)^{\frac{1}{2}} \left(\frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^1 q_i^0}\right)^{\frac{1}{2}}. \tag{9}$$

Hence, Fisher's price (quantity) index is the geometric mean of the price (quantity) indices of Laspeyres and Paasche. All above-mentioned indices satisfy the basic properties, so that they also satisfy the stronger property of 'proportionality' (Balk, 2008a; 2016). It is easily verified from (7) and (8) that the indices of Laspeyres and Paasche do ***not*** possess the property of time reversal. However, the indices of Fisher do possess it. It is also easily verified that:

$$RV = RP^L \times RQ^P; RV = RP^P \times RQ^L \text{ and } RV = RP^F \times RQ^F. \tag{10}$$

It follows from (10) that the index pair of Fisher is ideal, but that the pairs of Laspeyres and Paasche are not ideal.

2.3.3. Indicators

The Laspeyres and Paasche price and quantity indicators are the additive counterparts of the Laspeyres and Paasche price and quantity indices and defined as:

$$DP^L = \sum_{i=1}^N q_i^0(p_i^1 - p_i^0); DQ^L = \sum_{i=1}^N p_i^0(q_i^1 - q_i^0),$$
$$DP^P = \sum_{i=1}^N q_i^1(p_i^1 - p_i^0); DQ^P = \sum_{i=1}^N p_i^1(q_i^1 - q_i^0).$$

The additive counterpart of the Fisher indices are the indicators of Bennet (1920) defined as the arithmetic mean of the Laspeyres and Paasche price indicators:

$$DP^B = \frac{1}{2}\sum_{i=1}^N q_i^0(p_i^1 - p_i^0) + \frac{1}{2}\sum_{i=1}^N q_i^1(p_i^1 - p_i^0) = \sum_{i=1}^N \frac{(q_i^0 + q_i^1)}{2}(p_i^1 - p_i^0), \tag{11}$$
$$DQ^B = \frac{1}{2}\sum_{i=1}^N p_i^0(q_i^1 - q_i^0) + \frac{1}{2}\sum_{i=1}^N p_i^1(q_i^1 - q_i^0) = \sum_{i=1}^N \frac{(p_i^0 + p_i^1)}{2}(q_i^1 - q_i^0). \tag{12}$$

All above-mentioned indicators satisfy the basic properties (Balk, 2008a; 2016). It is easily verified that the indicators of Laspeyres and Paasche do ***not*** possess the property of time reversal. But Bennet's indicators do possess it. It is also easily verified that

$$DV = DP^L + DQ^P; DV = DP^P + DQ^L \text{ and } DV = DP^B + DQ^B. \tag{13}$$

It follows from (13) that Bennet's indicator pair is ideal, but Laspeyres's and Paasche's do not possess that property.

2.3.4. Fisher and Bennet: a combinatorial approach

De Boer (2009b) used the decomposition of the ratio change in value into the ratio changes in two factors (price and quantity) by means of the Fisher indices as a step towards the general case of n factors where he used the generalization of Siegel of the Fisher indices. Let V denote a certain aggregate to be decomposed and let x_1 and x_2 denote factors 1 and 2, respectively (in our example 'price' and 'quantity'), then we have:

$$V = \sum_{i=1}^{N} x_{1i}x_{2i}. \tag{14}$$

The Fisher index for factor 1 reads:

$$RX_1{}^F = \left(\frac{\sum_{i=1}^{N} x_{1i}^1 x_{2i}^0}{\sum_{i=1}^{N} x_{1i}^0 x_{2i}^0}\right)^{\frac{1}{2}} \left(\frac{\sum_{i=1}^{N} x_{1i}^1 x_{2i}^1}{\sum_{i=1}^{N} x_{1i}^0 x_{2i}^1}\right)^{\frac{1}{2}}. \tag{15}$$

The first term gives the change in factor 1, weighted by the magnitudes of factor 2 in the base period, and the second term the change in factor 1, weighted by the values of factor 2 in the comparison period. The number of duplicates of each term is 1 and the weight of each term is ½. This is summarized by De Boer (2009b) in his Table 1.

The Bennet indicator for factor 1 reads:

$$DX_1^B = \frac{1}{2}\sum_{i=1}^{N}(x_{1i}^1 - x_{1i}^0)x_{2i}^0 + \frac{1}{2}\sum_{i=1}^{N}(x_{1i}^1 - x_{1i}^0)x_{2i}^1. \tag{16}$$

The first term gives the change in factor 1 (prices), weighted by the values of the factor 2 (quantities) in the base period, and the second term the change in factor 1 (prices), weighted by the values of the factor 2 (quantities) in the comparison period. The number of duplicates of each term is 1 and the weight of each term is ½. Consequently, Table 1 can also be used for the Bennet indicators. De Boer (2009b) also supplied the tables with 3–6 factors; the one for 5 factors is used in our empirical application.

2.4. Logarithmic indices and indicators

2.4.1. The logarithmic mean

The indices and indicators that belong to this class are based on the logarithmic mean, which for two *positive* numbers a and b is defined as:

$$L(a,b) = \frac{a-b}{\ln\left(\frac{a}{b}\right)} \text{ and } L(a,a) = a. \tag{17}$$

The logarithmic mean is very convenient when switching from a ratio to a difference and vice versa (Balk, 2003). It follows straightforwardly from (17) that:

$$a/b = \exp\{(a-b)/L(a,b)\}, \tag{18}$$

Table 1. Summary for the case of two factors.

Number of ones	Combinations	Number of duplicates	Weight
0	{0}	1	1/2
1	{1}	1	1/2

$$(a-b) = L(a,b)\ln(a/b). \tag{19}$$

It is '*zero value robust*': in practice we can replace zeros by epsilon small positive numbers. (Ang and Liu, 2007a). If a and b are both positive, it can still be used. However, if there is a change in sign, that is to say when a is positive (negative) and b is negative (positive), the logarithmic mean (17) is not defined so that it is ***not*** '*change-in-sign robust*'. According to Ang and Liu (2007b) the logarithmic mean might handle changes in sign using the so-called 'Analytical Limit Strategy'. Their procedure has to be applied to each change in sign individually. In practice, this is so cumbersome that in case of the presence of changes in sign we do not advise to use methods based on the logarithmic mean.

2.4.2. The Montgomery indicator and Montgomery-Vartia index

Balk (2003) gave a simple derivation of the indicator of Montgomery (1929; 1937) that was used in the application of this indicator to SDA by de Boer (2008). Using, successively, definition (4) of the difference change in value, definition (17) of the logarithmic mean of the value in the comparison (v_i^1) and base period (v_i^0), and the definition (1) of v_i^1 and v_i^0 we obtain:

$$\begin{aligned} DV^M &= \sum_{i=1}^{N}(v_i^1 - v_i^0) = \sum_{i=1}^{N} L(v_i^1, v_i^0)\ln\left(\frac{v_i^1}{v_i^0}\right), \\ &= \sum_{i=1}^{N} L(v_i^1, v_i^0)ln\left(\frac{p_i^1}{p_i^0}\right) + \sum_{i=1}^{N} L(v_i^1, v_i^0)\ln\left(\frac{q_i^1}{q_i^0}\right), \end{aligned} \tag{20}$$

in which the first term after the second equality is the definition of the price indicator and the second term of the quantity indicator according to Montgomery.

Using the following definition of the weight for the Montgomery decomposition:

$$w_i^M = L(v_i^1, v_i^0), \tag{21}$$

the price and quantity indicators read:

$$DP^M = \sum_{i=1}^{N} w_i^M \ln\left(\frac{p_i^1}{p_i^0}\right) \text{ and } DQ^M = \sum_{i=1}^{N} w_i^M \ln\left(\frac{q_i^1}{q_i^0}\right). \tag{22}$$

These indicators exhibit the basic and desired properties of time reversal and being ideal, except the property of monotonicity. But, as argued by Balk (2003, Appendix A.2), this problem is unlikely to be of practical importance.

Using the definition of the logarithmic mean of the value in the comparison period (V^1) and the base period (V^0) and rewriting (19) we obtain:

$$\ln\left(\frac{V^1}{V^0}\right) = \frac{DV}{L(V^1, V^0)}. \tag{23}$$

We define the weight according to Montgomery-Vartia decomposition as:

$$w_i^{MV} = \frac{w_i^M}{L(V^1, V^0)}. \tag{24}$$

Then, the application of (23) to the Montgomery decomposition (20) leads to the Montgomery-Vartia decomposition:

$$\ln(RV^{MV}) = \frac{DP^M}{L(V^1, V^0)} = \sum_{i=1}^{N} w_i^{MV} \ln\left(\frac{p_i^1}{p_i^0}\right) + \sum_{i=1}^{N} w_i^{MV} \ln\left(\frac{q_i^1}{q_i^0}\right). \tag{25}$$

The first term after the second equality of (25) is the logarithm of the price index and the second one the logarithm of the quantity index of Montgomery-Vartia. By exponentiation it follows from (25) that:

$$RP^{MV} = \prod_{i=1}^{N} \left(\frac{p_i^1}{p_i^0}\right)^{w_i^{MV}} \text{ and } RQ^{MV} = \prod_{i=1}^{N} \left(\frac{q_i^1}{q_i^0}\right)^{w_i^{MV}}. \tag{26}$$

Like the Montgomery indicators, the indices of Montgomery-Vartia do not exhibit the property of global monotonicity, but as argued by Balk (2003, Appendix A.1) this problem is unlikely to be of practical importance. More importantly, contrarily to the Fisher indices, the indices of Montgomery-Vartia fail to exhibit the property of 'linear homogeneity in comparison prices (quantities)' and, consequently the property of 'proportionality'. Fulfilment requires the sum of the weights (24) being equal to one, but, using Jensen's inequality, Balk (2003) proved that this sum is smaller than one.

2.4.3. The Sato-Vartia index and the Additive Sato-Vartia indicator

Balk (2003) supplied a simple derivation of the indices of Sato-Vartia that was used in the application of this index to SDA by de Boer (2009a). From the logarithmic mean of the value shares of commodity i in total expenditure in comparison and base period, i.c. s_i^1 and s_i^0, we derive $L(s_i^1, s_i^0)\ln(s_i^1/s_i^0) = s_i^1 - s_i^0$. Summation over i results in:

$$\sum_{i=1}^{N} L(s_i^1, s_i^0)\ln(s_i^1/s_i^0) = \sum_{i=1}^{N} (s_i^1 - s_i^0) = 0, \tag{27}$$

where the last equality follows from the adding-up of the shares to one.

From (1) and (2) it follows that $s_i^1 = p_i^1 q_i^1/V^1$ and $s_i^0 = p_i^0 q_i^0/V^0$. Consequently,

$$\ln(s_i^1/s_i^0) = \ln(p_i^1/p_i^0) + \ln(q_i^1/q_i^0) - \ln(V^1/V^0). \tag{28}$$

Substitution of (28) into (27) leads to:

$$\ln(V^1/V^0) \sum_{i=1}^{N} L(s_i^1, s_i^0) = \sum_{i=1}^{N} L(s_i^1, s_i^0)\ln(p_i^1/p_i^0) + \sum_{i=1}^{N} L(s_i^1, s_i^0)\ln(q_i^1/q_i^0),$$

which after defining:

$$w_i^{SV} = \frac{L(s_i^1, s_i^0)}{\sum_{i=1}^{N} L(s_i^1, s_i^0)}, \tag{29}$$

can be rewritten to:

$$\ln(V^1/V^0) = \sum_{i=1}^{N} w_i^{SV} \ln(p_i^1/p_i^0) + \sum_{i=1}^{N} w_i^{SV} \ln(q_i^1/q_i^0). \tag{30}$$

The first term on the right-hand side of (30) is the logarithm of the price index and the second one the logarithm of the quantity index of Sato-Vartia. By exponentiation it follows from (30) that:

$$RP^{SV} = \prod_{i=1}^{N} \left(\frac{p_i^1}{p_i^0} \right)^{w_i^{SV}} ; RQ^{SV} = \prod_{i=1}^{N} \left(\frac{q_i^1}{q_i^0} \right)^{w_i^{SV}} . \tag{31}$$

Like the indices of Montgomery-Vartia, the indices according to Sato-Vartia do not exhibit the property of global monotonicity, but, as argued by Balk (2003, Appendix A.3), this problem is unlikely to be of practical importance either. More importantly, contrary to Montgomery-Vartia indices, due to the fact that the sum of the weights (29) is equal to one, the indices of Sato-Vartia exhibit the property of 'proportionality'.

The additive counterparts of the Sato-Vartia indices do not exist in index number theory. Therefore we call them 'Additive Sato-Vartia'. It was introduced in energy and environmental studies by Ang et al. (2003, Appendix B) under the name 'Additive LMDI II'. We rewrite (23) to:

$$DV = L(V^1, V^0)\ln \left(\frac{V^1}{V^0} \right). \tag{32}$$

Substitution of (30) into (32) and defining the weight of the Additive Sato-Vartia decomposition as:

$$w_i^{ASV} = L(V^1, V^0) w_i^{SV}, \tag{33}$$

we arrive at:

$$DV = \sum_{i=1}^{N} w_i^{ASV} \ln \left(\frac{p_i^1}{p_i^0} \right) + \sum_{i=1}^{N} w_i^{ASV} \ln \left(\frac{q_i^1}{q_i^0} \right),$$

so that the price indicators according to Additive Sato-Vartia are:

$$DP^{ASV} = \sum_{i=1}^{N} w_i^{ASV} \ln \left(\frac{p_i^1}{p_i^0} \right) \text{ and } DQ^{ASV} = \sum_{i=1}^{N} w_i^{ASV} \ln \left(\frac{q_i^1}{q_i^0} \right).$$

Like the other logarithmic indices and indicators, it will not exhibit the property of global monotonicity, either. But this problem is unlikely to be of much practical importance, as well.

2.5. A contribution of index number theory: consistency-in-aggregation

All indices (indicators) that we discussed above are one-stage indices (indicators). In practice, the computation of price and quantity indices might also be performed via a multistage process. As an example, Balk (2016) gives the computation of the price index of Laspeyres in two stages. The set of commodities A is partitioned into K disjoint subsets $A^k, (k = 1, \dots, K)$, that is to say: $A = \cup_{k=1}^{K} A^k$ with $A^k \cap A^l = \emptyset$ for $k \neq l$. In the first stage, the Laspeyres price index is computed for each subset A^k. In the second stage the Laspeyres price index of all first stage indices is computed. Balk (2016) proves that the Laspeyres index according to (7) is equal to Laspeyres index computed in the two-stage process. If at each stage the same type of index is used the index called 'consistent-in aggregation' (CIA). Consequently, the indices of Laspeyres are CIA. As Balk (2016, p. 7) states: 'However, this is the exception rather than the rule. For most indices, two-stage and one-stage variants do *not* coincide. Put otherwise, most indices are not consistent-in-aggregation.'

Balk (1996) formalized consistency-in-aggregation of a particular price index and proved that the 'pseudo Montgomery' price index (nowadays named 'Montgomery-Vartia') is CIA. Balk (2008a) reproduced this canonical form and the proof[1] in his section 3.7.2. The proofs that the indicators of Bennet and Montgomery[2] are CIA are given in his section 3.10.3. In Appendix A of the supplementary material, we give a numerical example from which it is immediately clear that the one-step indices according to Fisher and Sato-Vartia are not equal to the corresponding two-step indices. The same applies to the one- and two-step indicators according to the Additive Sato-Vartia indicator. As a consequence, 'Fisher', 'Sato- Vartia' and 'Additive Sato-Vartia' are ***not*** CIA.

2.6. Summary of relevant properties and formulas for IDA and SDA

2.6.1. Relevant properties

All pairs of indices and indicators share the properties of 'identity', 'linear homogeneity of degree zero in prices (quantities)', invariance to changes in the units of measurement', 'time reversal' and being 'ideal'. In Table 2 we summarize those theoretical properties, which are met with or not.

Table 2. Fulfilment of relevant properties of ideal indices and indicators.

	Fisher	Montgomery-Vartia	Sato-Vartia	Bennet	Montgomery	ASV
Monotonicity	Yes	No	No	Yes	No	No
Proportionality	Yes	No	Yes	N.A.	N.A.	N.A.
Consistency-in-aggregation	No	Yes	No	Yes	Yes	No

It follows from Table 2 that no ideal index meets all the three desired properties simultaneously. Fisher's is the only one that meets the property of monotonicity; Montgomery-Vartia is the only one that meets consistency-in-aggregation; whereas Fisher and Sato-Vartia both meet the requirement of proportionality, but not Montgomery-Vartia. With respect to the ideal indicators Additive Sato-Vartia is the only one that does not meet the requirement of consistency-in-aggregation so that it not advisable to use it in practice.

2.6.2. Summary of formulas

The relevant formulas of index number theory are summarized in Table 3. From the empirical point of view they are as easily implemented. But remember that index number theory deals with only two factors. In Section 4, we present the generalization to n factors. We find that methods based on the logarithmic mean are easier to implement than the combinatorial ones (Fisher and Bennet). But knowing that these two factors only take on nonnegative

[1] Without using the canonical form of Balk (1996, 2008a), Ang and Liu (2001) gave a direct proof that Montgomery-Vartia is CIA.

[2] Without using the canonical form of Balk (1996, 2008a) Ang and Wang (2015) give a direct proof that Montgomery is CIA.

Table 3. Decomposition of change in value into price and quantity effect of ideal methods.

Name	Weight	Price effect	Quantity effect
Fisher index	N.A.	$\left(\frac{\sum_{i=1}^N p_i^1 q_i^0}{\sum_{i=1}^N p_i^0 q_i^0}\right)^{\frac{1}{2}} \left(\frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^0 q_i^1}\right)^{\frac{1}{2}}$	$\left(\frac{\sum_{i=1}^N p_i^0 q_i^1}{\sum_{i=1}^N p_i^0 q_i^0}\right)^{\frac{1}{2}} \left(\frac{\sum_{i=1}^N p_i^1 q_i^1}{\sum_{i=1}^N p_i^1 q_i^0}\right)^{\frac{1}{2}}$
Bennet indicator	N.A.	$\sum_{i=1}^N \frac{(q_i^0 + q_i^1)}{2}(p_i^1 - p_i^0)$	$\sum_{i=1}^N \frac{(p_i^0 + p_i^1)}{2}(q_i^1 - q_i^0)$
Montgomery indicator	$w_i^M = L(v_i^1, v_i^0)$	$\sum_{i=1}^N w_i^M \ln\left(\frac{p_i^1}{p_i^0}\right)$	$\sum_{i=1}^N w_i^M \ln\left(\frac{q_i^1}{q_i^0}\right)$
Montgomery-Vartia index	$w_i^{MV} = \frac{w_i^M}{L(V^1, V^0)}$	$\prod_{i=1}^N \left(\frac{p_i^1}{p_i^0}\right)^{w_i^{MV}}$	$\prod_{i=1}^N \left(\frac{q_i^1}{q_i^0}\right)^{w_i^{MV}}$
Sato-Vartia index	$w_i^{SV} = \frac{L(s_i^1, s_i^0)}{\sum_{i=1}^N L(s_i^1, s_i^0)}$	$\prod_{i=1}^N \left(\frac{p_i^1}{p_i^0}\right)^{w_i^{SV}}$	$\prod_{i=1}^N \left(\frac{q_i^1}{q_i^0}\right)^{w_i^{SV}}$
Additive Sato-Vartia indicator	$w_i^{ASV} = L(V^1, V^0) w_i^{SV}$	$\sum_{i=1}^N w_i^{ASV} \ln\left(\frac{p_i^1}{p_i^0}\right)$	$\sum_{i=1}^N w_i^{ASV} \ln\left(\frac{q_i^1}{q_i^0}\right)$

numbers means that changes-in-sign cannot occur. In the practice of SDA, however, they might.[3]

3. Correspondence between index number theory and IDA and SDA: overview of literature

3.1. Combinatorial indices

3.1.1. Index decomposition analysis

Sun (1998) derived a complete additive decomposition model for n factors by a refinement of the Laspeyres's method. In it he assures that the residuals due to interactions are distributed equally among the main effects based on the principle of 'jointly created and equally distributed principle' (Sun, 1998, p. 88, citing Sun, 1996). Albrecht et al. (2002) used the Shapley value from noncooperative game theory to derive a complete additive decomposition model. Ang et al. (2003) proved that Sun's approach is equivalent to using the Shapley value and, hence, named the method 'Sun-Shapley'. But it was just the Bennet indicator applied to n factors. Ang et al. (2004) applied the multiplicative analogue of the Shapley value to the four-factor decomposition model of Chung and Rhee (2001). In their Appendix A they presented the decomposition formula using Shapley's (1953) generic formula. They called it the 'Generalized Fisher' method. Siegel (1945) had already supplied the resulting decomposition formula in his generalization of the two-factor Fisher index.

3.1.2. Structural decomposition analysis

Dietzenbacher and Los (1998) decomposed *additively* the change in labor costs of 214 sectors in the Netherlands between 1986 and 1992 into four components: the effects of

[3] To give but one example: in the seven-sector decomposition model of Chung and Rhee (2001) the sector 'petroleum, coal and town gas' has a negative value for final demand in the base period, but a positive one in the comparison period. Methods based on the logarithmic mean cannot be used, but the combinatory ones are applicable.

a change in the labor cost per unit of output; the effects of technical change; the effects of changes in the final demand mix; and the effects of the changes in the final demand levels. They computed the *arithmetic* mean of all $n! = 24$ elementary decompositions and of the two polar decompositions, and concluded that both means were rather close to each other. In the framework of the same example Dietzenbacher et al. (2000) gave a *multiplicative* decomposition and computed the *geometric* mean of all 24 elementary decompositions and of the two polar decompositions and reached the same conclusion as before. De Haan (2001) collected the duplicates of the additive decomposition of Dietzenbacher and Los and gave, in his Table 1, the weights attached to each of the eight combinations. These weights are equal to those given in Siegel (1945) for the multiplicative analogue. The formula used by de Haan is nothing but the Bennet indicator. Seibel (2003) extended De Haan's 'Dutch approach' and illustrated it for the five-factor case. But his approach is the same as Siegel's, which we show via application in Appendix B. In Appendix C we apply Shapley's approach and show the equivalence of the two approaches. De Boer (2009b) proved that the geometric mean of all elementary decompositions is equivalent to Siegel's generalization of the index of Fisher to n factors. Since Bennet's indicator is the additive counterpart of Fisher's index (Balk, 2003), De Boer's proof implies that the arithmetic mean of all elementary decompositions is equivalent to the Siegel's (1945) generalization of Bennet's indicator.

To summarize: all the approaches described above are either generalizations of the Fisher index or of the Bennet indicator applied to a decomposition into n factors. The generic formulae used are either Siegel's or Shapley's.

3.2. Logarithmic indices and indicators

3.2.1. Index decomposition analysis

Boyd et al (1987) introduced their so-called 'Divisia index approach', by assuming all variables are continuous and each is a function of time. The resulting equation is differentiated with respect to t, integrated over the time interval 0 to T, and the integral path is approximated using the arithmetic mean weight function. It resulted in the so-called AMDI method ('Artihmetic Mean Divisia Index'.) In the theory of indices and indicators, this method is called the 'Törnqvist index' (Törnqvist and Törnqvist, 1937), which is defined as the geometric mean of the Geometric Laspeyres and Geometric Paasche indices (Balk, 2008a). Boyd and Roop (2004) explicitly introduced the Fisher ideal index for the decomposition of structural change in energy intensity into two factors ('structure' and 'intensity') and compared the results with those obtained by application of AMDI. Since the latter is not an ideal index we do not consider it in this paper. Ang and Choi (1997) introduced 'a refined Divisia index method' by replacing the arithmetic mean by the logarithmic mean weight scheme proposed by Sato (1976). Ang and Liu (2001) renamed it the LMDI-II method but is clearly equivalent to the Sato-Vartia index. Ang et al. (1998) proposed 'a refined Divisia index method based on decomposition of a differential quantity'. This method is nothing but Montgomery's indicator. Ang and Liu (2001) proposed a so-called LMDI-I method, which is equivalent to the Montgomery-Vartia index. They also rename the method proposed by (Ang et al. 1998) to (additive) LMDI-I. The mathematical derivations of the multiplicative LMDI-I and LMDI-II methods and of the additive LMDI-I method (which are, respectively, the Montgomery-Vartia and Sato-Vartia indices, and the Montgomery indicator) are mathematically complicated. As shown in Section 2, Balk (2003) supplied far simpler derivations. The additive LMDI-II method is introduced in Ang et al. (2003, Appendix B). As noted earlier, this method is unknown in the theory of indices and indicators.

3.2.2. Structural decomposition analysis

De Boer (2008) showed the correspondence between the theory of indices and indicators and SDA. He applied the Montgomery indicator to the additive decomposition of the example of Dietzenbacher and Los (1998), replicated their results and showed that the Montgomery decompositions were very close to the arithmetic mean of all elementary decompositions. De Boer (2009a) applied the index of Sato-Vartia to the multiplicative decomposition in the framework of the same example (Dietzenbacher et al., 2000). He replicated their results and showed that the Sato-Vartia decompositions were very close to the geometric means of all elementary decompositions. Table 4 reports the equivalence between names of methods in index number theory and in index/structural decomposition analysis.

3.3. Summary of names of methods

Table 4. Summary of names of methods.

Ratio change		Difference change	
Index	Multiplicative decomposition	Indicator	Additive decomposition
Fisher	Dietzenbacher- Hoen-Los or Generalized Fisher	Bennet	Dietzenbacher- Los or Sun-Shapley
Montgomery-Vartia	LMDI- I	Montgomery	LMDI- I
Sato-Vartia	LMDI- II	Additive Sato-Vartia	LMDI- II

4. Ideal methods for decomposition: the case of *n* factors

4.1. Introduction

In Section 2, we decomposed the aggregate change in a variable V, i.e. value, into two factors: price and quantity. The decomposition took on two different forms: a ratio change $RV = V^1/V^0$ (*multiplicative decomposition*), or a *difference change* $DV = V^1 - V^0$ (*additive decomposition*.) We considered pairs of price and quantity indices (indicators) that are *ideal*, i.e, the decomposition is complete or, in other words, there is no residual term.

In this section, we deal with the decomposition of an aggregate change into n factors which can be written as[4]:

$$V = \sum_{\{I\}} \prod_{f=1}^{n} x_f, \tag{34}$$

where V: aggregate to be decomposed; $\{I\}$: set of summation indices; x_f: factor $f = 1, \ldots, n$.

The *multiplicative decomposition* reads:

[4] In applied research there are often two-step decompositions, such as the decomposition of the Leontief matrix (Xu and Dietzenbacher, 2014; Zhang and Lahr, 2014) and multiplicative attribution analysis (Choi and Ang, 2012; Su and Ang, 2014; Su and Ang, 2015; Wang et al. 2017a; Yan et al., 2018). In all these cases (34) is the first step. Such two-step decompositions are beyond the scope of the present paper.

$RV = \prod_{f=1}^{n} R_{x_f}$ with R_{x_f}: ratio change of factor f
and the additive decomposition:
$DV = \sum_{f=1}^{n} D_{x_f}$ with D_{x_f}: difference change of factor f

4.2. Fisher and Bennet

4.2.1. The generic formula of Siegel (1945)

For the multiplicative decomposition, Siegel (1945) reduced, by collecting duplicates, the calculation of $n!$ permutations (in SDA called 'elementary decompositions') to the calculation of 2^{n-1} combinations. Then, he proposed to take the *weighted geometric mean* of all combinations, the fraction of the number of duplicates in the total number of elementary decompositions being the exponent. This is, of course, equal to the geometric mean of all elementary decompositions.

The generic formula for the geometric mean is given in Appendix B, in which it is applied to the case of five factors. It amounts to the calculation of $2^4 = 16$ combinations, whereas the number of elementary decompositions is equal to $5! = 120$, which constitutes a considerable decrease in the number of computations. The Bennet decomposition is the additive counterpart to the Fisher decomposition (Balk, 2003). It is the *weighted arithmetic mean* with the same combinations, the weights being the same as the exponents of the Fisher decomposition.

4.2.2. The generic formula of Shapley (1953)

Although independent of Siegel,[5] Shapley (1953) followed an identical route for the additive decomposition. He reduced permutations to combinations and proposed taking the weighted *arithmetic* mean of all combinations, the divisor being again the fraction of the number of duplicates in the total number of permutations. In Appendix C we give the generic formula of Shapley and apply it to the case of five factors, as well. We present a table in which we prove that Siegel and Shapley yield exactly the result.

4.3. Logarithmic indices and indicators

Consider Equation 34 and define:

$$v = \prod_{f=1}^{n} x_f, \tag{35}$$

and

$$s = \frac{\prod_{f=1}^{n} x_f}{V} = \frac{v}{V}. \tag{36}$$

The four methods are summarized in Table 5 (compare Table 3).

It follows from Table 5 that all methods are a weighted mean of the logarithm of the relatives, i.e. the value of a factor in the comparison period relative to its value in

[5] In IDA and SDA literature the generic formula of Siegel has not as yet been presented. De Boer (2009b) used Siegel's formula but only gave a verbal description. Su and Ang (2014, Appendix B) use Shapley's generic formula but erroneously attribute it to Siegel.

the base period; the only difference being the weighting factor. It is easy to verify that the Montgomery-Vartia decomposition can be derived from Montgomery's using the transformation (cf. (18)):

$$R_{x_f}^{MV} = exp[D_{x_f}^{M}/L(V^1, V^0)]\ (f = 1, \ldots, n), \tag{37}$$

and that the Additive Sato-Vartia decomposition can be derived from Sato-Vartia's using the transformation (compare (19)):

$$D_{x_f}^{ASV} = L(V^1, V^0)\ln[R_{x_f}^{SV}]\ (f = 1, \ldots, n). \tag{38}$$

Equations 37 and 38 are used in the Matlab program given in Appendix E.

Table 5. Name of the method, the weight, the effect of a factor and the decomposition.

Name	Weight	Effect of factor	Decomposition
Montgomery (LMDI-I additive)	$w^M = L(v^1, v^0)$	$D_{x_f}^M = \sum_{\{I\}} w^M \ln\left(\frac{x_f^1}{x_f^0}\right)$	$\sum_{f=1}^n D_{x_f}^M$
Montgomery- Vartia (LMDI-I multiplicative)	$w^{MV} = \frac{w^M}{L(V^1, V^0)}$	$\ln(R_{x_f}^{MV}) = \sum_{\{I\}} w^{MV} \ln\left(\frac{x_f^1}{x_f^0}\right)$	$\prod_{f=1}^n R_{x_f}^{MV}$
Sato-Vartia (LMDI-II multiplicative)	$w^{SV} = \frac{L(s^1, s^0)}{\sum_{\{I\}} L(s^1, s^0)}$	$\ln(R_{x_f}^{SV}) = \sum_{\{I\}} w^{SV} \ln\left(\frac{x_f^1}{x_f^0}\right)$	$\prod_{f=1}^n R_{x_f}^{SV}$
Additive Sato-Vartia (LMDI-II additive)	$w^{ASV} = L(V^1, V^0)w^{SV}$	$D_{x_f}^{ASV} = \sum_{\{I\}} w^{ASV} \ln\left(\frac{x_f^1}{x_f^0}\right)$	$\sum_{f=1}^n D_{x_f}^{ASV}$

5. A hands-on toy model of decomposition of an aggregate change into five factors

5.1. The toy model and its decompositions

In our expository toy model we only deal with emissions of carbon dioxide of sectors of the Dutch economy and ignore the direct emissions of households. We dispose of two input–output tables and of the sectoral carbon dioxide emissions. The number of sectors is denoted by N and the number of final demand categories by m.

We define the following vectors and matrices:

$co2$: $N \times 1$ vector of sectoral emissions of carbon dioxide;

x: $N \times 1$ vector of sectoral outputs;

e: $N \times 1$ vector of sectoral emissions per unit of output;

$\hat{e}$: $N \times N$ diagonal diagonal matrix with e on the main diagonal;

A: $N \times N$ matrix of input–output coefficients a_{ij} measuring the input from sector i in sector j, per unit of sector j's output;

B: $N \times m$ matrix of bridge coefficients b_{jk} measuring the fraction of final demand in category k that is spent on products from sector j;

u: $m \times 1$ vector of shares u_k of final demand category k in total final demand; and

y: total final demand.

We consider the model:

$$co2 = \hat{e}x,$$

$$x = Ax + Buy,$$

of which the solution is:

$$co2 = \hat{e}DBuy, \tag{39}$$

with: $D = (I - A)^{-1}$ the Leontief inverse.

In sum notation (39) reads:

$$co2_i = \sum_{j=1}^{N} \sum_{k=1}^{m} e_i d_{ij} b_{jk} u_k y. \tag{40}$$

Consequently, the aggregate to be decomposed, V in (34), is 'carbon dioxide emissions of sector i' ($co2_i$), the set of summation indices is $\{I\} = j, k$; and the factors are: x_1 ('emission coefficients', e_i); x_2 ('production techniques', d_{ij}); x_3 ('final demand mix', b_{jk});x_4 ('demand structure', u_k); and x_5 ('size of economy', y), respectively. We want to decompose the change in carbon dioxide emissions from the base period, denoted by the superscript 0, to the comparison period, denoted by the superscript 1, into the changes of these five factors.

5.1.1. *Multiplicative (ratio) decomposition*

The ratio change in carbon dioxide emissions of sector i is defined to be:

$$RCO2_i = \mathrm{co2}_\mathrm{i}^1 / co2_i^0.$$

From (40) we obtain:

$$RCO2_i = \frac{\sum_{j=1}^{N} \sum_{k=1}^{m} e_i^1 d_{ij}^1 b_{jk}^1 u_k^1 y^1}{\sum_{j=1}^{N} \sum_{k=1}^{m} e_i^0 d_{ij}^0 b_{jk}^0 u_k^0 y^0}. \tag{41}$$

We want to decompose (41) into the ratio changes in emission coefficients (RE_i), production techniques (RD_i), final demand mix (RB_i), demand structure (RU_i) and size of the economy (RY_i), i.e.:

$$RCO2_i = RE_i \times RD_i \times RB_i \times RU_i \times RY_i.$$

5.1.2. *Additive (difference) decomposition*

The difference change in carbon dioxide emissions of sector i is defined to be:

$$DCO2_i = co2_i^1 - co2_i^0.$$

From (40) we obtain:

$$DCO2_i = \sum_{j=1}^{N} \sum_{k=1}^{m} (e_i^1 d_{ij}^1 b_{jk}^1 u_k^1 y^1 - e_i^0 d_{ij}^0 b_{jk}^0 u_k^0 y^0). \tag{42}$$

We want to decompose (42) into the difference changes in emission coefficients (DE_i), production techniques (DD_i), final demand mix (DB_i), demand structure (DU_i), and size of the economy (DY_i), i.e.:

$$DCO2_i = DE_i + DD_i + DB_i + DU_i + DY_i$$

Table 6. Summary for the case of five factors.

Appendix A Equation	Number of ones	Combinations				Number of duplicates	Weight
(A.1)	0	**{0,0,0,0}**				24	1/5
(A.2)	1	**{1,0,0,0}**	{0,1,0,0}	{0,0,1,0}	**{0,0,0,1}**	6	1/20
(A.3)	2	**{1,1,0,0}**	{1,0,1,0}	{1,0,0,1}		4	1/30
(A.4)	2	{0,0,1,1}	{0,1,0,1}	**{0,1,1,0}**		4	1/30
(A.5)	3	**{0,1,1,1}**	{1,0,1,1}	{1,1,0,1}	**{1,1,1,0}**	6	1/20
(A.6)	4	**{1,1,1,1}**				24	1/5

5.1.3. One- and two-step decomposition

Due to the presence of two common factors, the multiplicative decomposition (41) can easily be rewritten to a two-step procedure (de Boer, 2009b):

$$RCO2_i = \frac{e_i^1}{e_i^0}\frac{y^1}{y^0}\frac{\sum_{j=1}^{N}\sum_{k=1}^{m} d_{ij}^1 b_{jk}^1 u_k^1}{\sum_{j=1}^{N}\sum_{k=1}^{m} d_{ij}^0 b_{jk}^0 u_k^0}. \tag{43}$$

In the first step we compute the simple index numbers of the factors 'emission coefficients' and 'size of the economy' and in the second step the composite index numbers of the factors 'production techniques', 'final demand mix' and 'demand structure'. The decompositions according to Fisher and Sato-Vartia possess the property of proportionality. Then, the one-step decomposition (41) and the two-step (43) yield the same results for the simple index numbers of the factors 'emission coefficients' and 'size of the economy'. They are used as a assure that the Matlab programs performing the computation of multiplicative and additive decomposition at the same time are correct. The decomposition according to Montgomery-Vartia does not exhibit proportionality, so that the one- and two-step decompositions do not yield the same results. Unfortunately, such a simple device of two-step decomposition does not exist for additive decompositions. Evidently, Equation 42 cannot be rewritten to the sum of the two simple indicators for 'emission coefficients' and 'size of the economy' and the composite indicators of 'production techniques', 'final demand mix' and 'demand structure'. Moreover, there is no analogue of 'proportionality' for indicators.

5.2. The Fisher and Bennet decompositions

In Appendix B we use Siegel's generic formula, and in Appendix C Shapley's generic formula to arrive at the following summarizing table. In a different format it is also given in De Boer (2009b).

In the first row of Table 6, we give the combination with the values that the other four other factors take on in the base period. In SDA literature this is called the *Laspeyres perspective*. There are 24 duplicates so that on the 120 elementary decompositions the *first polar decomposition* has a weight of 1 over 5. In the last row we give the combination with the values of the four other factors in the comparison period, the *Paasche perspective*. The weight of the *second polar decomposition* is also 1 over 5. If the *mean of the two polar decompositions* is taken the combinations given in the rows (A.2) up to and including (A.5) are neglected and the weights are increased to 1 over 2. For a small number of factors you may expect it to be close to the mean of all decompositions.

From Table 6 we can derive the decomposition formulae for each of the five factors. For factor 1[6], 'ratio change in emission coefficients', it results in:

$$\begin{aligned}
R_{x_1}^F = &\left[\frac{\sum x_1^1x_2^0x_3^0x_4^0x_5^0}{\sum x_1^0x_2^0x_3^0x_4^0x_5^0}\right]^{1/5} \times \left[\frac{\sum x_1^1x_2^1x_3^0x_4^0x_5^0}{\sum x_1^0x_2^1x_3^0x_4^0x_5^0}\right]^{1/20} \times \ldots \times \left[\frac{\sum x_1^1x_2^0x_3^0x_4^0x_5^1}{\sum x_1^0x_2^0x_3^0x_4^0x_5^1}\right]^{1/20} \\
&\times \left[\frac{\sum x_1^1x_2^1x_3^1x_4^0x_5^0}{\sum x_1^0x_2^1x_3^1x_4^0x_5^0}\right]^{1/30} \times \ldots \times \left[\frac{\sum x_1^1x_2^0x_3^1x_4^1x_5^0}{\sum x_1^0x_2^0x_3^1x_4^1x_5^0}\right]^{1/30} \times \left[\frac{\sum x_1^1x_2^0x_3^1x_4^1x_5^1}{\sum x_1^0x_2^0x_3^1x_4^1x_5^1}\right]^{1/20} \\
&\times \ldots \times \left[\frac{\sum x_1^1x_2^1x_3^1x_4^1x_5^0}{\sum x_1^0x_2^1x_3^1x_4^1x_5^0}\right]^{1/20} \times \left[\frac{\sum x_1^1x_2^1x_3^1x_4^1x_5^1}{\sum x_1^0x_2^1x_3^1x_4^1x_5^1}\right]^{1/5}.
\end{aligned} \tag{44}$$

In the very same way, we use Table 6 for the Bennet decomposition. For factor 1, 'difference change in emission coefficients', we obtain:

$$\begin{aligned}
&\frac{1}{5}\left[\sum(\Delta x_1)x_2^0x_3^0x_4^0x_5^0\right] + \frac{1}{20}\left[\sum(\Delta x_1x_2^1x_3^0x_4^0x_5^0)\right] + \ldots + \frac{1}{20}\left[\sum(\Delta x_1)x_2^0x_3^0x_4^0x_5^1\right] \\
&+ \frac{1}{30}\left[\sum(\Delta x_1)x_2^1x_3^1x_4^0x_5^0\right] + \ldots + \frac{1}{30}\left[\sum(\Delta x_1)x_2^0x_3^1x_4^1x_5^0\right] \\
&+ \frac{1}{20}\left[\sum(\Delta x_1)x_2^0x_3^1x_4^1x_5^1\right] + \ldots + \frac{1}{20}\left[\sum(\Delta x_1)x_2^1x_3^1x_4^1x_5^0\right] \\
&+ \frac{1}{5}\left[\sum(\Delta x_1x_2^1x_3^1x_4^1x_5^1)\right].
\end{aligned} \tag{45}$$

In Appendix D the Matlab program is given that performs the Siegel and Bennet decompositions at the same time. As noted earlier, the results of the simple index numbers for the factors 'emission coefficients' and 'size of the economy' were used to check that the program yields the correct result for Siegel, so that the result according to Bennet is correct, as well.

5.3. Decompositions based on the logarithmic mean

According to Equations 35 and 40, we have

$$v_{ijk}^1 = e_i^1 d_{ij}^1 b_{jk}^1 u_k^1 y^1 \text{ and } v_{ijk}^0 = e_i^0 d_{ij}^0 b_{jk}^0 u_k^0 y^0, \tag{46}$$

so that the weight of the Montgomery decomposition (cf. Table 5) is equal to:

$$w_{ijk}^M = L(v_{ijk}^1, v_{ijk}^0). \tag{47}$$

Using the transformation (37), and the fact that the variable to be decomposed (V) in Equation 34 is the emission of carbon dioxide of sector i, $co2_i$, we arrive at the weight of

[6] The same formulae are used for the other factors.

the Montgomery-Vartia decomposition (cf. Table 5):

$$w_{ijk}^{MV} = w_{ijk}^{M}/L(co2_i^1, co2_i^0). \tag{48}$$

According to Equation 36 we have:

$$s_{ijk}^1 = v_{ijk}^1/co2_i^1 \text{ and } s_{ijk}^0 = v_{ijk}^0/co2_i^0.$$

Consequently, the weight of the Sato-Vartia decomposition (cf. Table 5) is:

$$w_{ijk}^{SV} = \frac{L(s_{ijk}^1, s_{ijk}^0)}{\sum_{j'=1}^{N}\sum_{k'=1}^{m} L(s_{ij'k'}^1, s_{ij'k'}^0)}. \tag{49}$$

Using the transformation (38) we derive from (49) that the weighting factors of the additive Sato-Vartia decomposition (cf. Table 5) read:

$$w_{ijk}^{ASV} = L(co2_i^1, co2_i^0)w_{ijk}^{SV}. \tag{50}$$

In Table 7, we summarize the formulas for the decompositions of the methods based on the logarithmic mean.

Table 7. Formulas for the methods based on the logarithmic mean.

Factor	Montgomery	Montgomery-Vartia	Sato-Vartia	Additive Sato-Vartia
E_i	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{M}\ln\left(\frac{e_i^1}{e_i^0}\right)$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{e_i^1}{e_i^0}\right]^{w_{ijk}^{MV}}$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{e_i^1}{e_i^0}\right]^{w_{ijk}^{SV}}$	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{ASV}\ln\left(\frac{e_i^1}{e_i^0}\right)$
D_i	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{M}\ln\left(\frac{d_{ij}^1}{d_{ij}^0}\right)$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{d_{ij}^1}{d_{ij}^0}\right]^{w_{ijk}^{MV}}$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{d_{ij}^1}{d_{ij}^0}\right]^{w_{ijk}^{SV}}$	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{ASV}\ln\left(\frac{d_{ij}^1}{d_{ij}^0}\right)$
B_i	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{M}\ln\left(\frac{b_{jk}^1}{b_{jk}^0}\right)$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{b_{jk}^1}{b_{jk}^0}\right]^{w_{ijk}^{MV}}$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{b_{jk}^1}{b_{jk}^0}\right]^{w_{ijk}^{SV}}$	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{ASV}\ln\left(\frac{b_{jk}^1}{b_{jk}^0}\right)$
U_i	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{M}\ln\left(\frac{u_k^1}{u_k^0}\right)$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{u_k^1}{u_k^0}\right]^{w_{ijk}^{MV}}$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{u_k^1}{u_k^0}\right]^{w_{ijk}^{SV}}$	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{ASV}\ln\left(\frac{y^1}{y^0}\right)$
Y_i	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{M}\ln\left(\frac{y^1}{y^0}\right)$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{y^1}{y^0}\right]^{w_{ijk}^{MV}}$	$\prod_{j=1}^{N}\prod_{k=1}^{m}\left[\frac{y^1}{y^0}\right]^{w_{ijk}^{SV}}$	$\sum_{j=1}^{N}\sum_{k=1}^{m} w_{ijk}^{ASV}\ln\left(\frac{y^1}{y^0}\right)$

We use just one Matlab program to perform all four decompositions at the same time. It is given in Appendix E. As said earlier, we used the results of the simple index numbers for the factors 'emission coefficients' and 'size of the economy' to check that the program yields the correct result for Sato-Vartia, so that the results according to Montgomery-Vartia, Montgomery and Additive Sato-Vartia are correct, as well.

6. Empirical results

6.1. Dataset[7]

The dataset consists of two Excel files. In the 'Base period' the emissions of carbon dioxide (in million kg) are given for 60 sectors of the Dutch economy in 2004, together with the 60×60 matrix of intermediate deliveries (in million €) and the 60×5 matrix of final deliveries (consumption, government consumption, investments, change in stocks, and exports.) In 'Comparison period' the same data are given for 2005, the matrices of intermediate and final deliveries are recorded in prices 2004.

In the last row of Table 8, we give the percentages of the five largest emitters in the total economy. Together they count for 68 % of the emissions while their share in total final demand is 10.4 %. Between brackets we give the percentages in the total economy of the largest emitter. Not unsurprisingly for the Netherlands it is sector number 25 'Electricity and gas supply'. It accounts for about one third of total emissions whereas its share in total final demand is only 1.3%. From the column 'DCO_2' we gather that the largest emitter accounts for 51.8% of the reduction of carbon dioxide emissions from 2004 to 2005.

Table 8. Carbon dioxide emissions (million kg), ratio and difference change, and total final demand (million €) for the five largest emitters and for the total economy.

#	Sector	CO_2 2004	CO_2 2005	RCO_2	DCO_2	Final 2004	Final 2005
25	Electricity and gas supply	56,538	55,076	0.974	−1,462	7,689	7,657
13	Chemicals; man-made fibres	15,149	15,215	1.004	66	19,978	20,605
12	Petroleum products, cokes, etc.	12,941	12,826	0.991	−115	13,217	13,097
36	Air transport	12,425	12,940	1.041	515	5,910	6,332
34	Land transport	8,821	8,478	0.961	−343	8,121	8,195
	Five largest Emitters	105,874	104,534		−1,340	54,915	55,866
	Total economy	171,419	168,599	0.984	−2,820	580,936	592,633
	% five largest in total economy	67.9 (33.0)	68.2 (32.7)		52.2 (51.8)	10.4 (1.3)	10.4 (1.3)

6.2. Empirical implementation of the decompositions based on the logarithmic mean

As stated before, the logarithmic mean is 'zero value robust'. That is to say, in practice we can replace zeros by epsilon small positive numbers (Ang and Liu, 2007a). In the Matlab program (Appendix D) 10^{-14} is used. But it is not 'change-in-sign robust'. We cannot apply the decompositions on our data set because of the presence of the final-demand category 'change in stocks'. As argued by De Boer (2008), however, this is not a genuine final-demand category. A correct treatment is the following: the final-demand matrix should include a column with the (nonnegative) 'addition to stocks' and the input–output table should have a row with the (nonnegative) 'depletion of stocks'. Due to problems of data collection, national account statisticians only include the balancing item 'change in stocks'. De Boer (2008) solved the problem of changes in sign for stocks by splitting them over the other items of a row according to the pertinent shares in total output. The column sums are no longer equal to total output, so he added a row (that plays no role in decompositions) in which he recorded the adjustment for the stocks. We applied this procedure to our example. As a consequence, the number of final-demand categories is reduced from five to four.

[7] The authors are indebted to Sjoerd Schenau of Statistics Netherlands for putting these two tables at their disposal.

6.3. Results for the multiplicative decompositions

As we conclude from Table 9 the decompositions are very close to each other. From an empirical point of view the split of 'change in stocks' over the other items of the pertinent row had no effect. If 'change-in-sign robustness' is required we need to apply the Fisher decomposition, but if it is not required, like in this example, we advise to use either the Montgomery- Vartia or the Sato-Vartia decomposition since the latter two are easier to program than Fisher's.

Table 9. Results of the multiplicative decompositions.

Sector	Method	RE	RD	RB	RU	RY
25	Fisher	0.9093	1.0670	0.9839	1.0010	1.0201
	Sato-Vartia	0.9093	1.0655	0.9853	1.0003	1.0201
	Montgomery-Vartia	0.9102	1.0646	0.9854	1.0003	1.0119
13	Fisher	0.9850	0.9899	1.0002	1.0088	1.0201
	Sato-Vartia	0.9857	0.9899	1.0005	1.0084	1.0201
	Montgomery-Vartia	0.9857	0.9899	1.0005	1.0084	1.0201
12	Fisher	0.9960	0.9962	0.9732	1.0073	1.0201
	Sato-Vartia	0.9949	0.9962	0.9734	1.0071	1.0201
	Montgomery-Vartia	0.9949	0.9962	0.9734	1.0071	1.0201
36	Fisher	0.9734	1.0070	1.0354	1.0065	1.0201
	Sato-Vartia	0.9734	1.0064	1.0356	1.0063	1.0201
	Montgomery-Vartia	0.9734	1.0064	1.0356	1.0063	1.0201
34	Fisher	0.9449	0.9920	0.9980	1.0024	1.0201
	Sato-Vartia	0.9449	0.9920	0.9974	1.0024	1.0201
	Montgomery-Vartia	0.9499	0.9920	0.9974	1.0024	1.0201

The decompositions according to Fisher and Sato-Vartia (cf. Table 2) satisfy 'proportionality', which implies that for all sectors the effect of the factor 'size of the economy' (y^1/y^0), in six decimal places, is equal to 1.020135. Because Montgomery-Vartia does not satisfy this property $RY^{MV} < 1$ .020135. For the abovementioned sectors 25, 13, 12, 36 and 34, we find 1.011939, 1.020128, 1.020125, 1.020129, and 1.020130, respectively, which are all very close to the upper bound. For all 60 sectors the minimum effect is equal to 1.019438; the maximum to 1.020134; while the mean effect is equal to 1.020098 with a standard deviation of 0.000119. If in this example we desire to have consistency-in-aggregation, we can easily take the nonfulfilment of 'proportionality' for granted and apply Montgomery-Vartia either as one-step decomposition or as a two-step one. If not, we advise to use Sato-Vartia since it satisfies 'proportionality'.

6.4. Results for the additive decompositions

Obviously, the results for the additive decompositions, presented in Table 10, show the same picture as those of the multiplicative decompositions with the same conclusion that the three methods yield the same results so that the split of 'change in stocks' over the other items of a row had no effect either. In subsection 2.6.1, we did not recommend the use of the Additive Sato-Vartia decomposition because it does not possess the theoretical property of consistency-in-aggregation. From the empirical point of view there is another serious drawback as pointed out by Ang et al. (2009). They provide a numerical example of an industry which is the aggregate of two sectors. At the aggregate level (industry) the

Table 10. Results of the three additive decompositions.

Sector	Method	DE	DD	DB	DU	DY
25	Bennet	−5,314	3,662	−912	17	1,115
	Montgomery	−5,253	3,492	−821	18	1,102
	Additive S-V	−5,304	3,540	−829	18	1,112
13	Bennet	−219	−154	3	132	303
	Montgomery	−218	−153	8	128	303
	Additive S-V	−219	−153	8	128	303
12	Bennet	−66	−49	−351	94	257
	Montgomery	−66	−49	−347	91	257
	Additive S-V	−66	−49	−348	91	257
36	Bennet	−342	81	441	82	253
	Montgomery	−341	81	443	80	253
	Additive S-V	−342	81	443	80	253
34	Bennet	−445	−69	−22	20	172
	Montgomery	−445	−70	−22	21	172
	Additive S-V	−445	−70	−22	21	172

decomposition is complete ('no residual term'), but that at the disaggregated level (sectors) the decomposition is not complete since the residuals are unequal to zero. It can be shown that in the framework of this example the decompositions according to Bennet and Montgomery are not only complete at aggregate level, but also at disaggregate level so that the use of one of these methods is recommended. If 'change-in-sign robustness' is required we need to apply Bennet's decomposition, but if it is not required, as in this example, we advise the use of Montgomery's decomposition because it is easier to program than Bennet's decomposition.

7. Summary of methods, their properties and our recommendations

In Table 11 below we summarize the various methods and their properties. Since the non-fulfilment of monotonicity of the methods based on the logarithmic mean plays no role of importance in practice (Balk, 2003) we do not list its fulfilment in Table 11.

In the following picture we summarize our recommendations:

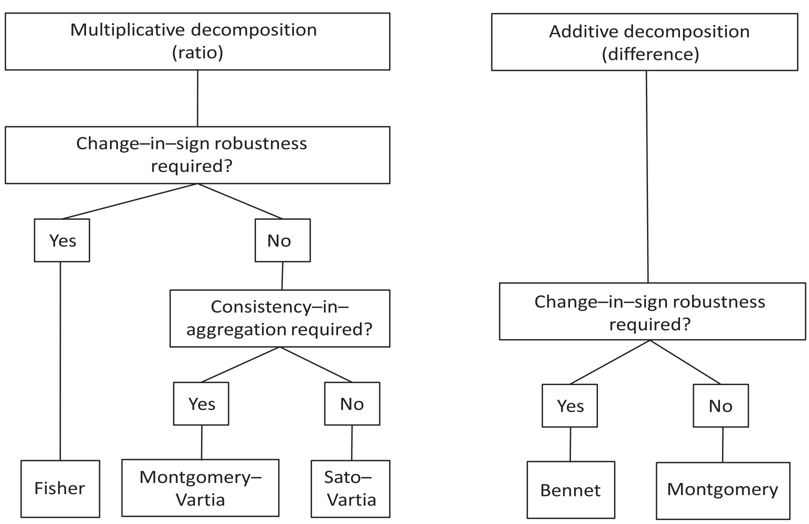


Table 11. Summary of methods and of their relevant properties.

Name	Change-in-sign robust	Proportionality*	Consistent- in-aggregation	Complete at disaggregate level	Simplicity of implementation
Fisher	Yes	Yes	No	Yes	Low
Montgomery- Vartia	No	No	Yes	Yes	High
Sato-Vartia	No	Yes	No	Yes	High
Bennet	Yes	N.A.	Yes	Yes	Low
Montgomery	No	N.A.	Yes	Yes	High
Additive Sato-Vartia	No	N.A.	No	No	High

*There is no analogue of the axiom of proportionality for indicators.

8. Concluding remarks

In this paper, we pay attention to six widely used decomposition methods, all of which share the properties of time reversal and of being ideal. On the basis of theoretical and empirical considerations we show when to use which method.

8.1. Multiplicative decomposition

We considered:

1. Fisher (SDA) is zero-value and change-in-sign robust, satisfies proportionality, but is ***not*** consistent-in-aggregation;
2. Montgomery-Vartia (multiplicative LMDI-I) is zero-value robust, but ***not*** change-in-sign robust, is consistent-in-aggregation, but does not satisfy proportionality; and
3. Sato-Vartia (multiplicative LMDI-II) is zero-value robust, but ***not*** change-in-sign robust, satisfies proportionality, but is ***not*** consistent-in-aggregation.

If there are changes in sign in the data set that cannot be resolved, so the only method that can be applied is Fisher. If the data set is 'change-in-sign robust' we can apply all three methods. If we wish to apply a method that is 'consistent-in-aggregation', we have to apply the Montgomery-Vartia decomposition and take the nonfulfilment of proportionality for granted. If we are not interested in consistency-in-aggregation, we can either apply Sato-Vartia or Fisher, both of which satisfy proportionality. Since the first method is simpler to implement than the latter, we recommend using Sato-Vartia.

8.2. Additive decomposition

We considered:

4. Bennet (SDA): consistent-in aggregation (CIA), zero-value and change-in-sign robust, and complete at aggregate and disaggregate level;
5. Montgomery (additive LMDI-I): CIA, zero-value robust, but ***not*** change-in-sign robust. It is complete at aggregate and disaggregate level;
6. Additive Sato-Vartia (additive LMDI-II): ***not*** CIA. It is zero-value robust, but ***not*** change-in-sign robust; and it is complete at aggregate level, but ***not*** at disaggregate level. These are the two reasons why we do not recommend this method.

If there are changes in sign in the data set which cannot be resolved, only Bennet can be applied. Otherwise, we can use either Montgomery or Bennet, but since the first method is simpler to implement we recommend the use of Montgomery.

8.3. Hands-on toy model

We applied all methods to an example in which the change from 2004 to 2005 in sectoral carbon dioxide emissions of the Netherlands are decomposed into five factors: emission coefficients, production techniques, final-demand mix, demand structure and size of the economy. The data set is ***not*** change-in-sign-robust because of the presence of 'change in stocks' which, as argued by de Boer (2008), is not a genuine final-demand category. It was resolved by spreading the change in stocks over the other items of the pertinent row in the input–output table. We applied the methods of Fisher and Bennet to the full data set, i.e. including the change in stocks, and the other methods which are based on the logarithmic mean to the data set where the number of final demand categories is reduced to four.

8.4. Multiplicative decomposition

From Table 9 it followed that all decompositions are very close to each other so that the split of 'change in stocks' over the other items of the pertinent row had no effect. We advise using either the Montgomery-Vartia or the Sato-Vartia decomposition since the latter two are easier to program than is Fisher's. The effect of the factor 'size of the economy' for the Montgomery-Vartia decomposition turned out to be very close to the effect according to the decompositions of Fisher and Sato-Vartia, both of which satisfy 'proportionality'. In the framework of this example, we can safely take its nonfulfillment for granted and apply, if desired, the two-step decomposition. In general, if one wishes to adopt a method that is consistent-in-aggregation Montgomery-Vartia should be applied, if not, Sato-Vartia is recommended since it satisfies 'proportionality'.

8.5. Additive decomposition

From Table 10 it followed again that the split of 'change in stocks' over the other items of a row had no effect. We advise to use the Montgomery decomposition because it easier to program than Bennet's.

Acknowledgements

This is a revised version of the paper presented at the 26th International Input Output Conference in Juiz de Fora (Brazil), June 2018, under the title 'Energy and environmental studies: When to use which method of decomposition?'. We are indebted to Bert Balk for the numerous discussions on index number theory and decomposition analysis and to the chairperson, Mike Lahr, for his encouraging and stimulating comments. The revised version greatly benefitted from remarks of anonymous reviewers.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

João F. D. Rodrigues http://orcid.org/0000-0002-1437-0059

References

Albrecht, J., D. François and K. Schoors (2002) A Shapley Decomposition of Carbon Emissions Without Residuals. Energy Policy, 30, 727–736.

Ang, B.W. (2004) Decomposition Analysis for Policymaking in Energy: Which is the Preferred Method? Energy Policy, 32, 1131–1139.

Ang, B.W. (2015) LMDI Decomposition Approach: A Guide for Implementation. Energy Policy, 86, 233–238.

Ang, B.W. and K.H. Choi (1997) Decomposition of Aggregate Energy and Gas Emissions Intensity for Industry: A Refined Divisia Index Method. Energy Journal, 18, 59–73.

Ang, B.W. and F.L. Liu (2001) A New Energy Decomposition Method: Perfect in Decomposition and Consistent in Aggregation. Energy, 26, 537–548.

Ang, B.W., F.L. Liu and E.P. Chew (2003) Perfect Decomposition Techniques in Energy and Environmental Studies. Energy Policy, 31, 1561–1566.

Ang, B.W., F.L. Liu and H.-S. Chung (2004) A Generalized Fisher Index Approach to Energy Decomposition Analysis. Energy Economics, 26, 757–763.

Ang, B.W., H.C. Huang and A.R. Mu (2009) Properties and Linkages of Some Index Decomposition Analysis Methods. Energy Policy, 37, 4624–4632.

Ang, B.W. and N. Liu (2007a) Handling Zero Values in Logarithmic Mean Divisia Index Decomposition Approach. Energy Policy, 35, 238–246.

Ang, B.W. and N. Liu (2007b) Negative-value Problems of the Logarithmic Mean Divisia Index Decomposition Approach. Energy Policy, 35, 739–742.

Ang, B.W. and H. Wang (2015) Index Decomposition Analysis with Multidimensional and Multilevel Energy Data. Energy Economics, 51, 67–76.

Ang, B.W., F.Q. Zhang and K.-H. Choi (1998) Factorizing Changes in Energy and Environmental Indicators Through Decomposition. Energy, 23, 489–495.

Balk, B.M. (1996) Consistency-in-Aggregation and Stuvel Indices. Review of Income and Wealth, 42, 353–363.

Balk, B.M. (2003) Ideal Indices and Indicators for Two or More Factors. Journal of Economic and Social Measurement, 28, 203–217.

Balk, B.M. (2008a) Price and Quantity Index Numbers: Models for Measuring Aggregate Change and Difference. CambridgeUK, Cambridge University Press.

Balk, B.M. (2008b) Searching for the Holy Grail of Index Number Theory. Journal of Economic and Social Measurement, 33, 19–25.

Balk, B.M. (2016) A Review of Index Number Theory. Wiley StatsRef: Statistics Reference Online.

Bennet, T.L. (1920) The Theory of Measurement of Changes in the Cost of Living. Journal of the Royal Statistical Society, 83, 455–462.

Boyd, G.A., J.F. McDonald, M. Ross and D.A. Hanson (1987) Separating the Changing Composition of U.S. Manufacturing Production From Energy Efficiency Improvements: A Divisia Index Approach. Energy Journal, 8, 77–96.

Boyd, G.A. and J.M. Roop (2004) A Note on the Fisher Ideal Index Decomposition for Structural Change in Energy Intensity. Energy Journal, 25, 87–101.

Carli, G.R. (1764) Del Valore et Della Proporzione ‘de Metalli Monetati. Reprinted in: G.G. Destefanis (1804). Scrittori Classici Italiani di Economia Politica, Vol. 13. Milan, Opere Scelte di Carli, Milano, 297–366.

Choi, K. and B.W. Ang (2012) Attribution of Changes in Divisia Real Energy Intensity Index – an Extension to Index Decomposition Analysis. Energy Economics, 34, 171–176.

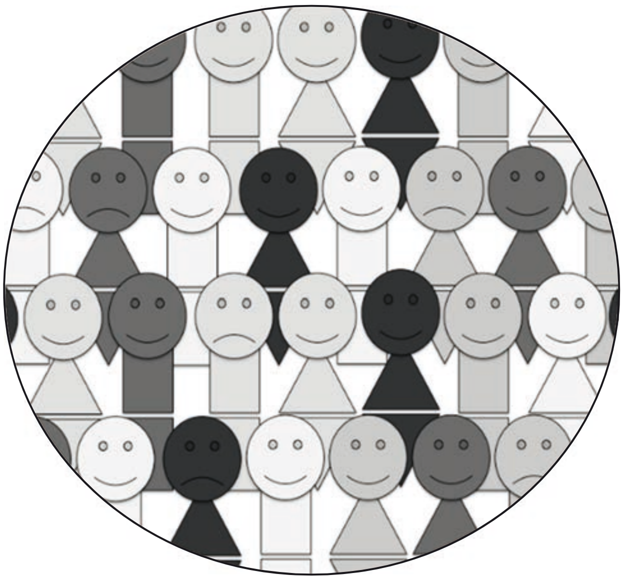
Chung H.S. and H.C. Rhee (2001) A Residual-free Decomposition of the Sources of Carbon Dioxide Emissions: A Case for the Korean Industries. Energy, 26, 15–30.

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, 97–109.
De Boer, P.M.C. (2009a), Multiplicative Structural Decomposition Analysis and Index Number Theory: An Empirical Application of the Sato-Vartia Decomposition. Economic Systems Research, 21, 163–174.
De Boer, P. (2009b) Generalized Fisher Index or Siegel-Shapley Decomposition? Energy Economics, 31, 810–814.
De Haan M. (2001) A Structural Decomposition Analysis of Pollution in the Netherlands. Economic Systems Research, 13,181–196.
Dietzenbacher, E. and B. Los (1998) Structural Decomposition Techniques: Sense and Sensitivity. Economic Systems Research, 10, 307–324.
Dietzenbacher, E., A.R. Hoen and B. Los (2000) Labor Productivity in Western Europe 1975–1985: An Intercountry, Interindustry Analysis. Journal of Regional Science, 40, 425–452.
Diewert, W.E. (2005) Index Number Theory Using Differences Rather Than Ratios. The American Journal of Economics and Sociology, 64, 311–360.
Dutot, C.F. (1738) Réflections Politiques sur les Finances et le Commerce, Volume I (les Frères Vaillants et N. Prevost, The Hague). Édition Intégrale Publiée Pour la Première Fois par Paul Harsin (1935) Bibliothèque de Faculté de Philosophie et Lettres de l'Université de Liège, Fascicule LXVII, Paris- Liège.
Eichhorn, W. (1978) Functional Equations in Economics, Reading, MA, Addison-Wesley.
Eichhorn, W. and J. Voeller (1976) Theory of the Price Index. Lecture Notes in Economics and Mathematical Systems, 140, Berlin, Springer-Verlag.
Fisher, I. (1922) The Making of Index Numbers. Boston, Houghton Mifflin.
Hoekstra, R. and J.C.J.B. van den Bergh (2003) Comparing Structural and Index Decomposition Analysis. Energy Economics, 25, 39–64.
Laspeyres, E. (1871) Die Berechnung Einer Mittleren Waarenpreissteigerung. Jahrbücher für Nationalökonomie und Statistik, 16, 296–314.
Montgomery, J.K. (1929) Is There a Theoretically Correct Price Index of a Group of Commodities? Rome, International Institute of Agriculture.
Montgomery, J.K. (1937) The Mathematical Problem of the Price Index. London, P.S. King & Son.
Paasche, H. (1874) Ueber die Preisentwicklung der Letzten Jahre Nach den Hamburger Börsennotierungen. Jahrbücher für Nationalökonomie und Statistik, 23, 168–178.
Sato, K. (1976) The Ideal Log-Change Index Number. The Review of Economics and Statistics, 58, 223–228.
Seibel, S. (2003) Decomposition Analysis of Carbon Dioxide Emission Changes in Germany: Conceptual Framework and Empirical Results. European Commissions Working Papers and Studies, THEME 2 Economy and finance, EUROSTAT publications, Luxemburg.
Shapley L. (1953) A Value for n-Person Games. In: H.W. Kuhn and A.W. Tucker (eds.) Contributions to the Theory of Games, vol. 2. Princeton, Princeton University, 307–317.
Siegel, I.H. (1945) The Generalized 'Ideal' Index-Number Formula. Journal of the American Statistical Association, 40, 520–523.
Sun, J.W. (1996) Quantitative Analysis of Energy Consumption, Efficiency and Savings in the World. Turku School of Economics and Business Administration, Technical Report Series A-4.
Sun, J.W. (1998) Changes in Energy Consumption and Energy Intensity: A Complete Decomposition Method. Energy Economics, 20, 85–100.
Su, B. and B.W. Ang (2012) Structural Decomposition Analysis Applied to Energy and Emissions: Some Methodological Developments. Energy Economics, 34, 177–188.
Su, B. and B.W. Ang (2014) Attribution of Changes in the Generalized Fisher Index with Application to Embodied Emission Studies. Energy, 69, 778–786.
Su, B. and B.W. Ang (2015) Multiplicative Decomposition of Aggregate Carbon Intensity Change Using Input-Output Analysis. Applied Energy, 154, 13–20.
Törnqvist, L. and E. Törnqvist (1937) Vilket är Förhållandet Mellan Finska Markens och Svenska Kronans Köpkraft? Ekonomiska Samfundets Tidskrift, 39, 1–39.

Vartia, Y.O. (1974) Relative Changes and Index Numbers. Licentiate Thesis, University of Helsinki, (The Research Institute of the Finnish Economy, Helsinki, 1976).
Vartia, Y.O. (1976) Ideal Log-Change Index Numbers. The Scandinavian Journal of Statistics, 3, 121–126.
Wang, H, B.W. Ang and B. Su (2017a) Multiplicative Structural Decomposition Analysis of Energy and Emission Intensities: Some Methodological Issues. Energy, 123, 47–63.
Wang, H, B.W. Ang and B. Su (2017b) Assessing Drivers of Economy-Wide Energy Use and Emissions: IDA Versus SDA. Energy Policy, 107, 585–599.
Yan, J., B. Su and Y. Liu (2018) Multiplicative Structural Decomposition and Attribution Analysis of Carbon Emission Intensity in China, 2002–2012. Journal of Cleaner Production, 198, 195–207.
Xu, Y. and E. Dietzenbacher (2014) A Structural Decomposition Analysis of the Emissions Embodied in Trade. Ecological Economics, 101, 10–20.
Young, A. (1812) An Inquiry Into the Progressive Value of Money in England as Marked by the Price of Agricultural Products. Hatchard, Piccadilly.
Zhang, H. and M.L. Lahr (2014) China's Energy Consumption Change From 1987 to 2007: A Multiregional Structural Decomposition Analysis. Energy Policy, 67, 682–693.

UNDERSTANDING SOCIAL CHANGE

A DECOMPOSITION APPROACH



PARFAIT M. ELOUNDOU-ENYEGUE
SARAH C. GIROUX
MICHEL TENIKUE

Preface

A story is often told of a drunk man who loses his key at night and only searches for it under a lamppost, because that is where the light is. Today, the social sciences may be in a similar predicament: Research on social change (a macro-phenomenon) often relies on (micro) survey data. Compared to macro-level regression, micro-level research is the 'methodological lamppost,' the better-lit area of quantitative social science, where analyses are deemed statistically most rigorous. Thus, to study how education affects national economies, a researcher might rely on estimates of how a worker's education level affects his or her income. In public health, to estimate how economic growth affects national obesity, our researcher might likewise extrapolate from the individual-level link between income and obesity. In demography, to capture the economic benefits of declining fertility, the researcher might explore how the economic fortunes of various family members depend upon the size of their family.

All these three examples show a discrepancy between the scale at which a problem occurs and the scale at which it is studied. Social change is a dynamic and macro-process, yet we often approach its study with micro cross-sectional data. This creates two fallacies. One, ecological, relates to unit of analysis, specifically, studying individual people rather than the entire society. The second, historical, fallacy amounts to "reading history sideways" i.e., reading cross-sectional data as indicating a historical trend. Most researchers acknowledge these fallacies: cross-sectional studies of interpersonal differences may reveal why some people have a higher propensity to experience a given condition (say, obesity) but they cannot explain why/how the social magnitude of this phenomenon changes over time. Nonetheless, many studies still directly infer macro-relationships from micro-results.

If an investigator wishing to avoid these biases turns to macro-level analysis, her peers may see this work as lacking statistical rigor or obscuring the differences between people in the same country (Rodrik 2005). The question, therefore, becomes, "How to avoid ecological and historical fallacies while also maintaining some statistical rigor?"

The challenge is to recognize within-country diversity but still show how the diverse behaviors of people in a country add up to produce a collective change. For this kind of analysis, decomposition is a useful tool. Even if it does not reveal ultimate causes, it points to main "sources" of change, i.e., the groups or proximate processes driving the social change.

Decomposition is widely applied across many disciplines, especially within the social sciences. However, its different variants remain insufficiently integrated and literature on the topic remains fragmented. Few textbooks offer a comprehensive introduction highlighting the method's wide range of possibilities. None to our knowledge clearly explains how to enrich elementary decompositions let alone combine them with other methods. We fill a gap with this monograph, which updates and augments an earlier version published in 2010.

Our work on this document was conducted under the auspices of the International Union for the Scientific Study of Population (IUSSP) with funding from the William and Flora Hewlett Foundation, and material support from Cornell University (USA), the Institute for Training and Demographic Research (IFoRD, Cameroon), and the Higher Institute of Population Sciences (ISSP, Burkina Faso). We received feedback from numerous colleagues, most prominently Françoise Vermeylen, Gervais Béninguissé, Jean-François Kobiané, Silvere Konan Yao, Crispin Mabika Mabika Gilles Gohi, Nina Trautmann Chaopricha and Latif Dramane as well as from dozens of participants in the various workshops conducted under this IUSSP project. We tested early versions of this monograph with support from Serge Bahoken, Firmin Zinvi Charles Mouté, Justin Dansou, Sandrine Mbele, and Habibou Ouedraogo. Several researchers and students — Yédodé Ahokpossi, Jacques Emina, Edzengte Pascal, Kouam Nadège and Bela Foe Chrystelle. — were early volunteers to apply the method in their own research. Their experiences informed subsequent revisions, as did the insights of several non-scientists (Pie-Marie Belibi, Stephanie Hissoak, and Vincent Onguéné) whose comments helped make the document more accessible to a broader audience. In the end, however, the authors are solely responsible for the remaining shortcomings of this document.

Ithaca, August 2017

Sommaire

Chapter I
Introduction

Can scientific research be a useful guide for social policy? Whether reducing unemployment, expanding school enrollment, containing social inequality, or reducing mortality, the constant challenge facing those who plan and implement policies is to promote desirable social change. To meet this challenge, policymakers must understand the levers of social change — and, in this context; scientific research has a key role to play.

Social science has made great strides over the last half-century yet these strides were greatest in micro-level studies. Thanks to advances in computing and communication technology, researchers can now collect, share, and process statistical data on millions of households and individuals. Researchers can explore the detail and mix of factors shaping individual behaviors.

However, an analyst of societal change will not be satisfied with this micro-level detail alone. She may find the detail useful but still need to convert the micro-level information into valid inferences about social change. Over half a century ago, Robinson (1950) warned against ecological bias, noting that relationships observed macroscopically need not match those occurring at the individual level and vice versa.[1] Thornton (2001) likewise drew attention to an equally harmful bias, the tendency to "reading history sideways." In essence, it is wrong to use cross-country comparisons to draw conclusions about development trajectories. Researchers must thus be careful in navigating the space between micro-data and macro-issues, building on the detail and robustness of microlevel statistics as they aggregate them to inform macro-level questions. The decomposition methods presented here can facilitate this conversion.

[1] For example, richer *countries* may have higher rates of obesity but that does not imply that richer people creases with the level of wealth of a country does not necessarily imply that within countries, richer *people* have a greater propensity to become obese than poor people. Similarly, national rates of divorces increase with national crime rates, but that does not necessarily mean that people from divorced families are more prone to crime.

I.1 What is decomposition?

I.1.1 Definitions

To decompose something is to break it into its elementary components. In biology, decomposition is a process of organic material decay. In chemistry, it is the process of bursting a complex molecule into simpler molecules or atoms. In physics, it can describe the trajectory of a projectile, projecting its path into a three-dimensional space that includes vertical, horizontal, and transversal components. In the social sciences, decomposition can serve to estimate how several elementary processes (or groups) fuel an aggregate social change. For example, how do different regions of a country contribute to national wealth, or how do women from different education backgrounds contribute to change national fertility?

Several variants of decomposition methods exist in the literature but the fundamental idea is the same: how to partition a functional set into its elementary constituents. To appreciate the uniqueness of decomposition vis-à-vis other methods, we specify below some situations in which it applies.

I.1.2 Application areas

Within the social sciences, decomposition analysis is particularly useful for studying aggregate social change considered as any transformation — induced or spontaneous — in the structure, functioning or performance of a social community. The most relevant changes are quantifiable transformations resulting from an aggregation of individual behaviors.

Decomposition is thus not applicable in studies where individuals are the unit of analysis. Decomposition is also of little help when studying societal outcomes that do not result from the aggregation of individual behavior. These two exceptions aside, the method applies to a wide range of social processes in demography, economics, political science, and sociology (Kitagawa 1955, Oaxaca 1973; DasGupta 1993; Vaupel and Romo 2003). The only requirements are to have quantifiable social outcomes that reflect an aggregation of individual outcomes.

Quantification: The outcome under study should be quantifiable, i.e., captured as an absolute number, an average, a percentage, a ratio, or a measure of inequality. This excludes qualitative processes such as a country's political transition from autocratic to democratic rule, or its economic transition from subsistence to market orientation, or its demographic transition from extended to nuclear family systems. Even in these cases, dichotomies such as autocracy/democracy, subsistence/market, or extended/nuclear can be expanded into a continuum. For example, one can replace the subsistence/market dichotomy by a continuous outcome such as the percentage of people working for an employer other than family. With such reformulation, phenomena that may a first seem qualitative lend themselves to decomposition analysis.

Aggregation of individual behavior: Sociologists distinguish between social outcomes that an intrinsic feature of the whole society versus aggregate outcomes. The former have no correspondent at the individual level. An example might be a country's laws on abortion. The latter resulting from an aggregation of individual characteristics. Examples may include the fertility level of a country or its average consumption of tobacco. This tobacco consumption is an aggregated (not intrinsic) feature of the society because a country does not smoke; individual citizens do.

Gradual change: Decomposition methods are seldom applicable to rare or sudden, accidental processes (e.g., studying the number of casualties in an earthquake or a sudden outbreak of cholera). Rather, they best apply to studies of processes that change gradually over time.

I.1.3 Mode of explanation

Decomposition analysis is about the sources rather than the ultimate causes of change. It is better at determining the origin of a change rather than why the change occurred. It mechanically accounts for the sources of change, specifically, how different processes or groups contribute to generate a social change. While full causal analysis seeks to reveal the ultimate causes, decomposition merely reveals the proximate processes or groups from which the change occurred.

From what? (Proximate processes). In a basic demographic decomposition, a national outcome is cast as the weighted average of behaviors observed among different sub-populations in the country. The total change is set as coming from two proximate sources, namely (a) compositional changes such as changes in population composition like the relative weight of constitutive sub-populations and (b) behavioral changes such as changes in group behavior or outcomes.[2]

Decomposition can also apply to outcomes resulting from a chain of events where, say, process #1 leads to process #2, then process #3 … For instance, the amount of locally-produced food available in markets depends on a sequence of three processes: local food production, allocation of harvest between personal consumption and market, and food transportation into markets. Any change in amounts of food available in markets will reflect a mix of changes in these three elementary processes.

By whom? In addition to revealing the processes by which change occurs, researchers may seek to estimate the relative contribution of each group to the total change. How for instance how much do people from various regions, age categories or educational categories (for instance) contribute to the change. The decomposition and its identification of proximate processes and groups can pave the way for a deeper exploration of the ultimate causes of change.

I.2 Decomposition versus other methods

Imagine a researcher wishing to explain a recent decline in mortality in her country. In all likelihood, decomposition will not be the first or only tool considered. The researcher has multiple options including qualitative analysis, trend analysis, and regression.

She could use a qualitative approach based on key informants, focus group discussions, or archival data to shed light on historical events occurring during the decline. Did the country make new investments or implement new projects? Did it make recent scientific discoveries or put new drugs on the market? Was there an outbreak of unusual events, a rise of effective leaders, or an improvement in the water supply and medical services?

Our researcher might also decide to monitor recent trends in mortality, to pinpoint the exact moment when the decline began and to review other key events preceding this decline. She will ask about social processes that appear to co-vary with mortality. This approach can complement a qualitative reliance on key informants by systematically testing the hunches from the informants with trend data. However, the conclusions remain subjective if all she did was to eye-ball and visually compare various trends without attempting a formal statistical test.

[2] We will return to these two sources and can already offer one example here. Imagine a country today that experienced war in 1970. One can well-imagine that the views about war –specifically, whether the country ought to embark on another war today — would likely vary by generation. For instance, people born before the 1970s who have lived through a previous war may be less likely to support another war than the younger generations. Therefore, the results of a poll about a country's readiness to go to war in 2000 and again in 2025 might change simply because the percentage of people born before 1970 would have dwindled as this group ages (a compositional effect). It could also be that the global political environment between 2000 and 2025 changes in ways that changes people's attitudes, perhaps making them less supportive of war (a behavioral effect). Overall, the total change in national attitudes will be a combination of these two effects.

More formally, our researcher could use formal correlation/regression analysis. Unfortunately, her macro-correlations would obscure the detail of individual responses. On the other hand, regression using individual data is more detailed and rigorous but it does not address the right level of analysis for someone interested in social change.

Micro and macroscopic approaches thus complement each other: one being more rigorous and detailed, and the other more relevant for national policy design. This complementarity permits two possible kinds of integration (Figure 1). The first looks at how the 'macro' (society) affects the 'micro' (individual) and is known as multilevel regression (Luke 2004). The second conversely looks at how micro behaviors aggregate to shape macro-level outcomes and change therein; this is the essence of decomposition.

Table 1 summarizes the differences between the four approaches discussed above. In particular, they differ in their presumed drivers of change. The drivers, in demographic decomposition, are either compositional forces or behavioral factors. In some qualitative analyses, change can come from a single person or a key event in the study community.

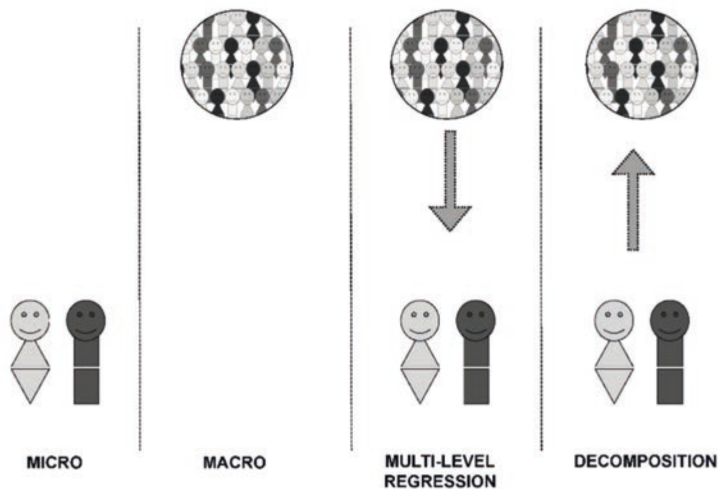


Figure 1. Four modalities of the integration of micro and macro analysis.

Approach	Qualitative analysis	Trend Analysis	Regression	Decomposition
Presumed driver of change	People of key events	Other social change	Factors statically associated with the phenomenon	Change in the size or behavior of social groups
Data sources	Focus groups, interviews, archives	National statistics	Individual level data	Individual level data (by groups)
Central questions	Why?	When?	Who?	How?
Strong points	Elucidate processes	Chronological sequencing	Rigorous statistical estimation of associations	Reliability, simplicity
Weak points	1, 2, 4	4	3, 4	5

Weak points 1. Not statistically representative; 2. Statistical associations are not established; 3. Chronology not established; 4. Competing explanations are not eliminated; 5. Causal relationships are not clarified.

Table 1. Comparison of four explanation approaches to social change.

I.2.1 Advantages and limitations

Unlike causal analysis, decomposition does not establish causation. Returning to our earlier example about mortality, decomposition may reveal which sectors, groups, or causes of death accounted for a rise in overall mortality, but the results will not say why these declines occurred. This shortcoming is severe, because sound policies require a clear understanding of cause and effect. However, by revealing the proximate sources of change, decomposition analysis can usefully guide social intervention and targeting. Moreover, it has the advantage of being simple, flexible, easy to interpret, and compatible with other methods.

Simplicity. Decomposition methods are simple to apply and interpret. Their application requires no fancy statistical analysis, complex calculations, or advanced software. Much of the work is doable with a spreadsheet program such as Excel, and the input data are often readily available in reports and online tabulations.[3] When the input data is publicly available, the analyses are made easier and the process transparent because other researchers can easily check both the input data and the results.

Flexibility. The decomposition approach is quite flexible in its application. A reasonably creative analyst can move from basic forms to complex combinations, depending on her needs. The presentation and illustrations in this monograph seek to highlight this flexibility.

Easy interpretation. The interpretation of decomposition results is intuitive and easy. Compared to the results from regression analysis (ordinary least squares, logit, and odds ratios), the statistics generated by a decomposition analysis are readily expressed in plain language. The results simply indicate the percentage of social change coming from a given process or group.

Compatibility. The decomposition method is compatible with many other methods, including micro-regressions, multilevel analysis, simulations, and qualitative analysis. It can help aggregate the results from micro-regressions. It can combine fruitfully with multilevel analysis. It can serve as a prelude to a qualitative analysis.[4] For all these reasons, decomposition can apply eclectically to a wide range of fields and methodological traditions. Importantly, it does not replace or compete with other methods but, rather, it complements them to improve the quality of the findings.

Transparency. This is a major strength of decomposition. Contrasting with the opacity of most other methods, the average reader can easily check the internal consistency of results from decomposition analysis. In a multivariate regression, for example, the reader often has few opportunities to check the accuracy of the regression coefficients presented and s/he must trust the researcher.[5] In a decomposition analysis, by contrast, s/he can check the accuracy of final results by reviewing the initial information and following, within the table of findings, the subsequent data transformations and the plausibility of final results.

---

[3] See, for instance, the Demographic and Health Survey (www.statcompiler.com) or the World Bank (http://data.worldbank.org/data-catalog/world-development-indicators).

[4] For example, in our previous example about attitudes vis-a-vis war, the researcher's decomposition results may show that, contrary to expectations, much of the change in national attitudes vis-à-vis war reflects a true behavioral change rather than a compositional change. In other words, many people in the country actually changed their attitudes about war. Alternatively, she may find that most of the change came from people between the ages of 45 to 60. Armed with this information, she can now go ahead and conduct more investigations into why these people changed their minds during that period.

[5] The situation is improving, as researchers must increasingly make their data publicly available. However, the reader must have access to the computing program used by the researchers and be able to retrace all steps, from coding to the final modeling, which is rarely convenient.

I.2.2 Political relevance and applications

Decomposition methods can apply to many of the social changes underway across the globe. In particular, they can inform the study of several development goals pursued under the United Nations' Sustainable Development agenda (SDG), including poverty, health, inequality, gender, or basic education, for instance. Achieving these ambitious goals requires an efficient use of countries' scarce resources, which implies a clear understanding of the drivers of socioeconomic change. The rapid and uneven demographic changes occurring in many countries create fast-changing and diverse societies that can no longer be fully understood without careful attention to disaggregated evidence. Decomposition methods can help.

I.3 The main types of decomposition

Decomposition is not a single method but set of related methods. Its variants appear separately in different fields and they have not been sufficiently integrated. Below, we review some of these variants and their differences based on four criteria: (a) the type of independent variable, (b) the type dependent variable, (c) the link function, and (d) the complexity of the analysis.

Criterion 1: The independent variable type

One can distinguish demographic, regression, and temporal decompositions, depending on whether the independent variable is nominal, interval or ordinal.

- In a demographic decomposition, the independent variable is nominal, e.g., country region, age group, ethnicity, or marital status. For instance, a national outcome can be viewed as the population-weighted average of outcomes across all of the countries' regions. For example, the national support for a given cause or person (e.g., the president) will be a weighted average of support across all regions. This support will change if the relative size of the regions happens to change over time or if the views of people within any of the regions change.
  Note that demographic decomposition also applies to an ordinal variable (e.g., socioeconomic status) if it is treated nominally, i.e., with no explicit attention to the order of categories.
- In a temporal decomposition, the independent variable is ordinal, and the order between categories is explicitly considered. For instance, if age group is the independent variable, we keep in mind that people aged 15-19 are younger than people aged 20-29 years, who are themselves younger than people aged 30-39 years.
- In a regression decomposition, finally, the independent variable is quantitative. Examples include a person's years of education, the number of siblings, or income in dollars. The starting point in regression decomposition is to have an estimate of the statistical effect of an independent variable (e.g., years of education) on some outcome (e.g., health). Health outcomes then are expected to change either because the amount of education changes or because the average effect of education changes. The decomposition in this case seeks to determine how much of the total change in health reflects the change in the quantity of education versus the effectiveness of education.

Criterion 2: The dependent variable

For the simplest basic decomposition, the dependent variable is an average or a percentage. In more complex (derived) decomposition, the dependent variable may be a more complex measure such as inequality.

Criterion 3: The functional relationship

The main criterion here is the type of relationship linking the independent to the dependent variable. The type and complexity of this functional relationship vary (see Figure 2), and we can specify three types:

- A demographic relationship: Basically, the Y value of the entire country is a weighted (by demographic weight) average of prevailing values in the various subpopulations of the country (y_i);
- A statistical relationship, specifically, a regression relationship between Y and X; and
- A mathematical relationship: in this case, the dependent and independent variables are linked by a simple mathematical relationship (which involves a quotient, sum, product, or log). Unlike the statistical relationship, it is a true relationship that does not vary across countries or situations. The only change is in the values of these variables.

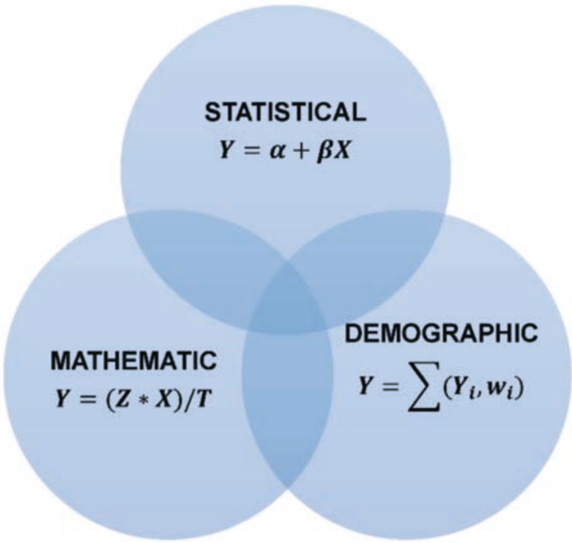


Figure 2. Basic types of decomposition.

Each of these basic functional relationships can become more complex. In regression analysis, for instance, one can move from a single independent variable and a linear specification to more elaborate forms that feature several independent variables, non-linear forms, or multi-level relations.

Criterion 4: The degree of complexity

This criterion distinguishes simple from advanced or mixed decompositions that combine two or more elementary forms. The basic forms shown in Figure 2 are thus the building blocks for generating complex forms. Therefore, as indicated in Figure 3, an elementary decomposition may yield a (derived) decomposition when its dependent variable or functional relationship becomes more complex. We obtain a mixed decomposition by combining two or three basic forms. This book's organization follows the typology shown in Figure 2. We begin with basic and derived forms and work our way up to mixed forms.[6]

[6] We should note here that a number of highly specialized decomposition formulas have been developed to study specific questions such as life expectancy (Vaupel, 2003), job discrimination (Oaxaca, 1973) and poverty or inequality (Shorrocks, 2013). The monograph does not show these detailed formulations but readers will easily see where these formulations fit into the taxonomy presented here.

I.4 Conventional notations

To facilitate the presentation, we use the following conventional notations for variables, indices, and historical change:

Variables. Following common practice in the social sciences, dependent and independent variables are indicated by the letters Y and X, respectively. In addition, we use capital letters when referring to national outcomes, and reserve lowercase letters when referring to individuals or sub-populations. Y thus designates a dependent, macro-level variable; y designates a dependent variable at the micro or meso-level; X designates an independent variable at the macro-level; and x designates an independent variable at the micro or meso-level. Therefore, for instance, if we study the effects of education on mortality, Y is the national mortality rate, while y is mortality rate in a subpopulation (e.g., for those between the ages of 15 to 19 years). In the study, X is the national education level, while x is the level of education in a subpopulation, e.g., for the 15-19 year olds.

Weighing. The letter w measuring the weights of different subgroups is mostly used in demographic decomposition or its close variants. The weights often reflect population size, i.e., the percentage of the national population in a given category.

Regression parameters. Regression decomposition includes conventional regression parameters such as:

- $\boldsymbol{\alpha}$, the intercept, which is the value of Y (or y) when X (or x) is 0;
- $\boldsymbol{\beta}$, the marginal increase of Y (or y) when X (or x) increases by one unit; the more complex regression analyses will integrate the case where β is a vector and take multiple values; and
- $\boldsymbol{e}$, the error term.

Indices. The presentation will also use the following indices:

- *j* indexes groups; e.g., y_j denotes the value of the dependent variable for the j group, while x_j indicate the value of the independent variable for the same group;
- *t* denotes time; e.g., Y_t indicate the value of the dependent variable for a given year t and for the entire population (e.g., the average mortality in Senegal as of 1990); and
- *a* indexes age; thus, Y_a is the value of the dependent variable for a given age group e.g., the specific fertility rate for the age group. Occasionally, we use the term + to show all ages above the referenced age; in fertility analysis Y_{39+} might be the average fertility for all ages above 39.

Historical change. Δ indicates the historical change. For example, ΔY represents the historical change in the outcome being studied, specifically, the difference in the Y values observed at two points in time (e.g., $Y_{t1} - Y_{t,}$ $t_1 > t_{...}$).

Averages. The annotations will distinguish between two types of averages, whether cross-sectional or historical. A cross-sectional average is the average in the population at a specific time t. It is calculated over several groups at one point in time. Since the dependent variable in most basic decompositions is an average, these averages will simply be denoted Y (or y). A historical average is the average for the same group over two periods. We signal it by adding a bar over the corresponding letter. Thus, $\bar{Y}$ is an average between two periods for the value of the dependent variable at the national level. Likewise, $\bar{y}$ indicates the average value of two years of the dependent variable for one subpopulation.

$$\bar{Y} = (Y_{t1} + Y_t)/2$$

I.5 Structure of the monograph

Our presentation progresses from simple to complex decompositions. We began by defining and comparing decomposition with other methods. We then offered a typology of decomposition methods, depending on the type of data and the complexity of the analysis. Decompositions are labeled 'demographic,' 'temporal,' and 'regression' depending on whether the independent variable is nominal, ordinal, or quantitative. These elementary types are then modifiable to yield nested decompositions. Once again, elaborations may reflect greater complexity in the type of dependent variable (e.g., a measure of inequality instead of an average) or the link function between independent and dependent variables (e.g., a curvilinear or multivariate relationships instead of a simple linear relationship).

Finally, one can combine elementary types to produce advanced decompositions (Figure 3). We present a few examples of such advanced decompositions, but investigators can use their own imagination to generate additional forms.

To facilitate understanding, the document mixes verbal descriptions with mathematical formulas, illustrative examples, annotated charts, graphs, and figures. This pedagogical approach hopefully makes the material accessible to a wide pool of readers with different learning styles.

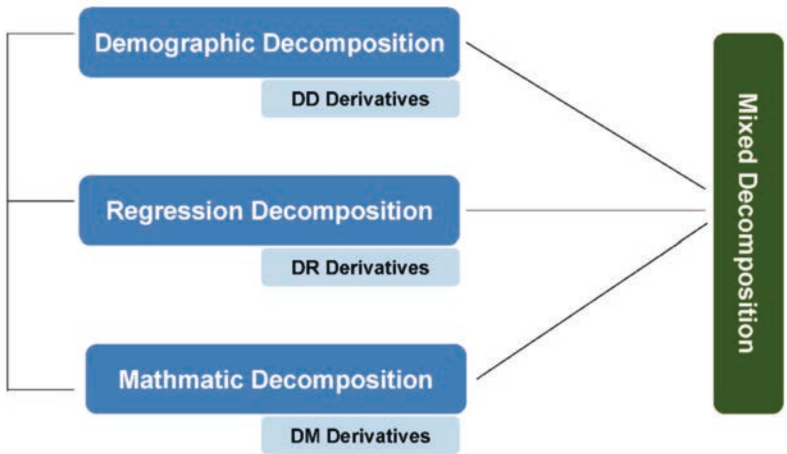
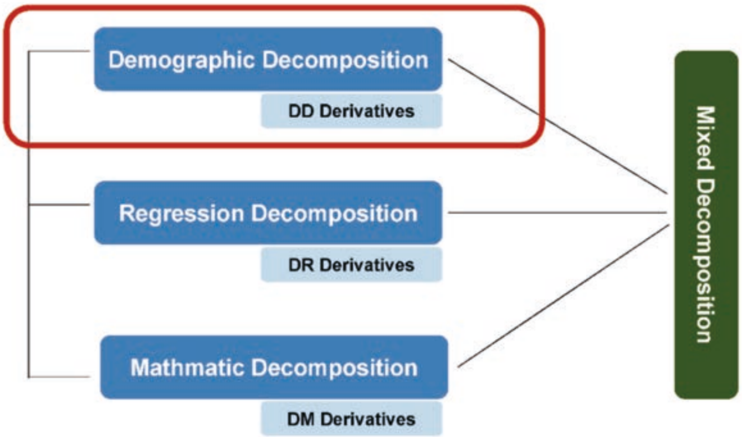


Figure 3. Decomposition types and combinations.

Chapter II
Demographic decomposition



II.1 Basic demographic decomposition

II.1.1 Problem Type

This first type of decomposition applies to national outcomes (Y) that are an aggregated outcome from several subpopulations (y_i), each weighted by its relative size (w_j). Formally,

$$Y = f\,(y_j, w_j) \qquad [II.1]$$

For example, the mortality rate of a country is the weighted average of rates in different regions or socioeconomic groups. In this formulation, Y, the dependent variable, is quantifiable and X, the independent variable, is measured nominally. The variable X and its categories must meet at least four criteria:

Exhaustiveness: The full set of categories for the independent variable must be such that each member of the population belongs to one and only one category j. In other words, the set of categories must cover the entire population, with categories also being mutually exclusive.

Distribution: The number of categories for the classification variable should be neither too small (> 2) nor too large. With too few categories, the analyses are not detailed enough to be informative. Conversely, having too many categories spreads the data too thinly. Thus, variables such as sex (with too few categories) or age (too many categories if measured in singe years) are not ideal as classification variables.

Variability: The size of the individual categories must fluctuate over time. Otherwise, the compositional effect in the decomposition analysis will remain zero. For this reason, gender is, once again, a poor classification variable unless the researcher is dealing with a very dynamic population with rapidly-changing sex ratios.

Relevance: A good classification variable should be theoretically relevant or policy-relevant. Variables such as "region" for instance usually meet the criterion of policy-relevance, whereas education for instance is theoretically relevant if the outcome being studied is expected to depend on one's education level.

II.1.2 Visual representation

As stated in the introduction, decomposition methods seek to identify the sources of change, whether substantive (with processes driving the change) or sociological (with groups or people driving the change).

The figure below illustrates a basic decomposition with a study of gender parity in education. The left-hand side of the diagram shows five squares capturing the trend in parity for the country as a whole, with darker colors reflecting greater parity. The chart shows a steady progress at the national level from high educational inequality between boys and girls (the white square on the far left) to parity (the black square on the far right). Obviously, the country became more gender-equitable but for several reasons, a researcher may wish to understand how this evolution occurred, specifically, how the country's various social classes contributed to it.

On the right side of the diagram (Frame B) are two possible, opposite, scenarios of convergence. The first (B1) shows a horizontal convergence, with the educational gap closing at the same rate for all groups. Conversely, Frame B2 shows a vertical convergence, with change starting among the top income group before gradually trickling down. In year 2, the top income group had already achieved parity, while gender inequality remained prevalent in lower-income groups. If we ask about the groups that led the change, we get different answers from the two scenarios: in the first case, all groups evolved simultaneously, while in the second case the change came from above.

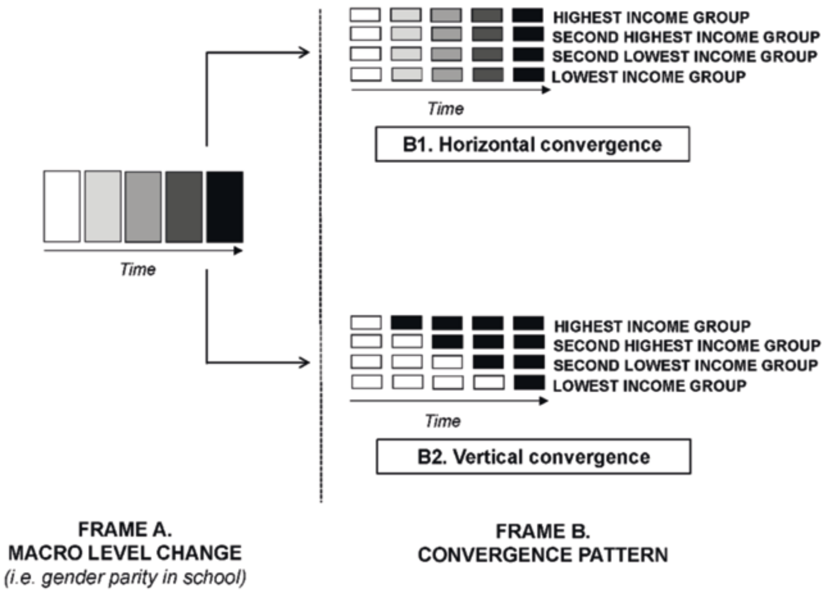


Figure 4. Vertical versus horizontal convergence in education.

II.1.3 Example

While the previous figure shows the groups leading the change, Table 2 below shows the leading processes. In particular, it shows the difference between compositional and behavioral effects. Imagine a country where the monthly income (in thousands of FCFA[7]) is initially 142.5. This income is the weighted average of incomes across all of the economic classes making up the national population, from the richest to the poorest. The table shows, for years 1 and 2, respectively, the average income and the size of each class. Thus, in year 1, the poorest category represented 20% of the population and earned an average of 50,000 FCFA.

[7] FCFA is the Cameroonian national currency, 1US$ ≈ 500 FCFA.

In year 2, the table shows two possible scenarios of change. Both scenarios lead to the same aggregate result, a growth in national income from 142.5 to 152.9. Yet they are qualitatively different. Under scenario 1, the average incomes of different classes did not change; what changed was the relative size of the different classes, including the poor's share of the national population, which fell from 20% to 15%. In contrast, scenario 2 does not involve a compositional change; what changed instead were the average incomes of some groups. The richest economic group (350 on average in year 1) became even richer (400), which fully explains the rise in the country's average income.

The two scenarios in Table 2 show extreme, textbook, illustrations (100% composition versus 100% behavior). In practice, compositional and behavioral change often occur simultaneously. One may thus end up with a situation where composition explains 30% of the change while behavioral change accounts for the rest. The value of a demographic decomposition is precisely to quantify these relative contributions.

TIME 1

Economic Class	Average Income	% of Total Population
Highest	350	10%
Second highest	200	15%
Average	150	30%
Second lowest	90	25%
Lowest	50	20%
AVERAGE	**142.50**	

TIME 2: SCENARIO 1

Economic Class	Average Income	% of Total Population
Highest	350	10%
Second highest	200	20%
Average	150	35%
Second lowest	90	20%
Lowest	50	15%
AVERAGE	**153.93**	

100 % Compositional Change

TIME 2: SCENARIO 2

Economic Class	Average Income	% of Total Population
Highest	400	10%
Second highest	225	15%
Average	155	30%
Second lowest	90.7	25%
Lowest	50	20%
AVERAGE	**153.93**	

100 % Behavioral Change

Table 2. Compositional vs behavioral change: An extreme case.

II.1.4 Mathematical formulation

In this first case, we focus on a national average; Y is expressed as a weighted average (by w_j) of the values of individual subpopulations (y_j).

$$\boldsymbol{Y_t = \sum w_{jt} * y_{jt}} \qquad \text{[II.2]}$$

In this formula, a national change can be broken down into two components:

$$\boldsymbol{\Delta Y = \underbrace{\sum \bar{y}_j * \Delta w_j}_{\textit{Compositional effect}} + \underbrace{\sum \bar{w}_j * \Delta y_j}_{\textit{Behavioral effect}}} \qquad \text{[II.3]}$$

This basic decomposition identifies two sources of change. The first, the compositional effect, measures the change in the relative representation of the various subpopulations. This change in composition affects the national average through a mechanical change in the weights and the importance of different subpopulations.

The second source of change, behavior, is less mechanical. It shows a real change of mortality within one or more groups. If the mortality of a group increases, other things being equal, the national mortality will increase. What changes here is not the relative size of groups, but rather the mortality levels within some or all subpopulations.

II.1.5 Application

In practice, one implements a decomposition analysis in four main steps: defining the problem; calculating national averages and change therein; decomposing the total change; and presenting/ discussing the findings.

Defining the problem: Here, the researcher specifies the nature of the substantive (dependent) variable, the classification (independent) variable, and the period. In our example below, the substantive variable is child mortality, the classification variable is socioeconomic status, and the study period is 1991-2001.

Calculating the national averages: The national averages are calculated using formula #1, the year-specific information about group size (w_j), and the value of the dependent variable (y_j). These data should be available for first and the last year of the period at least. The statistical procedures to generate these results are a simple frequency analysis (for w_j) and an equally simple comparison of means (for y_j). These basic data can come directly from published reports or automatic compilations available online (e.g., www.Statcompiler.com).

Doing the actual decomposition: Using the available annual information, one can simply apply the formulation in equation 3. Given the repetitive nature of the calculations, we recommend the use of spreadsheet software like Excel but other software or personal programs can be used. Table 3 summarizes the basic data for the calculations.

	CAMEROON 1991		CAMEROON 2011			DECOMPOSITION		CONTRIBUTION BY SES GROUP	
Economic Class	Infant Mortality Rate	% of Total Population	Infant Mortality Rate	% of Total Population		Compositional effect	Behavioral Effect	Aggregate total	%
Highest	201	0.18	184	0.15		-6.92	-2.77	-9.69	37%
Second highest	171	0.23	144	0.20		-4.85	-5.79	-10.65	41%
Average	142	0.20	120	0.23		4.09	-4.77	-0.68	3%
Second lowest	120	0.20	90	0.22		2.14	-6.23	-4.09	16%
Lowest	82	0.19	72	0.21		1.17	-1.98	-0.81	3%
	AVERAGE IMR- 142.50		AVERAGE IMR- 117.68		Total Change in IMR 1991-2011 -25.9	-4.38	-21.54		
						17%	83%		
						Decline due to change in composition	Decline due to change in behavior		

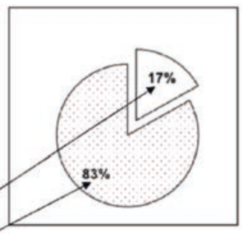


Table 3. Decomposing a change in child mortality, Cameroon 1991-2011.

Again, the first four columns include the basic information needed for the analysis, including the child mortality rate for each of the social classes and the percentage of children in these classes for the years 1991 and 2011. Using these data, one can easily calculate the national average for each year and the change over the study period, from 143.6 in 1991 to 117.7 in 2011. The difference here is -25.9.

The final step is to explain this decline in mortality. We do so by applying formula #3. The results in this case show that 17% of the decline stemmed from compositional change. Note, for example, how the percentage of children in poorer families dropped by 18 to 15%, while the percentage of children in the richest families rose from 19 to 21%. The remainder of the change (83%) reflects a behavioral effect. Notice in particular how mortality rates declined across all social classes, including the poorest.

Presenting the findings

For a scientific audience, the presentation can simply use a table like the one in Section II.2, with the discussion focusing on both the processes and groups driving the change.

Leading processes: The percentages at the bottom of the table indicate the contributions of the two competing processes (composition and behavior) to the total change in national mortality. In this case, these relative contributions are 17% for composition and 83% for behavior.

Leading groups. In addition to clarifying the contributions of various processes, a decomposition analysis reveals which groups drove the change. In our case, the contribution of the lower income group is -9.69, 37% of the total change. This total contribution reflected effects stemming from changes in group size (-6.92) and behavior (-2.77). The other income groups contributed 41%, 3%, 16%, and 3%, respectively.

In a decomposition analysis, the sum of these contributions is always 100%. However, individual contributions can be negative (less than 0%) or greater than 100%. A negative percentage indicates a contribution that goes in the opposite direction of the general change. For example, a group making a negative contribution to the national decline in mortality means that this group's effect worked to increase mortality, i.e., it worked against the prevailing trend. In contrast, a percentage greater than 100% indicates that group or process accounted for more than the total change observed at the national level. The national change would have been even greater if it had depended only on this group, and if it had not been offset by the opposite influences of other groups.

For a non-scientific audience, it might be useful to rely on graphs rather than a table. For instance, one might use pie charts or 100% (stacked area) histograms. Such diagrams offer nice summaries that clearly identify the dominant drivers of change.

Policy implications

The policies to recommend will depend on whether a social change reflects a compositional or behavioral effect. If mortality is driven by a compositional effect, the appropriate response is to target the vulnerable segments of the population. If, on the other hand, the change reflects behavioral effects, targeting becomes less appropriate, and one would consider broader-based interventions to all families and their children. Later, we will see how to refine policy recommendations with more detailed decompositions.

II.1.6 Application to the demographic dividend

Basic demographic decomposition can apply to the study of demographic dividend. Indeed, it is the logic behind the National Transfer Accounts (NTA) methodology often used in this field (Mason and Lee 2005). At the heart of this method is the simple observation that income, consumption, and savings vary with age (Figure 5). In particular, the income/consumption balance tends to be negative among younger and older populations but positive among the middle-aged adult population. Intuitively, therefore, the greater the share of adults in the national population, the higher the national income.

The parallel between NTA and demographic decomposition is obvious. One only needs to note here that 'age group' is the classification variable and income-consumption is the substantive variable (see Table 4 below). Thus, the composition effect is the demographic dividend, or at least the mechanical component of demographic dividend. Indeed, the application of decomposition could enhance the standard NTA analysis by (a) considering cases where consumption/income profiles do change with age, rather than assuming that they remain fixed; (b) quantifying the dividend both in absolute terms and also in relative terms by comparing it to changes induced by historical evolution in income and consumption; and (c) exploring the contribution of different age groups to the dividend.

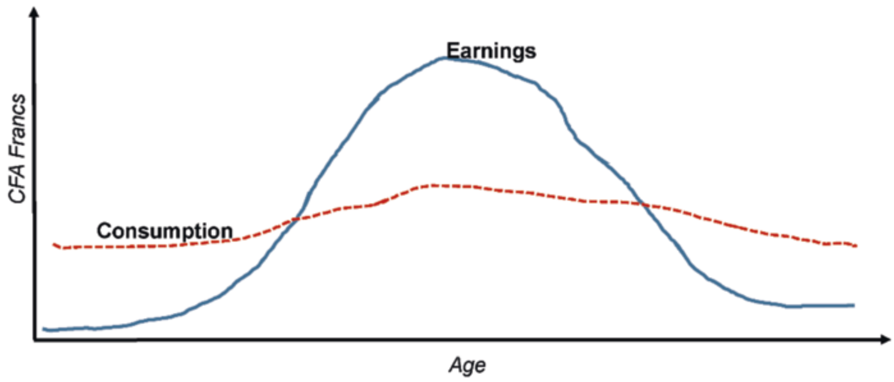


Figure 5. Income and consumption profiles by age. Hypothetical data.

	TIME 1		TIME 2		DECOMPOSITION		CONTRIBUTION BY AGE GROUPS	
Age Group	Net Revenue ($)	% of Total Population	Net Revenue ($)	% of Total Population	Compositional effect	Behavioral Effect	Aggregate total	%
0-14	-15	.30	-20	.20	1.75	-1.25	.50	4%
15-29	-3	.25	-5	.20	.20	-.45	-.25	-2%
20-49	35	.25	50	.30	2.13	4.13	6.25	48%
50-64	50	.15	75	.20	3.1	4.38	7.5	58%
65+	-10	.05	-15	.10	-.63	-.38	-1.00	-8%
	AVERAGE REVENUE $10.50		AVERAGE REVENUE $23.50		6.58	6.43		
					51%	49%		

Total Change in REVENUE $13.00

Demographic Dividend= Economic gain due to change in age composition

Table 4. Estimating the demographic dividend with a demographic decomposition approach.

II.2 Derived demographic decompositions

II.2.1 Decomposing a difference

Let us return to our case study on mortality. Imagine a researcher who is not interested in historical change but, rather, in the mortality difference between two countries or provinces. Fortunately, the same approach works, as long as the research has data on the size and mortality rate of each social class within each of the two provinces. The calculations proceed in the same way as previously described.

II.2.2 Decomposing inequality

In our mortality case study, the dependent variable was an average—specifically, the average level of child mortality in the country. There may be cases where a researcher is interested not in the average but in the mortality inequality across regions, economic classes, education levels, or racial groups, for instance. At the global level, a researcher may seek to account for the historical change in global income inequality, and she can use basic demographic decomposition to do so. In Firebaugh and Goesling (2014), for instance, formulate economic inequality as a function of differences in national populations (p) and per capita income (i). Taking the mean log deviation (MLD) as a measure of inequality, they express the global level of inequality as follows:

$$\boldsymbol{I}_t = \sum w_{jt} * ln(1/r_{jt}) \qquad \text{[II.4]}$$

In this context, the historical change in global inequality is:

$$\Delta \boldsymbol{I} = \underbrace{\sum(\bar{r}_j - \overline{\ln r_j})\Delta w_j}_{\textit{Compositional effect}} + \underbrace{\sum(\overline{w_j r_j} - \overline{w_j})\Delta ln(r_j)}_{\textit{Behavioral effect}} \qquad \text{[II.5]}$$

We use this approach to analyze the historical change in income inequality between African countries from 1985 to 2005 (Table 5). Between these two years, economic inequality, as measured by the Mean Log Deviation in GDP per capita, increased from 0.391 to 0.425, or by about 9%. We apply a decomposition analysis to identify the sources of this economic divergence, looking at the respective contributions of compositional and behavioral forces. The analysis can also show individual country contributions, highlighting the countries that made the greatest contributions to increasing inequality and those that instead worked to reduce it. The results of this analysis are in Table 5.

The first four columns in the table show national data on GDP per capita and total population for 1985 and 2005, respectively. These nominal data are transformed into relative values: raw population numbers, for instance, are turned into the share of the total African population living in a country as obtained by dividing the population of each country's by the total population of Africa during the indexed of year. Similarly, GDP is converted into relative GDP by dividing the country's GDP by the weighted average of the African GDP during that year. Countries where the relative GDP value is greater than 1 are richer than average; they are poorer than average if that value is below 1. Using these relative values, one can calculate inequality, and the values obtained here are 0.39 and 0.43, i.e., a nominal increase of about 0.04. This nominal change is then decomposed into a compositional effect (-18%) and a behavioral effect (118%).

The compositional effect suggests that, during the investigated period, African populations grew at different rates, and this altered the relative sizes of individual countries. The pattern of these changes contributed to reduce inequality, suggesting that, perhaps, the countries at economic extremes (the very rich or the very poor) grew more slowly than others. At any rate, this differential population growth helped reduce economic inequality. However, this influence was overshadowed by the changes in economic circumstances, with countries' GDPs becoming even more unequal in 2005 than they were in 1985. At this stage, it is not possible to say why these differences in economic performance or population growth occurred. As we will see later, more detailed decompositions can shed light on these questions. For now, this first analysis is a good start.

It is also useful to examine the individual contributions of each country to the change in inequality as shown in the last column of Table 5.

	PIB par habitant		Population totale		PIB relatifs		Parts de population		Niveau d'inegalite (MLD)		DECOMPOSITION DU CHANGT			
PAYS	1985	2005	1985	2005	1985	2005	1985	2005	1985	2005	Effet de composition	Effet de comportement	Contribut. nationales	
Congo, Dem. Rep.	715	273	32443229	58740547	0.34	0.12	0.064	0.069	0.07	0.15	0.010	0.0561	0.0657	183%
Guinee Eq	2227	24770	314190	608807	1.06	10.47	0.001	0.001	0.00	0.00	0.000	0.0079	0.0083	23%
Niger	696	584	6708883	13264190	0.33	0.25	0.013	0.016	0.01	0.02	0.004	0.0030	0.0067	19%
Madagascar	1044	882	9778464	17614261	0.50	0.37	0.019	0.021	0.01	0.02	0.002	0.0033	0.0051	14%
Egypt, Arab Rep.	2954	4319	50654901	77154409	1.40	1.83	0.100	0.091	-0.03	-0.05	-0.010	0.0152	0.0048	13%
Kenya	1291	1349	19673682	35598952	0.61	0.57	0.039	0.042	0.02	0.02	0.003	0.0012	0.0047	13%
Botswana	4620	12088	1173803	1835938	2.19	5.11	0.002	0.002	0.00	0.00	0.000	0.0049	0.0046	13%
Malawi	717	648	7264558	13226091	0.34	0.27	0.014	0.016	0.02	0.02	0.002	0.0023	0.0041	11%
Ethiopia	500	633	41049476	74660901	0.24	0.27	0.081	0.088	0.12	0.12	0.011	-0.0076	0.0038	10%
Cote d'Ivoire	2155	1560	10475579	19244866	1.02	0.66	0.021	0.023	0.00	0.01	0.002	0.0016	0.0037	10%
Liberia	1384	323	2214623	3334222	0.66	0.14	0.004	0.004	0.00	0.01	-0.001	0.0039	0.0031	9%
Tunisia	3905	6445	7260360.969	10029000	1.85	2.72	0.014	0.012	-0.01	-0.01	-0.004	0.0063	0.0026	7%
Zambia	1333	1127	6784944	11738432	0.63	0.48	0.013	0.014	0.01	0.01	0.001	0.0017	0.0022	6%
Togo	886	772	3344926	5992080	0.42	0.33	0.007	0.007	0.01	0.01	0.001	0.0011	0.0017	5%
Angola	3109	3611	9331250	16617589	1.48	1.53	0.018	0.020	-0.01	-0.01	0.001	0.0003	0.0016	4%
Benin	1240	1309	4122108	7867626	0.59	0.55	0.008	0.009	0.00	0.01	0.001	0.0002	0.0015	4%
Burundi	469	340	4884506	7378129	0.22	0.14	0.010	0.009	0.01	0.02	-0.002	0.0033	0.0015	4%
Mauritius	4184	9975	1016000	1243253	1.99	4.22	0.002	0.001	0.00	0.00	-0.001	0.0025	0.0014	4%
RCA	899	644	2678023	4191429	0.43	0.27	0.005	0.005	0.00	0.01	0.000	0.0015	0.0010	3%
Chad	1044	1468	5227487	10145609	0.50	0.62	0.010	0.012	0.01	0.01	0.002	-0.0011	0.0008	2%
Senegal	1463	1614	6513728	11281296	0.69	0.68	0.013	0.013	0.00	0.01	0.000	0.0001	0.0005	1%
Gambie	1141	1142	735039	1526138	0.54	0.48	0.001	0.002	0.00	0.00	0.000	0.0001	0.0005	1%
Namibia	4371	5361	1130855	2019677	2.08	2.27	0.002	0.002	0.00	0.00	0.000	0.0002	0.0004	1%
Guinea-Bissau	653	497	919005	1472626	0.31	0.21	0.002	0.002	0.00	0.00	0.000	0.0005	0.0004	1%
Swaziland	2519	4335	705657	1124410	1.20	1.83	0.001	0.001	0.00	0.00	0.000	0.0003	0.0002	1%
Mauritania	1599	1684	1715028	2963105	0.76	0.71	0.003	0.004	0.00	0.00	0.000	0.0001	0.0002	0%
Guinea	860	1056	5267169	9220768	0.41	0.45	0.010	0.011	0.01	0.01	0.001	-0.0005	0.0001	0%
Cape Verde	1589	2695	318417	477438	0.75	1.14	0.001	0.001	0.00	0.00	0.000	0.0000	-0.0001	0%
Cameroon	2716	1959	10514988	17795149	1.29	0.83	0.021	0.021	-0.01	0.00	0.000	-0.0005	-0.0003	-1%
Sierra Leone	720	640	3631130	5106977	0.34	0.27	0.007	0.006	0.01	0.01	-0.002	0.0011	-0.0007	-2%
Burkina Faso	711	1026	7709074	13933363	0.34	0.43	0.015	0.016	0.02	0.01	0.002	-0.0024	-0.0008	-2%
Congo, Rep.	4033	3497	2116659	3416654	1.91	1.48	0.004	0.004	0.00	0.00	0.000	-0.0007	-0.0009	-3%
Lesotho	835	1266	1472306	1980831	0.40	0.54	0.003	0.002	0.00	0.00	-0.001	-0.0004	-0.0011	-3%
Rwanda	768	793	6111361	8992140	0.36	0.34	0.012	0.011	0.01	0.01	-0.002	0.0006	-0.0014	-4%
Mali	707	1004	6793924	11611090	0.34	0.42	0.013	0.014	0.01	0.01	0.000	-0.0020	-0.0016	-4%
Uganda	525	901	14795432	28699255	0.25	0.38	0.029	0.034	0.04	0.03	0.007	-0.0091	-0.0022	-6%
Nigeria	1306	1731	81598130	141356083	0.62	0.73	0.162	0.167	0.08	0.05	0.006	-0.0088	-0.0030	-8%
Gabon	16557	13029	791848	1369229	7.86	5.51	0.002	0.002	0.00	0.00	0.000	-0.0032	-0.0030	-8%
Ghana	816	1193	13005766	21915168	0.39	0.50	0.026	0.026	0.02	0.02	0.000	-0.0038	-0.0036	-10%
Morocco	2485	3589	21779134	30142708.8	1.18	1.52	0.043	0.036	-0.01	-0.01	-0.008	0.0033	-0.0046	-13%
Sudan	889	1601	24051873	38698472	0.42	0.68	0.048	0.046	0.04	0.02	-0.002	-0.0100	-0.0122	-34%
Algeria	6847	7176	22097343	32854159	3.25	3.03	0.044	0.039	-0.05	-0.04	-0.010	-0.0061	-0.0160	-44%
Mozambique	312	677	13324105	20532675	0.15	0.29	0.026	0.024	0.05	0.03	-0.004	-0.0131	-0.0169	-47%
South Africa	8100	8504	31307880	46892428	3.84	3.59	0.062	0.055	-0.08	-0.07	-0.016	-0.0108	-0.0266	-74%
MOYENNE OU TOTAL	2106.6	2365.8	504,806,844	845,868,171					0.39	0.43	-0.006	0.0423		
										0.04	-18%	118%		

Table 5. Decomposition of change in economic inequality between African countries (1985-2005).

II.2.3 Ordinal decomposition

The ordinal decomposition is very similar to the demographic decomposition, except that the classification variable is ordinal. We illustrate this variant with a study of total fertility rate (TFR). TFR values, roughly the average number of births per woman in a country, are the sum of all of the fertility rates found in all of the relevant age groups—in this case, the ages between 15 and 49 during which women are assumed to be of reproductive age:

$$Y_t = \Sigma\, y_{at} \tag{7}$$

Suppose a country's TFR were to fall over time. A researcher may wish to understand why the decline occurred, specifically whether the decline occurred among all groups (a quantum effect) or whether women mostly altered the ages at which they bear children, perhaps postponing childbearing (a tempo effect).

Using temporal decomposition, the researcher can express age-specific fertility rates relative to the fertility observed in the oldest age group . For example, the fertility among the 15-19 year olds will be expressed in terms of the average fertility for all women aged 20 to 49 ($y_{15-19} = r_{15-19} * y_{20-49}$);

$$y_a / y_{a+} = r_a \rightarrow y_a = r_a * y_{a+}$$

The TFR can be re-expressed as follows:

$$Y_t = \Sigma\, r_{at} * y_{a+t} \tag{II.6}$$

This new formulation helps decompose the change in the TFR into two terms that reflect the quantum and tempo, respectively.

$$\Delta \boldsymbol{Y} = \sum \overline{\boldsymbol{y}}_{\boldsymbol{a+}} * \Delta \boldsymbol{r_a} + \sum \overline{\boldsymbol{r}}_{\boldsymbol{a}} * \Delta \boldsymbol{y_{a+}} \tag{II.7}$$

Illustration

A paper by Ouedraogo (2012) (data not shown here) looked at patterns of change in age-specific fertility curves in Cameroon from 1991, 1998, and 2004. An examination of these curves shows some parallelism. If the change occurs at the same rate across all age groups, the fertility curves should be perfectly parallel. Otherwise, it becomes difficult to succinctly describe the differences between the various curves. For a researcher wishing to assess how much the decline in births occurs disproportionately among younger age groups, the temporal decomposition is a good tool.

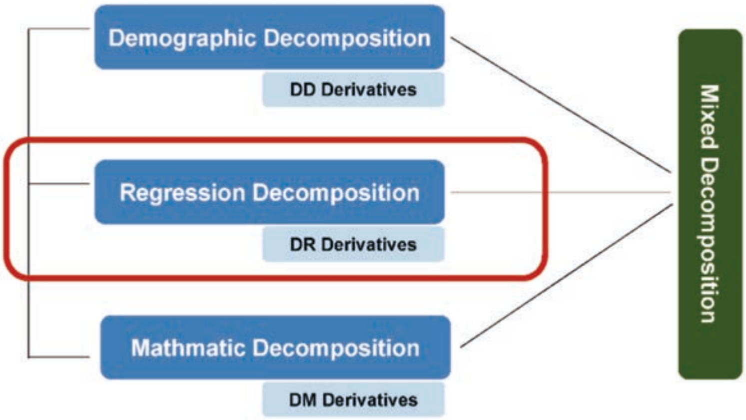
Calculations here show a decline driven by a quantum effect (90%): the decline affected all age groups, but, as the chart shows, it was larger among younger women, suggesting some postponement of births (10%). Note that this analysis could be repeated using SES, for example, as a classification variable to study the dispersion of reproduction across social classes.

II.2.4 Nested decomposition

This extension offers more detail by expressing demographic change (the composition component) in terms of its constitutive elementary processes such as fertility, mortality, and migration. To apply this decomposition, one obviously needs much more data than for a simple decomposition.

In addition, one can also envision an extension based on nesting categories of two classification variables. For instance, the researcher might be interested in the intersection of age and SES and would group people in terms of these two variables. Thus, instead of distinguishing only between the poor, middle class, and rich, one would further distinguish between the poor who are young, middle-aged, or old. Thus, if each of the two classification variables had three categories, the researcher would end up with a categorization of her population including six categories.

Chapter III
Regression decomposition



III.1 Simple regression decomposition

III.1.1 Problem type

The main difference between demographic decomposition and simple regression decomposition is in the kind of independent variable involved. It is nominal in a demographic decomposition but quantitative here. Moreover, the independent variable is linked to the dependent variable through a statistical relationship (derived from regression analysis) rather than an arithmetic one. The equation has a regression coefficient (the effect of the independent variable on Y) and an intercept. The error term is omitted here for now.

III.1.2 Formulation

The generic form is Y = f (α, β, X) but a more explicit formulation is as follows:

$$y_t = \alpha_t + \beta_t X_t \qquad \text{[III.1]}$$

In this case, the decomposition seeks to explain the change in the dependent variable based on the change in the various parameters of the regression equation. This change is expressed as follows:

$$\Delta Y = \underbrace{\Delta\alpha}_{\text{Change in baseline}} + \underbrace{\bar{\beta}\Delta X}_{\text{Change in effect magnitude of X level}} + \underbrace{\bar{X}\Delta\beta}_{\text{Change in effect of X}} \qquad \text{[III.2]}$$

Once again, the same procedure can apply to both cross-sectional analysis (the difference between two groups in a given year) and longitudinal analysis (the change experienced by one group between years). The approach is the same; only the interpretations differ.

Illustration

A classic realm of application for regression decomposition is the study of income/wage differentials (say, between men and women) and the extent to which they reflect discrimination in the labor market. Differences could stem from discrimination (different returns to education for men and women) but it is also possible that men and women enter the workforce with different education levels. One must therefore estimate how the wage differentials reflect discrimination rather than differences in human capital. These two possibilities are explored below. The formal analysis (III.3) consists of writing the earning equations for males (h) and females (f) and then taking the difference between these two equations.

$$Y_h = \alpha_h + \beta_h * X_h$$
$$Y_f = \alpha_f + \beta_f * X_f$$

$$\Delta Y = \Delta\alpha + \bar{\beta}\Delta X + \bar{X}\Delta\beta \qquad \text{[III.3]}$$

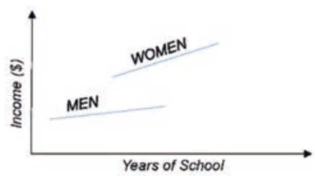
III.1.3 Application

Table 6 illustrates with a numerical example showing earnings by level of education. In this case, the average return on a year of education is 4,000 FCFA for men and 5,000 FCFA for women. The base salaries are 20,000 and 25,000 for men and women, respectively. With the average education levels being 10 and 12 years for men and women, respectively, the average salary of a man is 60,000, while the average woman would earn 85,000 FCFA, yielding an average salary difference of 25,000 between women and men. A decomposition analysis will show that this pay gap reflects a complex mix of forces:

- 20% of the gap comes from the gender difference in base salaries (Δα),
- 36% of the gap comes from the gender difference in levels of education, and
- 44% of the gap comes from the gender difference in returns to education.

In other words, even if one achieved parity in education, only 36% of the current wage inequality between men and women would disappear. The rest of the inequality that comes from discriminatory inequalities in basic salary and returns to schooling would remain.

	Expression
Regression for women	$Y_w = \alpha_w + \beta_w * X_w$
Regression for men	$Y_m = \alpha_m + \beta_m * X_m$
Regression decomposition	$\Delta Y = \Delta\alpha + (\bar{\beta} * \Delta X) + (\bar{X} * \Delta\beta)$



Application Example

Components	Women	Men
α	25,000	20,000
β	5,000	4,000
$\bar{X}$ (Average years of schooling)	12	10

1. Average salary for women = 25,000 + (5,000*12)= 85,000
2. Average salary for men = 20,000 + (4,000*10)= 60,000
3. Decomposition of salary difference (25,000)
4. ΔY = (5,000) + (4,500*2) + (11*1000)
   = (5,000) + 9,000 + 11,000
   = 20% + 36% + 44%

Wage gap due to differences in **baseline salaries.**

% of wage gap due to differences in **average years of schooling.**

% of wage gap due to differences **returns to schooling.**

Table 6. Regression decomposition regression for analyzing wage differentials between men and women.

III.2 Other regression decompositions

III.2.1 Curvilinear regression

Most processes of interest to social scientists do not fit the linear bivariate pattern posited in the previous section. A curvilinear, rather than a linear, specification is often more realistic. Fortunately, regression decomposition analysis easily extends to these situations. For instance, earnings can be written as a quadratic (rather than linear) function of education:

$$Y = \alpha + \beta_1 X + \beta_2 X^2 \qquad \text{(III.4)}$$

The decomposition analysis is as follows:

$$\Delta Y = \Delta\alpha + (\bar{X} * \Delta\beta_1) + (\overline{\beta_1} * \Delta X) + (\overline{X^2} * \Delta\beta_2) + (\overline{\beta_2} * 2\Delta X)$$
$$\Delta Y = \Delta\alpha + (\bar{X} * \Delta\beta_1) + (\overline{X^2} * \Delta\beta_2) + ((\overline{\beta_1} + 2\overline{\beta_2}) * \Delta X) \qquad \text{(III.5)}$$

To give an example, one can return to our study of wage differences between men and women, only this time, we set the effects of education on income to be curvilinear.

III.2.2 Multivariate regression

In other cases, the researcher wishes to estimate the contributions of multiple independent variables. This is feasible. The decomposition will simply include more than a single factor (and the additional factors will be other independent variable rather than the square term of the main independent variable). In the simplest of these situations, the researcher has two variables, rather than a single one. For instance, she might wish to examine how income is affected by both the level of education and the number of years of experience:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 \qquad \text{(III.6)}$$

The decomposition then writes as follows:

$$\Delta Y = \Delta\alpha + (\overline{X_1} * \Delta\beta_1) + (\overline{\beta_1} * \Delta X_1) + (\overline{X_2} * \Delta\beta_2) + (\overline{\beta_2} * \Delta X_2) \qquad \text{(III.7)}$$

III.2.3 Multilevel regression

This type of regression involves factors at two or more levels, e.g., individual and community levels. Research on schooling might for instance examine how the academic performance of students depends on their individual characteristics (level 1) but also the characteristics of the schools (level 2). The equations to estimate this model are as follows:

Level 1: $Y_{jk} = \beta_{0k} + \beta_{1k} * x_{jk} + r_{jk}$

Level 2: $\beta_{0k} = \gamma_{00} + \gamma_{01} Z_k + \mu_{0k}$

$B_{1k} = \gamma_{10} + \gamma_{11} Z_k + \mu_{1k}$

An integration of the values from level 2 into the level 1 equation yields a mixed equation that expresses individual performance based on the individual characteristics and those of the group at a particular time and their interactions:

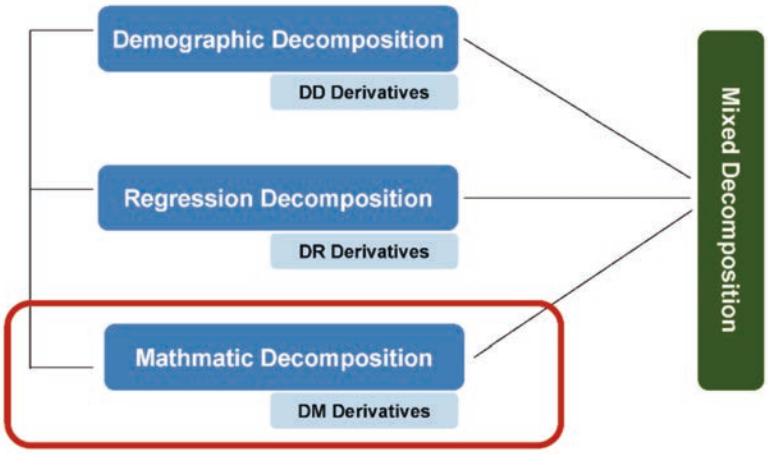
$$Y_{jk} = \gamma_{00} + \gamma_{10} x_{jk} + \gamma_{01} Z_k + \gamma_{11} Z_k x_{jk} + \mu_{0k} + \mu_{1k} x_{jk} + r_{jk} \qquad \text{(III.8)}$$

To investigate the change of Y_{jk} over time, one simply needs to differentiate the above formula and incorporate it into the second term of the equation [1].

III.3 Application to the demographic dividend

The study of demographic dividends is about how changes in national birth rates affects national levels of schooling. In essence, the researcher must investigate a relationship between two macro-outcomes. For instance, imagine a country where average family size declines from 6 to 4 children between 1990 and 2010 at the same time as the percentage of 6-10 year olds enrolled in school rises from 60 to 75%. A question one might ask is whether or how much these improvements in education came from the observed reduction in average family size. It is possible that the smaller family sizes played a role, but other factors could have also contributed. A macro-level correlation cannot apply because we are dealing with a single country with only two points in time. Furthermore, even if a researcher could implement, the macro-regression is open to criticism for lacking in detail and rigor. Instead, one can use decomposition to look at how the percentages of children living with different sibsizes has changed but also how the effects of sibsize itself has changed. Through decomposition analysis, one can estimate the relative contributions of several factors, but the researcher will need reliable information on how the statistical relationship between education and family size changed over time. In a recent study (Eloundou-Enyegue and Giroux 2012), we show how to combine the micro relation between education and fertility with information about fertility change to estimate the implications of demographic change on national school enrollment.

Chapter IV
Mathematical decomposition



IV.1 Simple mathematical decomposition

We have so far covered demographic and regression decompositions. In demographic decomposition, the dependent outcome is an average, and the function linking it to the independent variable is simply a weighing function. In regression decomposition, the dependent variable is a relationship linked to the independent variables through a statistical relationship, estimated via regression analysis. Next, we want to explore yet another case, where there is an exact mathematical relationship between the independent variables and the outcome. For example, the average per-capita income in a country equals the total income divided by the total population. A change in average income can only result from a change in total income (numerator) or in the size of the population (denominator).

IV.1.1 Problem type

This type of analysis applies to processes that involve a mathematical relationship between two or more social variables. GDP per capita is one example. It combines an economic component (GDP) and a demographic component (number of inhabitants), and it can be broken down into simple terms that show how the GDP per capita changes as either one of these two components varies. Other similar variables measure individual welfare by relating available resources to the population served.

IV.1.2 Mathematical formulation

For this example, we will use a slightly more complex example than GDP per capita. Consider the public education spending per child (r). This expenditure is positively related to the total level of national resources (g) and the percentage of resources allocated to education (k). On the other hand, it is inversely proportional to the number of school-aged children in the country (p).

$$r = {}^{gk}/_{p} \qquad \text{(IV.1)}$$

In this case, the historical change in this expense can be broken down as follows:

$$\Delta r \cong -[\overline{(kp/p^2)} * \Delta p] \;+ [\overline{(k/p)} * \Delta g] + [\overline{(g/p)} * \Delta k] \qquad \text{(IV.2)}$$

Effect of change in population | *Effect of change in national income* | *Effect of change in share of budget to education*

IV.1.3 Application

Using World Bank statistics (2014), Eloundou, Tenikue, and Ryu (2014) decompose the changes in public expenditure per child among Southern Africa countries between 1990 and 2010 and compare the results to the results of South Korea between 1975 and 1995. The results show that, over their respective study periods, South Korea's economy cumulatively grew by 250%, much faster than the growth observed in Southern African nations (3-94%). Its ratio of youth to adults also fell faster, from 0.65 to 0.23, compared to South Africa's decline from 0.67 to 0.46. Despite such demographic and economic differences, the relative contributions of demographic change to the gains in r values were similar (60% for Korea versus 70 % and 61% in South Africa and Swaziland, respectively). The study also noted an often-overlooked fact: many African countries allocate a larger portion of their national budget to education (e.g., 10% in Lesotho and almost 6% in Botswana versus 2% in Korea). It is therefore unsurprising that budget decisions made larger contributions to improving the r values in Lesotho (32%) and Botswana (24 %) than in South Korea (12%).

	GDP per adult		Share of budget spent on education		Children per adult		Public education spending per child			Share of the total change in r associated with changes in		
	g		k		p*		r		Δ r	Age Dependency	Income	Budget Commitment
	1990	2010	1990	2010	1990	2010	1990	2010				
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Botswana	6257	9890	0.057	0.078	0.85	0.55	421.2	1409.4	988.20	41.7%	34.4%	24.0%
Mauritius	4581	8889	0.032	0.037	0.44	0.30	330.8	1093.5	762.70	37.1%	50.6%	12.2%
South Africa	8390	8906	0.053	0.060	0.67	0.46	659.7	1166.8	507.10	69.9%	9.9%	20.3%
Swaziland	4099	4229	0.055	0.069	0.98	0.67	230.0	436.2	206.20	61.1%	4.7%	34.2%
Lesotho	1074	1514	0.101	0.130	0.79	0.65	137.3	303.2	165.90	26.0%	42.3%	31.8%
	1975	*1995*	*1975*	*1995*	*1975*	*1995*	*1975*	*1995*	*1975-1995*			
Korea	4,885	17,254	0.021	0.032	0.654	0.230	$158.5	$2,415.9	$2,257.4	60%	28%	12%

Table 7. Mathematical decomposition of trends in public spending per child (South Korea versus the 'vanguard' countries of Africa).

IV.2 Derived mathematical decomposition

IV.2.1 Extended mathematical chain

Let us return to the mathematical decomposition of GDP per capita as introduced in section IV.1.1. This first decomposition usefully describes the contributions two components (population and GDP). However, is not very informative because its two components do not refer to key decision variables but also because these variables themselves need fuller exploration. With only a light transformation of these initial mathematical expressions, we can get a formula that is slightly longer but conceptually richer. Specifically, the initial formula of per capita GDP

$Y = {}^{G}/_{P}$ was differentiated as

$$\Delta Y = \left(\bar{G} * \Delta(1/P)\right) + (\left(\overline{1/P}\right) * \Delta G) \tag{IV.3}$$

which can be rewritten more interestingly as

$$Y = \frac{G}{P} = \frac{G}{A} * \frac{A}{P} = \pi * \alpha \tag{IV.4}$$

where G and P represent mean the national income and the total national population, respectively. The new term introduced is A, the active (adult) population in the country. In this new formula, G/A (or) refers to adult productivity, and it is a conceptually interesting variable. It is frequently cited in economic growth theories and may be shaped through specific policies to raise productivity through education, research, and technological development. Also, the new term A/P (or α) refers to the population's age structure—specifically, the ratio of adults to the total population, a core variable in demographic dividend theory. Thus, our analyst now has two very interesting variables (and α), and she can break down the changes in per capita GDP in terms of these two variables.

$$\Delta Y = (\bar{\pi} * \Delta\alpha) + (\bar{\alpha} * \Delta\pi) \tag{IV.5}$$

Obviously, this new expression can itself expand further to yield a more detailed decomposition, with better potential to inform policy decision-making. For example, adult productivity can be split into two components: the adult unemployment rate and the productivity of adult workers:

$$Y = \frac{G}{P} = \frac{G}{E} * \frac{E}{A} * \frac{A}{P} = \rho * \varepsilon * \alpha \tag{IV.6}$$

And its historical change can be represented as follows:

$$\Delta Y = (\bar{\varepsilon}\bar{\rho} * \Delta\alpha) + (\bar{\rho}\,\bar{\alpha} * \Delta\epsilon) + (\bar{\varepsilon}\,\bar{\alpha} * \Delta\rho) \tag{IV.7}$$

Similarly, the adult ratio of the total population can be expressed as a function of the youth and the elderly populations, respectively. How far the researcher extends this development depends on how much detail she needs, and on data availability, the policy relevance and the tractability of the added terms.

Students of economic growth can apply the same logic in decomposing popular formulations of economic performance such as the Cobb-Douglas function, which expresses growth based on physical capital (K), human capital (h), employment (L), and total factor productivity (A):

$$Y = AK^{\alpha}(hL)^{\beta} \tag{IV.8}$$

From this formula, one can decompose economic growth as

$$\Delta y_t \cong \Delta A_t + \alpha\Delta k_t + (1-\alpha)\Delta h_t + \Delta l_t + \Delta w_t \tag{IV.9}$$

where y_t is the GDP per capita (Y/P)

k is the stock of physical capital per employee (K/L)

h is human capital

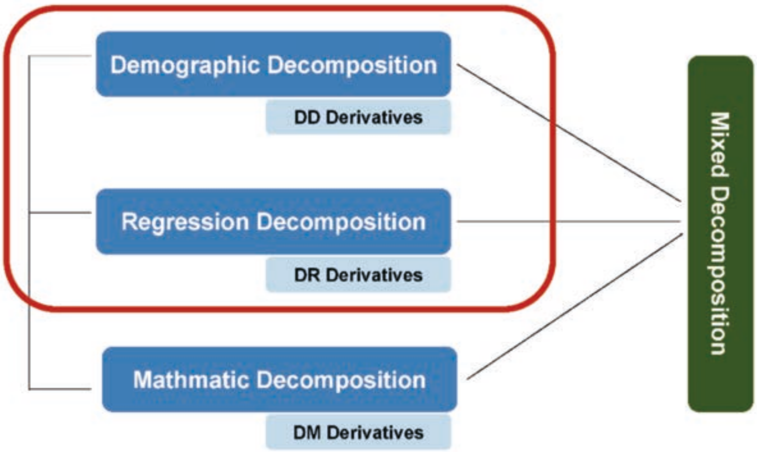
l is the employment rate (L/W)

w is the population's age structure (W/P)

A is the total factor productivity (TFP)

Clearly, our analyst has several options to decompose economic growth and education. As long as these decompositions include the effects of population age structure, then they help assess the magnitude of the demographic dividend.

Chapter V
Combination of demographic and regression decompositions



Elementary decompositions have the merit of being simple, but they lack in detail. This problem is solved with mixed decompositions combining two or more elementary forms of decomposition. One can combine demographic and regression decompositions to get a more detailed form as described in this chapter. Starting from a simple demographic decomposition, we will see how to expand each of the main components, starting with the behavioral effect (Δy_j) and then the composition effect (Δw_j).

V.1 Extension of behavior effect

V.1.1 General presentation

Let us start with the basic formula (3), which expresses a change in any national outcome in terms of the composition and behavior of various groups:

$$\Delta \boldsymbol{Y} = \underbrace{\left[\sum \bar{y}_j * \Delta w_j\right]}_{\textit{Compositional effect}} + \underbrace{\left[\sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} * \Delta \boldsymbol{y}_{\boldsymbol{j}}\right]}_{\textit{Behavioral effect}}$$

This formula's extension can express the behavior of any given group (y_j) as a function of one or more other variables. In a simple regression analysis,

$$y_j = \alpha + \beta x_j + \mu^j \quad \text{where,}$$

- α (intercept) is the baseline performance, when x = 0,
- β is the increased mortality associated with a unit increase in X, and
- μ_j is the error term (the over-performance or under-performance of the group or the residual effect of factors other than X, that are not considered in the analysis).

The change in the value of y_j between two periods is thus:

$$\Delta y_j = \Delta\alpha + \bar{\beta}\Delta x_j + \bar{x}_j\Delta\beta + \Delta\mu_j \qquad [V.1]$$

By inserting [V.1] in the basic equation (3), we get

$$\Delta \boldsymbol{Y} = \underbrace{\underbrace{\sum \bar{\boldsymbol{y}}_j \Delta \boldsymbol{w}_j}_{A}}_{\textit{Compositional Effect}} + \underbrace{\underbrace{\sum \bar{\boldsymbol{w}}_j \Delta\boldsymbol{\alpha}}_{B1} + \underbrace{\sum \bar{\boldsymbol{w}}_j\, \bar{\boldsymbol{\beta}}\Delta \boldsymbol{x}_j}_{B2} + \underbrace{\sum \bar{\boldsymbol{w}}_j\, \bar{\boldsymbol{x}}_j \Delta\boldsymbol{\beta}}_{B3} + \underbrace{\sum \bar{\boldsymbol{w}}_j\, \Delta\boldsymbol{\mu}_j}_{B4}}_{\textit{Behavioral effect}} \qquad [V.2]$$

- A (the composition effect) remains unchanged from the previous situation.
- B (the behavioral effect) is now divided into four sub-components that respectively reflect changes in
  - baseline performance (B1),
  - the level of the independent variable (B2),
  - the effect of the independent variable (B3) , and
  - the residual effect (B4).

V.1.2 Illustration

To illustrate this mixed decomposition, we can return to our first example in which we used simple demographic decomposition to study changes in infant mortality between 1990 and 2011 (Table 2). To recall, this simple decomposition showed a nearly 26-point decline in infant mortality, which reflected a mix of compositional change (17%) and behavioral change (83%). We can refine this analysis by considering that mortality varies with one's education level. We can thus express the mortality within each group as a function of the group's average education level. If one can estimate this relationship, and get a reliable estimate for the values of α (baseline mortality), β (the effect of education on mortality), and μ (the residual term), it becomes possible to refine our initial decomposition and get more detail about the forces driving change at the national level. This decomposition has the potential to reveal, in greater detail, the policy areas that were most influential in driving the change. Thus, in equation V.2,

- B1 (the baseline performance) reflects the improvement in the public sanitation and public health conditions that raise the minimum health standard of the population, regardless of education level;
- B2 measures the health improvements associated with the gains in the national level of education, assuming that the payoffs to education remain the same;
- B3 measures the improvement in the educational effects on health; and
- B4 measures the residual effect of other variables not considered.

Although this first example limits itself to a linear bivariate regression model (a single independent variable modeled linearly), one can easily imagine how this analysis can extend to cases where multiple independent variables or curvilinear influences are considered. The only difference is that the corresponding equations become longer and longer!

V.1.3 Comparison with national transfer accounts

The method of National Transfer Accounts is a core method for studying demographic dividends (Mason and Lee 2005). It builds on the simple idea that economic behavior (income, savings, transfers) varies systematically with age, even if the exact age profiles of economic behavior vary across countries (Figure 5). One can use a country's economic profile to estimate the balance between income and consumption at each age and combine the data for all age groups to calculate a cumulative balance for the entire country. Assuming a constant age pattern of economic behavior, any change in a population's age structure automatically changes the national balance between income and consumption.

However, it is not realistic to assume a constant age profile of economic behavior, especially during a demographic transition, given the large changes during a life course (e.g., age at marriage, duration of schooling, and the time spent bearing children versus participating in the labor force). It is therefore useful to consider situations where a country experiences changes in both the age structure and the consumption profile. Such situations are easily managed in a decomposition framework, and a full NTA approach can be seen as an example of mixed decomposition (Eloundou, Tenikue, Giroux 2014). Specifically, the national surplus in a given year is the weighted average of the specific surpluses for each age group.

$$\boldsymbol{S_t} = \sum \boldsymbol{w_{jt} s_{jt}} \qquad \text{[V.3]}$$

Change is thus expressed as

$$\Delta \boldsymbol{S} = \sum \bar{\boldsymbol{s}}_{\boldsymbol{j}} \Delta \boldsymbol{w_j} + \sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} \Delta \boldsymbol{s_j} \qquad \text{[V.4]}$$

Now, we can additionally consider, as indicated above, that consumption/income profiles can change over time, perhaps even in response to a change in the age structure. It is therefore useful to explore the additional possibility that a change in the age structure can also have an effect on behavior. If the analyst can obtain, through regression analysis, a reliable estimate of the effect of the age structure of (J) on the economic behavior of each age group [Equation V.5 below], then the change in the national surplus can be estimated through a mixed decomposition.

$$S_{JT} = a_{jt} + b_{jt}J_t + e_{jt} \qquad \text{[V.5]}$$

$$\Delta \boldsymbol{s_j} = \Delta \boldsymbol{a_j} + \bar{\boldsymbol{J}} \Delta \boldsymbol{b_j} + \bar{\boldsymbol{b}} \Delta \boldsymbol{J} + \Delta \boldsymbol{e_j} \qquad \text{[V.6]}$$

$$\Delta \boldsymbol{S} = \underbrace{\sum \bar{\boldsymbol{s}}_{\boldsymbol{j}} \Delta \boldsymbol{w_j}}_{m\text{-dividend}} + \sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} \Delta \boldsymbol{a_j} + \sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} \bar{\boldsymbol{J}} \Delta \boldsymbol{b_j} + \underbrace{\sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} \bar{\boldsymbol{b}} \Delta \boldsymbol{J}}_{s\text{-dividend}} + \sum \bar{\boldsymbol{w}}_{\boldsymbol{j}} \Delta \boldsymbol{e_j} \qquad \text{[V.7]}$$

In this last equation, the change in the age structure has both a mechanical effect (the first term) and a substantive effect on behavior (the penultimate term). The question, of course, is whether one can get a reliable estimate for the coefficient b—the effect of age structure on economic behavior.

V.2 Extension of composition effect

Just as with the behavior effect, the composition effect can also be disaggregated. The disaggregation can focus on primary demographic groups, on the age structure of the sub-populations, or on the demographic processes shaping the size of the groups (fertility, mortality and migration).

(A) Extension according to primary groups

In this case, the size of the study group is expressed as a function of the size of a primary group that generates the members of the group being studied. As one example, the number of children in poor families (w_j) might be expressed as a function of the number of poor families (n_j) and the relative fertility of poor families (f_j), i.e., the fertility of the poor compared to the national average.

$$w_j = n_j * f_j \qquad [V.8]$$

The change in the percentage of children from each class will depend on the change in (1) the proportion of families that belong to that social class and (2) the relative fertility of those families.

$$\Delta w_j = \bar{f}\Delta n_j + \bar{n}\Delta f_j \qquad [V.9]$$

One can thus insert [V.9] in [3], yielding

$$\Delta \boldsymbol{Y} = \underbrace{\underbrace{\left[\sum \bar{\boldsymbol{y}}_j * \bar{\boldsymbol{f}}\Delta \boldsymbol{n}_j\right]}_{A1} + \underbrace{\left[\sum \bar{\boldsymbol{y}}_j * \bar{\boldsymbol{n}}\,\Delta \boldsymbol{f}_j\right]}_{A2}}_{\boxed{\textit{Compositional effect}}} + \underbrace{\left[\sum \bar{\boldsymbol{w}}_j * \Delta \boldsymbol{y}_j\right]}_{\substack{B \\ \boxed{\textit{Behavioral effect}}}} \qquad [V.10]$$

In this equation, A1 represents the change in the percentage of poor families in a society, and A2 represents the change in the relative fertility of the various social classes. These two variables are conceptually richer, more informative, and more policy-relevant. The first (A1) might guide policies to stimulate growth and reduce poverty. The second might be related to the ability of family planning programs to reduce fertility inequality, especially if that inequality stems from differential access to family planning.

IV.3 Double extension

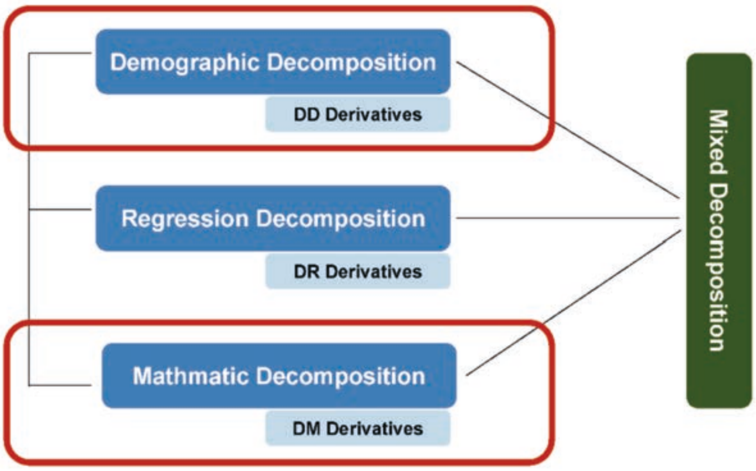
Finally, one can merge equations [V10] and [3] and get a detailed system that refines the analysis of both the compositional and behavioral sides. The result of this combination (below) yields an even more detailed expression where

- A1 = change in the distribution of mothers by social class
- A2 = change in the relative fertility of mothers
- B1 = change in baseline health
- B2 = change in income levels
- B3 = change in the health benefits of income
- B4 = residual effect of the factors omitted from the regression equation.

$$\Delta Y = \underbrace{\underbrace{\left[\sum \bar{w}_j\,\Delta\alpha\right]}_{\boxed{B1}} + \underbrace{\left[\sum \bar{w}_j\,\bar{\beta}_j\Delta x_j\right]}_{\boxed{B2}} + \underbrace{\left[\sum \bar{w}_j\,\bar{x}_j\Delta\beta\right]}_{\boxed{B3}} + \underbrace{\left[\sum \bar{w}_j\,\Delta\mu_j\right]}_{\boxed{B4}}} + \underbrace{\underbrace{\left[\sum \bar{y}_j\bar{f}_j\,\Delta n_j\right]}_{\boxed{A1}} + \underbrace{\left[\sum \bar{y}_j\bar{n}_j\Delta f_j\right]}_{\boxed{A2}}}$$

Again, both of the primary components of a basic demographic decomposition (compositional or behavioral) can be expanded for greater detail. With this more detailed accounting, a planner can get more nuanced insights into policy priorities.

Chapter VI
Combination of demographic and mathematical decompositions



This variant also begins with a simple demographic decomposition, but the expansion of the behavioral component uses an exact mathematical (rather than statistical) expression. To illustrate, let us consider a study of GDP change in Africa. The region's GDP per capita is a weighted average across all African countries, and the decomposition of its change can be expressed as usual.

$$\Delta \boldsymbol{Y} = \underbrace{\left[\sum \bar{y}_j * \Delta w_j\right]}_{\textit{Compositional effect}} + \underbrace{\left[\sum \bar{\boldsymbol{w}}_j * \Delta \boldsymbol{y}_j\right]}_{\textit{Behavioral effect}}$$

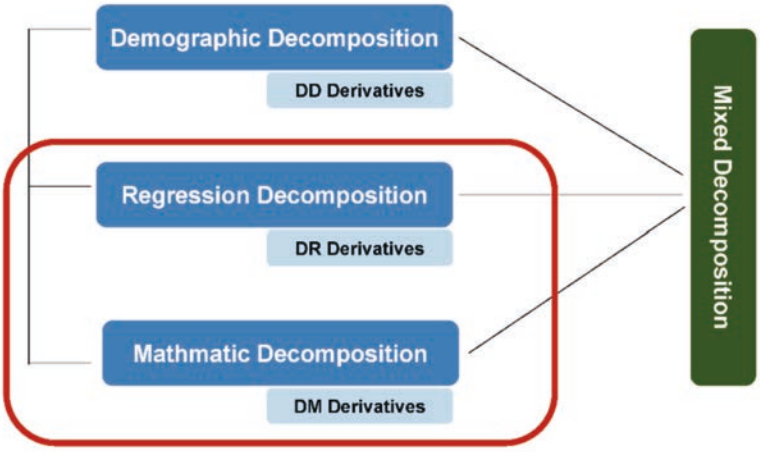
Next, we now express the countries' GDPs as a function of adult productivity (π) and age structure (α). The change of individual country GDP is as follows:

$$\Delta y_j = \left(\bar{\pi}_j * \Delta \alpha_j\right) + \left(\bar{\alpha}_j * \Delta \pi_j\right) \qquad \text{[VI.1]}$$

By inserting VI.1 into the basic equation of the demographic decomposition (II.3), the change in the average GDP of Africa becomes

$$\Delta \boldsymbol{Y} = \underbrace{\left[\sum \bar{y}_j \Delta w_j\right]}_{\textit{Effect of population size}} + \underbrace{\left[\sum \bar{\boldsymbol{w}}_j \bar{\boldsymbol{\pi}}_j \Delta \boldsymbol{\alpha}_j\right]]}_{\textit{Effect of age structure}} + \underbrace{\left[\sum \bar{\boldsymbol{w}}_j \bar{\boldsymbol{\alpha}}_j \Delta \boldsymbol{\pi}_j\right]}_{\textit{Effect of adult productivity}} \qquad \text{[VI.2]}$$

Chapter VII
Combination of regression and mathematical decompositions



The combination covered here is less common. Consider the link between money and happiness. Assume, for the sake of simplicity, that this relationship is linear.

$$Y = \alpha + \beta\$ + e \qquad [VII.1]$$

Next, consider that people have multiple income sources, perhaps wages and transfers, and that they derive different levels of satisfaction from each. If that is true, then the formula VII.1. is simplistic and should ideally incorporate this differential satisfaction.

Alternatively, one can begin with a mathematical relationship and incorporate a regression relationship. Consider, for instance, the factors shaping education spending per child. Suppose, as is likely, that the share of a family budget allocated to its children education depends on parental income , so the researcher can refine the analysis by integrating details from the regression study.

The possibilities depend only on the researcher's own imagination. Still, even where further expansion is possible, the researcher must compromise between detail and parsimony. The combination opportunities suggested are not ready-made recipes to apply mechanically. Rather, they are tools to be used selectively by researchers in their quest to understand social change.

Chapter VIII
Summary and conclusions

This monograph offers an introduction to decomposition, a cluster of methods that can enrich the methodological toolbox of social scientists. The adoption of these methods advances the study of social change beyond the current limits of micro and macro-regressions. While decomposition does not address causation, it can reliably locate the social or geographic sources of social change. In doing so, decomposition reduces the margin of error in understanding societal change and in designing related policies. Within these limitations, decomposition is a reliable and transparent tool that, alone or with other methods, can usefully guide the allocation of policy resources. Policymakers faced with several options can use a decomposition approach to narrow down the list of most promising options.

The flexibility of the decomposition method makes it possible to imagine creative extensions. The few options presented here are not an exhaustive list, and researchers are encouraged to imagine other possibilities. In addition, many of the examples used here draw from the fields of population or economics, but the methods apply to a wide range of social phenomena.

One strength of the decomposition approach is its compatibility with other methods. When carefully combined with other methods, it can enrich understanding. It is therefore not a substitute but a complement. It complements other methods by identifying key processes and groups. At the same time, it leaves room for other methods (say causal analysis or informant interviews) to elucidate questions about causation, processes, key events, and key actors. Such complementarities facilitate a more complete understanding of social change. In a sense, and going back to our initial story of the proverbial drunk looking exclusively under the lamp-post, the contribution of decomposition is twofold: it widens the search area by considering all possible sources of change; and it can help begin the search closer to where the key—rather than the light—is. Having identified the groups of processes that lead the change, researchers can launch additional investigations into the reasons why these groups did change. Ultimately, those who seek to change the world can do so more effectively with a better understanding of the drivers of social change.

References

Adams Jr, R. H. (1994). Non-farm income and inequality in rural Pakistan: A decomposition analysis. *The Journal of Development Studies*, *31*(1), 110–133.

Barsky, R., Bound, J., Charles, K. K., & Lupton, J. P. (2002). Accounting for the Black–White Wealth Gap. *Journal of the American Statistical Association*, *97*(459), 663–673. https://doi.org/10.1198/016214502388618401

Bauer, T., Göhlmann, S., & Sinning, M. (2007). Gender differences in smoking behavior. *Health Economics*, *16*(9), 895–909.

Bechtel, W., & Richardson, R. C. (2010). *Discovering complexity: Decomposition and localization as strategies in scientific research*. MIT Press.

Bertrand, M., & Hallock, K. F. (2001). The gender gap in top corporate jobs. *ILR Review*, *55*(1), 3–21.

Blau, F. D., & Kahn, L. M. (1992). The Gender Earnings Gap: Learning from International Comparisons. *The American Economic Review*, *82*(2), 533–538.

Blau, F. D., & Kahn, L. M. (2006). The US gender pay gap in the 1990s: Slowing convergence. *ILR Review*, *60*(1), 45–66.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human Resources*, 436–455.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica: Journal of the Econometric Society*, 901–920.

Bourguignon, F., & Ferreira, F. (2005). Decomposing changes in the distribution of household incomes: methodological aspects. *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America, Ed. by F. Bourguignon, FH Ferreira, and N. Lustig*, 17–46.

Bourguignon, F., Ferreira, F. H., & Leite, P. G. (2008). Beyond Oaxaca–Blinder: Accounting for differences in household income distributions. *The Journal of Economic Inequality*, *6*(2), 117–148.

Cho, L.-J., & Retherford, R. D. (1973). Comparative analysis of recent fertility trends in East Asia.

Eloundou-Enyegue, P. M., & Giroux, S. C. (2012). Fertility transitions and schooling: From micro-to macro-level associations. *Demography*, *49*(4), 1407–1432.

Fairlie, R. W. (2005). An extension of the Blinder-Oaxaca decomposition technique to logit and probit models. *Journal of Economic and Social Measurement*, *30*(4), 305–316.

Firebaugh, G., & Goesling, B. (2004). Accounting for the recent decline in global income inequality. *American Journal of Sociology*, *110*(2), 283–312.

Fortin, N., Lemieux, T., & Firpo, S. (2011). Chapter 1 - Decomposition Methods in Economics. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics* (Vol. 4, pp. 1–102). Elsevier. https://doi.org/10.1016/S0169-7218(11)00407-2

Gage, T. B. (1993). The decline of mortality in England and Wales 1861 to 1964: Decomposition by cause of death and component of mortality. *Population Studies*, *47*(1), 47–66.

Gómez-Redondo, R., & Boe, C. (2005). Decomposition analysis of Spanish life expectancy at birth: evolution and changes in the components by sex and age. *Demographic Research*, *13*, 521–546.

Gupta, P. D. (1978). A general method of decomposing a difference between two rates into several components. *Demography*, *15*(1), 99–112.

Gupta, P. D. (1991). Decomposition of the difference between two rates and its consistency when more than two populations are involved. *Mathematical Population Studies*, *3*(2), 105–125.

Gupta, P. D. (1993). *Standardization and decomposition of rates: A user's manual*. US Department of Commerce, Economics and Statistics Administration, Bureau of the Census.

Horiuchi, S., Wilmoth, J. R., & Pletcher, S. D. (2008). A decomposition method based on a model of continuous change. *Demography*, *45*(4), 785–801.

Kaufman, R. L. (1983). A structural decomposition of black-white earnings differentials. *American Journal of Sociology*, *89*(3), 585–611.

Kim, C. (2010). Decomposing the change in the wage gap between white and black men over time, 1980-2005: An extension of the Blinder-Oaxaca decomposition method. *Sociological Methods & Research*, *38*(4), 619–651.

Kim, Y. J., & Strobino, D. M. (1984). Decomposition of the difference between two rates with hierarchical factors. *Demography*, *21*(3), 361–372.

Kim, Y.-M., & Shirahase, S. (2014). Understanding intra-regional variation in gender inequality in East Asia: Decomposition of cross-national differences in the gender earnings gap. *International Sociology*, *29*(3), 229–248.

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, *50*(272), 1168–1194.

Krieg, J. M., & Storer, P. (2006). How much do students matter? Applying the Oaxaca decomposition to explain determinants of adequate yearly progress. *Contemporary Economic Policy*, *24*(4), 563–581.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill? *American Economic Review*, *96*(3), 461–498.

Liao, T. F. (1989). A flexible approach for the decomposition of rate differences. *Demography*, *26*(4), 717–726.

Luke, D. A. (2004). Multilevel Modeling, Quantitative Applications in the Social Sciences. *London: A Sage University Paper Series*.

Martikainen, P., Valkonen, T., & Martelin, T. (2001). Change in male and female life expectancy by social class: decomposition by age and cause of death in Finland 1971–95. *Journal of Epidemiology & Community Health*, *55*(7), 494–499.

McEwan, P. J., & Marshall, J. H. (2004). Why does academic achievement vary across countries? Evidence from Cuba and Mexico. *Education Economics*, *12*(3), 205–217.

Mookherjee, D., & Shorrocks, A. (1982). A decomposition analysis of the trend in UK income inequality. *The Economic Journal*, *92*(368), 886–902.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 693–709.

Oaxaca, R. L., & Ransom, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics*, *61*(1), 5–21.

Oaxaca, R. L., & Ransom, M. R. (1999). Identification in detailed wage decompositions. *Review of Economics and Statistics*, *81*(1), 154–157.

Page, H. J. (1977). Patterns underlying fertility schedules: A decomposition by both age and marriage duration. *Population Studies*, *31*(1), 85–106.

Pollard, J. H. (1988). On the decomposition of changes in expectation of life and differentials in life expectancy. *Demography*, *25*(2), 265–276.

Ponnapalli, K. M. (2005). A comparison of different methods for decomposition of changes in expectation of life at birth and differentials in life expectancy at birth. *Demographic Research*, *12*, 141–172.

Powers, D. A., & Yun, M.-S. (2009). Multivariate decomposition for hazard rate models. *Sociological Methodology*, *39*(1), 233–263.

Pradhan, J., & Arokiasamy, P. (2010). Socio-economic inequalities in child survival in India: a decomposition analysis. *Health Policy*, *98*(2), 114–120.

Retherford, R. D., & Ogawa, N. (1978). Decomposition of the Change in the Total Fertility Rate in the Republic of Korea, 1966–1970. *Social Biology*, *25*(2), 115–127.

Retherford, R. D., & Rele, J. R. (1989). A decomposition of recent fertility changes in South Asia. *Population and Development Review*, 739–747.

Robinson, W. S. (1961). Ecological correlations and the behavior of individuals. *Studies in Human Ecology (Row Peterson, Evanston, Ill.)*, 115–120.

Rodrik, D. (2012). Why we learn nothing from regressing economic growth on policies.

Sandberg, J. F., & Hofferth, S. L. (2001). Changes in children's time with parents: United States, 1981–1997. *Demography*, *38*(3), 423–436.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica: Journal of the Econometric Society*, 613–625.

Shorrocks, A. F. (1984). Inequality decomposition by population subgroups. *Econometrica: Journal of the Econometric Society*, 1369–1385.

Shorrocks, A. F. (2013). Decomposition procedures for distributional analysis: a unified framework based on the Shapley value. *Journal of Economic Inequality* 11, 99-126

Shorrocks, A., & Wan, G. (2005). Spatial decomposition of inequality. *Journal of Economic Geography*, *5*(1), 59–81.

Smith, H. L., Morgan, S. P., & Koropeckyj-Cox, T. (1996). A decomposition of trends in the nonmarital fertility ratios of blacks and whites in the United States, 1960–1992. *Demography*, *33*(2), 141–151.

Solow, R. M. (1957). Technical change and the aggregate production function. *The Review of Economics and Statistics*, 312–320.

Thornton, A. (2001). The developmental paradigm, reading history sideways, and family change. *Demography*, *38*(4), 449–465.

UNICEF. (2008). *The State of Africa's Children, 2008: Child Survival*. UNICEF.

Vaupel, J. W., & Romo, V. C. (2003). Decomposing change in life expectancy: A bouquet of formulas in honor of Nathan Keyfitz's 90th birthday. *Demography, 40*(2), 201–216.

Wan, G., & Zhou, Z. (2005). Income Inequality in Rural China: Regression-based Decomposition Using Household Data. *Review of Development Economics, 9*(1), 107–120.

Yamaguchi, K. (2017). Decomposition Analysis of Segregation. *Sociological Methodology, 47*(1), 246–273.

Industrial Marketing Management 107 (2022) 315–322

Contents lists available at ScienceDirect

Industrial Marketing Management

journal homepage: www.elsevier.com/locate/indmarman

Variance decomposition analysis: What is it and how to perform it – A complete guide for B2B researchers

Ghasem Zaefarian [a,*], Viacheslav Iurkov [b], Mariia Koval [b]

[a] Leeds University Business School, University of Leeds, Maurice Keyworth Building, Leeds LS2 9JT, United Kingdom
[b] Grenoble Ecole de Management, 12 Rue Pierre Semard, 38000 Grenoble, France

ARTICLE INFO

Keywords:
B2B firms
Variance decomposition
Multilevel methods
Step-by-step guide

ABSTRACT

Variance decomposition analysis allows partitioning the total variance in an outcome variable, e.g., firm performance, into several components. Such partitioning allows identifying groups of factors (e.g., firm-, industry-, and country-specific) that explain a significant portion of the variation in firm performance, thus helping researchers, managers, and policymakers better understand the sources of competitive advantage. The present study aims to inform scholars, particularly those in business-to-business (B2B) marketing, about the benefits of utilizing variance decomposition analysis and draw scholarly attention to the relevant statistical techniques needed to produce accurate estimates. We specifically point to multilevel modeling techniques due to their significant advantages over other approaches to decompose the variance in a given outcome variable. We provide a detailed step-by-step guide as well as the related Stata codes on conducting variance decomposition analysis with multilevel modeling techniques. Using a 10-year (2009–2018) dataset comprising 7281 distinct European B2B firms operating in 348 industries and 29 countries, we empirically examine the relative importance of firm, industry, country, year, and residual effects in driving firm performance for B2B firms. Our analysis shows that firm-specific factors have the highest relative importance for B2B firms' performance, followed by home country and industry effects.

1. Introduction

Variance decomposition analysis is a statistical technique that allows partitioning the total variance in an outcome variable, for example, firm financial performance, into several components (groups of factors), such as firm, industry, and country (e.g., Guo, 2017; Makino, Isobe, & Chan, 2004; McGahan & Porter, 1997; Rumelt, 1991). Being able to identify effects that explain a significant portion of the variation in firm behavior and performance, variance decomposition analysis helps shed light on areas researchers should focus their attention to explain the phenomena. Such analysis can provide managers and policymakers with guidance regarding the most important sources of competitive advantage. Variance decomposition techniques have been widely utilized in social science disciplines, including strategic management (e.g., Guo, 2017; McGahan & Porter, 1997; Misangyi, Elms, Greckhamer, & Lepine, 2006), international business (e.g., Ma, Tong, & Fitza, 2013; McGahan & Victer, 2010), and economics (e.g., Schmalensee, 1985; Tarziján & Ramirez, 2010), enabling researchers to study the relative importance of various effects for behavior and outcomes of economic actors.

Rumelt (1991) and McGahan and Porter (1997) were some of the first to apply the variance decomposition approach to examine the relative importance of different groups of effects, such as business segment, corporate parent, and industry, on firm financial performance. Since then, researchers have increasingly been applying this methodology, adding new factors, such as country (Chan, Isobe, & Makino, 2008; Makino et al., 2004; McGahan & Victer, 2010) or strategic groups (Chang & Hong, 2002), to explain variation in firm performance. In parallel, there has been an advancement in methodological approaches from more basic ones, such as analysis of variance (ANOVA) (McGahan & Porter, 1997, 2002), to multilevel modeling with maximum likelihood estimation (Hough, 2006; Karniouchina, Carson, Short, & Ketchen, 2013; Misangyi et al., 2006) or more complex estimation methods involving Bayesian Markov chain Monte Carlo (MCMC) algorithms (Castellaneta & Gottschalg, 2016; Guo, 2017).

While being widely used in strategic management and other management subfields, variance decomposition analysis has remained overlooked in marketing. Among the few applications in marketing is a study by Zhang, Hult, Ketchen, and Calantone (2020). The authors argue

* Corresponding author.
E-mail addresses: G.Zaefarian@leeds.ac.uk (G. Zaefarian), viacheslav.iurkov@grenoble-em.com (V. Iurkov), mariia.koval@grenoble-em.com (M. Koval).

https://doi.org/10.1016/j.indmarman.2022.10.020
Received 19 March 2022; Received in revised form 17 October 2022; Accepted 21 October 2022
Available online 29 October 2022

that innovation as a strategic resource exists within (firm-level innovation) and beyond organizational boundaries (industry- and country-level innovation). Their variance decomposition analysis accounts for firm-, industry- and country-level factors to explain innovation-related variance in firm performance. Their analysis revealed that industry- and country-level innovations are far more important than firm-level innovation in explaining firm performance, suggesting scholars and managers to be more attentive to the former.

This paper is an attempt to (re)introduce the variance decomposition analysis as an effective technique to marketing scholars. We provide a step-by-step guide as well as the related Stata codes for (marketing) scholars that explains how to estimate the relative contribution of different types of effects on firm performance in Stata using multilevel modeling techniques. We illustrate the process by implementing the technique in an exemplative study that examines the relative importance of firm, industry, country, year, and residual effects on firm performance for B2B firms. Our study also shows how one can model data structures that are not perfectly hierarchical but rather cross-nested (e.g., firms are simultaneously nested within counties and industries). Such data structures are common in management and marketing research (Andersson, Cuervo-Cazurra, & Nielsen, 2014).

Examining the relative importance of firm, industry, country, and year on the financial performance of B2B firms, the present study relies on a 10-year (2009–2018) dataset comprising 7281 distinct European B2B firms operating in 348 industries and 29 countries. We decompose the variance in firm return on assets (ROA) accounting for a cross-nested data structure: repeated observations of ROA are nested within firms, which are cross-nested within both industries and countries. We utilize two of the most recent methodological approaches in variance decomposition research, namely, multilevel modeling with maximum likelihood estimation (e.g., Hough, 2006; Misangyi et al., 2006) and multilevel modeling with MCMC estimation in a Bayesian framework (e.g., Castellaneta & Gottschalg, 2016; Guo, 2017). We find that for B2B firms, firm-specific effects account for around 56.88% of the variation in ROA, while industry and country effects constitute only 2.76% and 3.98%, respectively. Similar to other variance decomposition studies, the proportion of residual variation constitutes 36.26%, with year effects remaining negligible (0.12%).

Our findings shed light on significant sources of variability in B2B firm performance. While the results point to a substantial role of firm-specific effects (i.e., firm resources and capabilities have been claimed instrumental for B2B firm performance), the impact of the external environment (especially industry) remains rather limited. The findings imply that scholars and practitioners should focus on firm-specific factors as the main driver of competitive advantage. The sources of the substantial amount of residual variation could be further explored by introducing other levels of analysis. More generally, utilizing variance decomposition study in marketing provides ramifications for marketing theory and practice: it highlights crucial yet less researched areas in extant marketing literature and provides practitioners with guidance regarding which factors affect performance (of different kinds) the most. Finally, our study adds to a series of methodological papers in Industrial Marketing Management (e.g., Davvetas, Diamantopoulos, Zaefarian, & Sichtmann, 2020; Ullah, Zaefarian, Ahmed, & Kimani, 2021; Ullah, Zaefarian, & Ullah, 2021) by enhancing marketing scholars' understanding of the exact procedures to implement variance decomposition analysis using common statistical packages.

2. Variance decomposition analysis

2.1. Overview, development over the years, and theoretical value

Variance decomposition analysis is a widely used methodology in strategic management, international business, and economics. The methodology implies splitting the total variance in an outcome variable, such as firm performance, into various components or groups of factors, such as country factors that explain differences in institutions and cultures, industry factors characterizing the competitive landscape in the industry, and firm factors capturing unique firm resources and capabilities. The methodology implies computing the relative importance of different groups of factors and understanding which ones drive firm performance the most.

Rumelt (1991) was among the first to conduct a variance decomposition study to understand the sources of variation in the rate of return on assets of business units of American manufacturing corporations observed over 1974–1977. Rumelt was particularly interested in comparing industry and business-unit effects to challenge the then-dominant view that industry structure is a central determinant of firm performance instead of unique resources and capabilities. Rumelt disaggregated business-unit profitability into components associated with the industry, corporate parent, and business unit effects. While substantial industry effects would support industrial organization tradition, sizeable business-unit effects would instead favor the propositions of the resource-based view. Rumelt (1991) found relatively minor industry and corporate parent effects, yet significant business unit effects. Specifically, while industry effects accounted for 4% of the variance in business-unit profitability, business unit effects were equal to 44%. These results confirmed firms' resources and capabilities to be important determinants of heterogeneity in firm performance. Rumelt's study gave rise to fruitful research in strategic management, further refining his original findings.

Building on Rumelt's (1991) work, McGahan and Porter (1997) examined the contribution of industry, year, corporate parent, and business unit effects in the variation of U.S. public corporations' profitability. Using data from Compustat's Business Segment Reports, which covered corporate parents and their business units in numerous industries during 1981–1994, McGahan and Porter found that industry, corporate parent, and business unit effects accounted for 19%, 4%, and 32% in variation of business unit profitability. They also found that the contribution of each type of effects varied significantly across economic sectors. For example, industry effects accounted for a smaller portion of variation in profitability in manufacturing sectors, while a more significant portion was observed in services, wholesale and retail trade, transportation, and entertainment sectors.

Both Rumelt (1991) and McGahan and Porter (1997) suggested that time-varying effects need to be distinguished from time-invariant ones when explaining firm performance. The majority of studies investigating the relative importance of business unit, corporate parent, and industry effects also introduced so-called "year effects," or, in other words, the effects of "macroeconomic fluctuations that affect all business segments to the same degree in a particular year" (McGahan & Porter, 1997: 24). These effects stem from a reduction in residual variation when year fixed-effects are incorporated in a model (Misangyi et al., 2006).

Since McGahan and Porter's (1997) work, there has been a surge in variance decomposition research in strategic management and other fields. For example, Chang and Hong (2002) extended Rumelt (1991) and McGahan and Porter (1997) to the Korean context. They decomposed the variance in the profitability of companies associated with Korean business groups into the business group-, industry-, and company-specific factors, which roughly matched corporate parent-, industry-, and business unit-specific factors in Rumelt and McGahan and Porter's studies, respectively. Chang and Hong (2002) found substantial business group effects (around 9%), controlling for industry and company effects. The findings indicated that business groups could play an essential role in developing countries with higher market inefficiencies.

Makino et al. (2004) significantly extended prior variance decomposition studies that focused on a single country by adding a new level, namely the host country. Specifically, the authors focused on multinational corporations and examined the relative importance of host country effects in explaining variation in foreign affiliates' performance. They relied on a database consisting of a panel of foreign affiliates of Japanese multinational corporations observed over 1996–2001. Makino and colleagues found that country factors (equal to nearly 6%) were as vital as industry factors (7%), following affiliate (31%) and corporate parent (11%) factors. The results also suggested that corporate parent and affiliate factors explained the most variance in the foreign affiliates' performance in the developed countries, equating to 13% and 28%, respectively. In contrast, host country (4%) and industry (5%) factors were more critical in the subsample of developing countries. Makino et al. (2004) concluded that internal (affiliate and corporate) effects tend to play a relatively more salient role than external (industry and country) effects in the developed country context because "countries with advanced economies are more integrated in terms of market transactions, infrastructure, institutional rules and enforcement mechanisms" (p. 1038).

In a more recent study, Bamiatzi, Bozos, Cavusgil, and Hult (2016) attempted to reconcile institutional theory with the resource-based view and industrial organization economics by decomposing variation in firm performance in recessionary and expansionary economic periods. Utilizing a sample of more than 15,000 firms from ten emerging and ten developed economies and operating in 779 industries, Bamiatzi and colleagues found that the relative importance of firm effects on firm performance is higher during periods of economic recession as opposed to economic expansion. This result helped solve the existing contention about whether firm heterogeneity should become more pronounced during recessionary periods when the rules of the game are fluid, and there are imperfections in strategic factor markets (Oliver, 1997).

Hence, variance decomposition analysis has proven helpful in refining our knowledge of the drivers of firm performance heterogeneity and better understanding the boundaries of existing theories of competitive advantage.

2.2. Key analytical techniques to conduct variance decomposition study

2.2.1. Overview of earlier techniques

Earlier studies utilizing variance decomposition analysis to estimate components of firm performance relied on two approaches: variance components analysis (VCA) or analysis of variance (ANOVA) (e.g., McGahan & Porter, 1997, 2002; Rumelt, 1991). Although these studies provided a step forward in enhancing our understanding of the "general importance of industry, corporate, and business effects on firm performance" (McGahan & Porter, 2002: 835), VCA and ANOVA have been found to have several critical limitations (Bowman & Helfat, 2001; Brush & Bromiley, 1997).

The advantage of VCA is that it allows decomposing variance for calculating the relative importance of different effects. However, this technique requires certain assumptions to be met (Garson, 2012): (1) effects of different components must be independent and identically distributed, (2) residuals must be uncorrelated and normally distributed, and (3) residuals must have constant variance. If these assumptions are violated, estimates will be biased and unreliable. VCA also does not allow modeling interaction effects. Thus, estimates may be biased when cross-level interactions exist.

In turn, ANOVA progressively adds components into the model allowing calculating incremental R^2. Respective incremental changes to R^2 show the relative importance of various components. ANOVA requires the following assumptions to be met (Hedeker & Gibbons, 2006): (1) sphericity of variance-covariance, and (2) the residuals are independent and normally distributed. Compared to VCA, ANOVA is more robust to deviations from these assumptions, producing more consistent and reliable empirical findings (Garson, 2012). At the same time, the main limitations of ANOVA are that it provides no explicit variance decomposition (i.e., the result differs depending on how the effects enter the model) and does not allow to capture interaction effects adequately (Hoffman, 2015). Hence, if interactions exist, the unmodeled effects may be confounded with other effects, resulting in potentially biased results.

Since both ANOVA and VCA assume independence between effects (Bowman & Helfat, 2001; Brush & Bromiley, 1997), both methods present difficulties in calculating the correct size of the effects. Due to the inability of the methods to incorporate the relationships that could exist between effects, McGahan and Porter (2002: 850) have concluded that "while there are ways to continue to learn from this research, its limits suggest that the time has come to explore whole new approaches." VCA and ANOVA have subsequently been replaced by multilevel modeling.

2.2.2. Cross-nested data structures and the limitations of earlier variance decomposition techniques

A traditional stream in variance decomposition research studies the performance of business units that are nested, or "embedded," within both corporate parents and industries (McGahan & Porter, 1997, 2002; Rumelt, 1991). Rumelt (1991: 171) acknowledged that "both industries and corporations are considered to be sets of business units." Although economics often views business units as atomistic actors, they are still interrelated with industries, as firm conduct is reciprocally related to industry conditions (Henderson & Mitchell, 1997; Porter, 1980). Corporations are not strictly nested within industries as they often have multibusiness operations, and industries are not nested within corporations. At the same time, corporations and industries are interrelated (McGahan & Porter, 2002: 838) because "the covariance between industry and corporate-parent effects is potentially important because, for example, a diversified firm may be more likely to expand into particular types of industries." Therefore, business units, corporate parent, and industry effects are not independent, producing biased results in ANOVA and VCA techniques (Misangyi et al., 2006). Similarly, international business research has acknowledged that it is critical to account for the simultaneous embeddedness of firms in industry and country contexts when exploring drivers of firm performance (Andersson et al., 2014).

To conclude, firm performance varies across levels that are not characterized by perfect hierarchical nesting (Andersson et al., 2014; Guo, 2017; Misangyi et al., 2006). As mentioned above, business unit performance over time is nested within business units, and business units may be cross-nested within corporate parents, industries, and countries. To appropriately model such complex structures, scholars have advocated for *multilevel modeling* that provides a number of advantages over other techniques to partition the variation in firm performance or other outcome variables.

2.2.3. Multilevel modeling (MLM) approaches to variance decomposition

Multilevel modeling (MLM), also known as hierarchical linear modeling, allows addressing the discussed limitations of both VCA and ANOVA (Hoffman, 2015). First, MLM directly decomposes variance in the outcome variance into each level (Hoffman, 2015), relaxing the assumption of independence of lower-level units nested in higher-level units. Second, MLM ensures efficient estimation of unbalanced data, preventing the loss of information (Raudenbush & Bryk, 2002). Third, MLM permits modeling complex data structures, such as interactions, cross-nesting, and multiple-membership (Browne, Goldstein, & Rasbash, 2001), thus addressing the issue of collinearity among different types of the effects, such as business units, corporate parents, and industries (Guo, 2017; Hough, 2006; Misangyi et al., 2006). Fourth, while VCA and ANOVA employ only categorical independent variables (Brush & Bromiley, 1997; McGahan & Porter, 1997, 2002), MLM allows explaining variance in a more nuanced way, including both categorical and continuous independent variables that could be either time-invariant or time-varying (Guo, 2017; Misangyi et al., 2006). Finally, MLM can account for autocorrelation in longitudinal data (Guo, 2017). Therefore, MLM is a significant step forward in the variance decomposition research that allows moving away from descriptive models toward inferential models to examine relationships and test multiple theories (Andersson et al., 2014; Guo, 2017; Hough, 2006; Mathieu & Chen, 2011; Misangyi et al., 2006).

Recent variance decomposition studies (e.g., Castellaneta & Gottschalg, 2016; Guo, 2017; Karniouchina et al., 2013; Meyer-Doyle, Lee, &

Helfat, 2019; Misangyi et al., 2006) have shown the advantages of using MLM as opposed to VCA and ANOVA. In particular, scholars have modeled the cross-nested structure of corporate parent and industry and obtained a more nuanced picture compared to studies utilizing earlier techniques (e.g., McGahan & Porter, 1997, 2002; Rumelt, 1991).

2.3. Use of the variance decomposition analysis in marketing

Despite being common in strategic management and other related disciplines, marketing scholars have only recently started to employ variance decomposition analysis to examine the relative role of different factors in driving firm behavior and outcomes. One of the few examples is Zhang et al. (2020) who applied the variance decomposition approach to study the contribution of the firm-, industry- and country-level innovation in explaining variance in firm performance. Zhang et al. (2020) compiled a dataset comprising 4530 firms operating in 794 industries with headquarters in 39 countries. They utilized hierarchical linear multilevel modeling techniques to test the effects of innovation at the three levels on performance. The technique was appropriate due to the hierarchical nature of the study's data and the possibility of simultaneously partitioning the variance-covariance components. Specifically, firms were nested within industries, which were nested within countries. Having decomposed variance in firm performance, the authors found that industry- and country-level innovations were the most important drivers of firm performance by explaining 34% and 40% of the variance in firm performance, respectively, contrary to firm-level innovation explaining only 26% of the total variance.

In the following section, we provide a step-by-step procedure to conduct variance decomposition analysis. We examine the relative importance of firm, industry, country, year, and residual effects in explaining the performance of B2B firms. We utilize multilevel modeling to estimate the effects at each level of analysis. Specifically, we utilize two existing MLM approaches to decompose the variance in firm performance: a more common approach relying on multilevel modeling with maximum likelihood estimation (e.g., Hough, 2006; Misangyi et al., 2006) and a more recent approach involving multilevel modeling with MCMC estimation in a Bayesian framework (e.g., Castellaneta & Gottschalg, 2016; Guo, 2017). We demonstrate how a variance decomposition analysis could be done in common statistical software packages, such as Stata.

3. Step-by-step procedure for the variance decomposition analysis

3.1. Step 1: Defining categories of explanatory variables

The steps follow several guidelines established in the previous variance decomposition studies (Guo, 2017; McGahan & Porter, 1997; McGahan & Victer, 2010; Misangyi et al., 2006). First, we have to clearly define categories of explanatory variables. The aim of variance decomposition studies is mainly to capture categorical effects, e.g., industry, country, corporate parent, and business unit as a whole (McGahan & Porter, 1997, 2002). We define a *firm* as an independent entity as indicated in the Orbis database (i.e., no shareholder with more than 25% direct or total ownership). Such firms are not prone to external interference in their decision-making, having more freedom in their strategy development and implementation. The firm's *industry* membership is based on its primary 4-digit NACE Rev. 2 code.[1] The *home country* is defined as the country of the firm's headquarters.

3.2. Step 2: Initial data selection

We then continue by constructing a dataset. Variance decomposition studies normally require a sizeable amount of lower-level observations (e.g., firms nested in industries or countries). Also, to separate residual effects from firm effects, the data should have a panel structure (Guo, 2017; Misangyi et al., 2006). In variance decomposition studies, residual effects can also be referred to as "error" or "unexplained variance," i.e., "the performance variance potentially attributable to transient factors" (Misangyi et al., 2006: 580). We obtained all data for the analyses from the Orbis Europe database, which identifies both listed and nonlisted European firms (both within and outside of the European Union), the core industry in which they operate, and their home country. The period considered for this study is 2009–2018.

3.3. Step 3: Identifying possible subsamples

As discussed above, we investigate the relative importance of firm, industry, home country, and year effects on firm performance for B2B firms. A B2B firm is defined in line with Delgado and Mills (2020). The authors classify all industries into two broad categories: those selling primarily to businesses or government (i.e., business-to-business industries) and those selling primarily to consumers (i.e., business-to-consumer industries). If a firm's primary activity is in the former type of industry, it is classified as a B2B firm and is of interest to our study.[1 and 2] More specifically, Delgado and Mills' classification is based on the percentage of output sold to Personal Consumption Expenditure (PCE). "The PCE is a final use item in the [Input-Output] IO Accounts that captures the value of the goods and services that are purchased by households, such as food, cars, and college education" (Delgado & Mills, 2020, p. 4). Delgado and Mills classify an industry as B2B if it sells less than 35% of its output to PCE, with the rest being classified as B2C. To date, this is one of the most systematic and comprehensive classifications of industries into B2B and B2C.

3.4. Step 4: Data cleaning

Our total initial sample consisted of 1,322,458 observations of yearly firm performance for the years 2009–2018. This dataset is further screened following the steps initially reported in McGahan and Porter (1997) and then adopted by subsequent variance decomposition studies (e.g., Guo, 2017; Karniouchina et al., 2013; Misangyi et al., 2006). From our original total, we drop: (a) 62,102 observations of firms conducting financial and insurance activities and (b) 30,024 observations of firms with activities in public administration and defense, compulsory social security, and not elsewhere classified. Variance decomposition research often excludes firms from these industries as their returns are not comparable with those of other industries according (McGahan & Porter, 1997). We further proceed to eliminate (c) 1,161,740 observations of small firms with sales or assets less than €10 million and (d) 1390 observations of firms that have been in this dataset for only one year. Single-year appearances may have anomalous performance and are normally excluded from the analysis (McGahan & Porter, 1997, 2002; Misangyi et al., 2006). Unless small firms are of interest to a researcher, they are also often excluded from the analysis for a reason similar (McGahan & Porter, 1997, 2002). We also drop (e) 787 observations of firms that are the only ones in an industry in a given year, as these entities may be analogous to monopolies (McGahan & Porter, 1997). In addition, (f) 619 firm observations are removed due to a limited number of them in a given country as opposed to other countries. Keeping these observations in the analysis can affect the power of

[1] NACE is the Statistical Classification of Economic Activities in the European Community. The current version is revision 2. It is the European implementation of the United Nations International Standard Industrial Classification of All Economic Activities (ISIC), revision 4.

[2] Delgado and Mills' classification is provided with respect to the North American Industry Classification System (NAICS). We use conversion tables to translate it to NACE Rev. 2.

higher-level effects (Hofmann, 1997; Peterson, Arregle, & Martin, 2012), such as home country effects. Finally, we remove (g) 14,272 observations from our sample as they do not fall under our definition of a B2B firm. Our final sample consists of 51,524 firm-year observations. This screened 10-year dataset represents 7281 different firms operating in 348 industries and 29 countries. This sample's mean return on assets (ROA) is 5.45%, with a variance of 10.88%.

3.5. Step 5: Model specification and estimation

The next step is to specify and estimate a model. We proceed by decomposing variance in firm ROA by fitting multilevel models to the dataset. In the dataset, repeated observations of ROA (Level 1) are nested within firms (Level 2), which are cross-nested within both industries (Level 3) and countries (Level 3). There are certain specifics in modeling such data structures as they are not characterized by perfect hierarchical nesting. Industries and countries represent imperfect hierarchies because lower-level units (firms) simultaneously belong to multiple higher-level units (industries and home countries). In other words, industries and home countries are not hierarchically nested, as not all firms from a particular home country compete in the same industry, nor do all firms competing in a particular industry originate from the same home country. Cross-nested data structures like this are common to marketing and management research and need to be appropriately modeled (Andersson et al., 2014).

We utilize two existing approaches to performance variance decomposition: multilevel modeling with maximum likelihood estimation (e.g., Hough, 2006; Misangyi et al., 2006) and multilevel modeling with MCMC estimation in a Bayesian framework (e.g., Castellaneta & Gottschalg, 2016; Guo, 2017). The former approach has been more standard in the variance decomposition research, and it can be easily implemented using standard statistical software, such as using the "mixed" command in Stata. The latter approach has been introduced to the management literature relatively recently; one of its benefits is that it allows obtaining reference statistics, such as standard error, for both the absolute effects (variances) and relative effects (percentages) of different variance components (Browne, 2017).

3.5.1. Multilevel modeling with maximum likelihood estimation

We start by following Misangyi et al. (2006) to perform *multilevel modeling with maximum likelihood estimation*. It involves estimating a series of equations that nest repeated observations of firm ROA within firms and cross-nest firms within both industries and countries. In line with prior research, our models also attempt to capture so-called year effects or the general impact of macroeconomic fluctuations in business activity (McGahan & Porter, 1997, 2002). First, an unconditional three-level model (i.e., a model with no predictors) is estimated (Model 1). The model separates the variation in firm ROA into three components. At Level 1, firm ROA is modeled as the following:

$$ROA_{tij} = \alpha_{0ij} + e_{tij}, \tag{1.1}$$

where ROA_{tij} is firm ROA at time t in firm i in industry j; α_{0ij} is the over-time mean ROA of firm i in industry j; e_{tij} is the time-level random error. The model assumes $e_{tij} \sim N(0, \sigma_e^2)$. σ_e^2 is therefore denoted as *residual* (also referred to as "across-time") *variance*.

At Level 2, the over-time mean ROA, α_{0ij}, is simultaneously modeled as an outcome varying randomly around the industry mean:

$$\alpha_{0ij} = \beta_{00j} + u_{ij}, \tag{1.2}$$

where β_{00j} is the mean ROA of all firms in industry j and u_{ij} is the between-firm residual that is distributed as $u_{ij} \sim N(0, \sigma_u^2)$. σ_u^2 thus denotes *between-firm variance*.

At Level 3, the mean ROA of all firms in industry j, β_{00j}, is simultaneously modeled as an outcome varying randomly around the grand mean:

$$\beta_{00j} = \gamma_{000} + v_j, \tag{1.3}$$

where γ_{000} is the grand-mean ROA of all firms in the dataset and v_j is the between-industry residual distributed as $v_j \sim N(0, \sigma_v^2)$. σ_v^2 is *between-industry variance*.

To model the cross-nesting of country effects on firm ROA, we incorporate these effects at the firm level (Misangyi et al., 2006). Year effects are incorporated at the time level of analysis (Guo, 2017; Misangyi et al., 2006). To calculate year effects, we need to estimate Model 2 below:

$$ROA_{tij} = \alpha_{0ij} + \alpha_{1ij} Year_{tij} + e_{tij}, \tag{2.1}$$

$$\alpha_{0ij} = \beta_{00j} + u_{ij}, \tag{2.2}$$

$$\beta_{00j} = \gamma_{000} + v_j. \tag{2.3}$$

In turn, Model 3 is used to derive home country effects:

$$ROA_{tij} = \alpha_{0ij} + \alpha_{1ij} Year_{tij} + e_{tij}, \tag{3.1}$$

$$\alpha_{0ij} = \beta_{00j} + \beta_{01j} Home\ country_{ij} + u_{ij}, \tag{3.2}$$

$$\beta_{00j} = \gamma_{000} + v_j. \tag{3.3}$$

In Models 2 and 3, α_{1ij} denotes year effects (the impact of macroeconomic fluctuations in business activity); *Year* is a matrix of dummy variables coded for each of the years for each firm i in industry j. α_{0ij} now stands for across-time mean ROA for firm i in industry j adjusted for year effects. In Model 3, β_{01j} represents (stable) home country effects, i.e., the effect home country affiliation on mean firm ROA; *Country* is a matrix of dummy variables capturing home country affiliation of firm i in industry j. β_{00j} is now the mean ROA of firms nested in industry j adjusted for country effects.

Having specified Models 1, 2, and 3 above, we can now calculate the relative importance of firm, industry, country, and year effects on firm ROA, as well as the residual variation. Estimating the unconditional model comprising Eqs. (1.1), (1.2), and (1.3), the percentage of total variance attributable to each level is calculated as: $\sigma_e^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)$ is the proportion of residual variance, $\sigma_u^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)$ is the proportion of variance between firms, and $\sigma_v^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)$ is the proportion of variance between industries.

The percentage of total variance explained by *year effects* can be obtained by comparing the change in residual variance between Model 1 (unconditional model) and Model 2 comprising Eqs. (2.1)–(2.3) where year effects enter at the time level (Guo, 2017; Misangyi et al., 2006). Formally, year effects are calculated as follows: $(\sigma^2_{e,\ Model\ 1} - \sigma^2_{e,\ Model\ 2})/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)_{Model\ 1}$.

As follows from Model 3, Eqs. (3.1)–(3.3), the cross-nesting of home country effects involves their introduction at the firm level and may account for both between-firm and between-industry variance (recall that intercept β_{00j} represents the mean ROA of firms nested in industry j adjusted for country effects, and it is also modeled as the outcome at the industry level). In line with Misangyi et al. (2006), we can calculate home country effects by observing the decrease in the variance at the firm and industry levels as a proportion of total variance when home country effects are included, i.e., by comparing the Model 2 and Model 3 estimates of σ_u^2 and σ_v^2. Formally: $(\sigma^2_{u,\ Model\ 2} - \sigma^2_{u,\ Model\ 3})/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)_{Model\ 2} + (\sigma^2_{v,\ Model\ 2} - \sigma^2_{v,\ Model\ 3})/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2)_{Model\ 2}$.

To obtain the final results, we need to adjust residual, firm, and industry effects estimated in Model 1 (unconditional model) by year effects and the respective parts of home country effects. Table 1. below specified the Stata commands that can be used to estimate Models 1, 2, and 3 to obtain the variance components and calculate the relative importance of the effects.

Note that in Stata there are workarounds to decompose variance in a cross-classified model with maximum likelihood estimation in one step

Table 1
Stata commands to fit Models 1, 2, and 3 using maximum likelihood estimation.

Code	Explanation
mixed ROA || industry: || firm:	This command fits Model 1—a simple variance-components model. Note that we have two specifications for the random part, one at the industry level (Level 3), and one at the firm (Level 2). The lowest level (residual variance at Level 1) is not specified. These specifications are both null (intercept-only) models, so we are simply estimating an error term at each level. The levels need to be specified going down the hierarchy. By default, the *-mixed-* command in Stata 16 performs maximum likelihood estimation (MLE). To estimate models with restricted maximum likelihood estimation (REML), the *-reml-* option needs to be specified. Both REML and MLE estimations give almost identical result in our case.
mixed ROA i.year || industry: || firm:	This command fits Model 2. Year fixed-effects are entered in the fixed part of the model. As previously, the random part of the model estimates three variance components: residual, between-firm, and between-industry variance.
mixed ROA i.year i.country || industry: || firm:	This command fits Model 3. Year and country fixed-effects are entered in the fixed part of the model. The random part stay the same.

(Rabe-Hesketh & Skrondal, 2012). Table 2 reports Stata codes and the respective explanations for these workarounds.

Because the procedure in Table 2 is now allowing to fit a cross-classified model directly, the model specification can be rewritten as follows (e.g., Guo, 2017):

$$ROA_{tijk} = \alpha_{0ijk} + e_{tijk}, \tag{4.1}$$

$$\alpha_{0ijk} = \beta_{00jk} + u_{ijk}, \tag{4.2}$$

$$\beta_{00jk} = \gamma_{000} + v_j + \delta_k, \tag{4.3}$$

where $e_{tijk} \sim N(0, \sigma_e^2)$, $u_{ijk} \sim N(0, \sigma_u^2)$, $v_j \sim N(0, \sigma_v^2)$, and $\delta_k \sim N(0, \sigma_k^2)$. All terms except for δ_k and σ_k^2 are defined as above. δ_k is between-home country residual normally distributed with a mean of zero and variance of σ_k^2. The percentage of total variance attributable to a given type of effect is therefore calculated as: $\sigma_k^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2 + \sigma_k^2)$ for home country effects, $\sigma_v^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2 + \sigma_k^2)$ for industry effects, $\sigma_u^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2 + \sigma_k^2)$ for firm effects, and $\sigma_e^2/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2 + \sigma_k^2)$ for residual effects.

In line with Guo (2017), year effects can be calculated by entering these effects at the time level, as indicated in Model 5 below, and then comparing the changes in residual variance between this and the unconditional model. Model 5 is therefore specified as:

$$ROA_{tijk} = \alpha_{0ijk} + \alpha_{1ijk} Year_{tijk} + e_{tijk}, \tag{5.1}$$

Table 2
A Stata workaround to fit the cross-classified model with maximum likelihood estimation.

Code	Explanation
mixed ROA || _all: R.country || industry: || firm:	To estimate a cross-nested model in Stata, it is possible to use a work-around solution that creates a "fake" four-level model in which repeated observations are nested in firms, which nested in industries, which are then nested in an entire set but with random effects for dummy variables for countries. That effectively can give us variances for each level. Therefore, for countries, we use the following additional specification in the random part of the model: *-_all: R.country-*.
mixed ROA i.year || _all: R. country || industry: || firm:	Year effects are entered in the fixed part of the model, while the random part stays as above.

$$\alpha_{0ijk} = \beta_{00jk} + u_{ijk}, \tag{5.2}$$

$$\beta_{00jk} = \gamma_{000} + v_j + \delta_k. \tag{5.3}$$

Year effects are calculated as before, i.e., $(\sigma^2_{e,\ Model\ 4} - \sigma^2_{e,\ Model\ 5})/(\sigma_e^2 + \sigma_u^2 + \sigma_v^2 + \sigma_k^2)_{Model\ 4}$. Residual effects have to be subsequently adjusted on this value.

3.5.2. Multilevel modeling with MCMC estimation in a Bayesian framework

We now shift our attention to an alternative estimation approach that has been recently adopted in variance decomposition research, namely, *multilevel modeling with MCMC estimation in a Bayesian framework*. This approach has the same advantages as multilevel modeling with maximum likelihood estimation (Hough, 2006; Karniouchina et al., 2013; Misangyi et al., 2006) since both build on the assumptions of the normal joint distribution of residuals and independence of random effects (Guo, 2017). Both approaches allow modeling cross-classified structures and explicit variance decomposition, resulting in their recent application in variance decomposition research (Castellaneta & Gottschalg, 2016; Guo, 2017; Meyer-Doyle et al., 2019). The additional advantage of a Bayesian MCMC estimation approach is that it reports the means and standard deviations of the parameter monitoring chains (Browne, 2017; Goldstein, 2011), thus providing inference statistics for estimated absolute (variances) and relative (percentages) effects. Importantly, this approach can provide more precise estimates in the presence of cross-classified structures (Rasbash & Goldstein, 1994). Finally, it is possible to implement in Stata quickly, yet it requires the MLwiN software (Rasbash, Charlton, Browne, Healy, & Cameron, 2009) to be additionally installed.

We fit Models 4 and 5 using the MCMC estimation procedure with a Bayesian estimation (Browne, 2017) by running the MLwiN software (Rasbash et al., 2009) in Stata using the "runmlwin" command (Leckie & Charlton, 2012) with the cross-classification option. Because the software requires us to specify starting values for the model parameters, we first estimate a naive hierarchical linear model by iterative generalized least squares (IGLS). Even though the resulting IGLS estimates are not substantively interpretable because the effects are not perfectly nested, they typically provide good starting values for correctly fitting a cross-nested model. The procedure to estimate variances and their percentages, therefore, replicates Guo (2017). Table 3 below specified the Stata commands that can be used to estimate Models 4 and 5 to obtain the variance components.

4. Discussion of the results

We first discuss the maximum likelihood estimation results obtained by using Stata codes in Table 1 (e.g., Misangyi et al., 2006). The results of the estimation of the unconditional model, i.e., Eqs. (1.1), (1.2), and (1.3), are reported in the top panel of Table 4. The percentage of total variance attributable to each level are: 36.45% is the proportion of residual variance (σ_e^2), 60.35% is the proportion of variance between firms (σ_u^2), and 3.20% is the proportion of variance between industries (σ_v^2). We can additionally conduct likelihood-ratio tests to determine whether the random intercepts at Level 2 (versus linear model) and Level 3 (versus two-level model) improve model fit. We find that there is significant variance across firms (p = 0.000) and across industries (p = 0.000).

As noted above, the percentage of total variance explained by year effects can be obtained by comparing the change in residual variance between Model 1 and Model 2 (Guo, 2017; Misangyi et al., 2006). We find that year effects account for 0.12% of the total variance in firm ROA (as reported in the top middle panel of Table 4). Home country effects result from a decrease in the variance at the firm and industry levels as a proportion of total variance when home country effects are included. Using the results in the two middle panels of Table 3, home country effects constitute 4.18% of the total variance in firm ROA.

Table 3
Stata commands to fit Models 4 and 5 using MCMC estimation in a Bayesian framework.

Code	Explanation
runmlwin ROA cons, level4(country: cons) level3(industry: cons) level2(firm: cons) level1(year: cons)	In Stata 16, the following syntax is first employed to estimate a naive unconditional hierarchical linear model using IGLS, i.e., a naive Model 4. *cons* is a variable that is equal to 1 for all observations. It is generated as *-gen cons = 1-*. For information, this *-runmlwin-* command produces results identical to: *-mixed ROA || country: || industry: || firm:-*
runmlwin ROA cons, level4(country: cons) level3(industry: cons) level2(firm: cons) level1(year: cons) mcmc(cc on) initsprevious	We then fit Model 4—the unconditional cross-classified model—with a Bayesian MCMC estimation using the IGLS estimates from the previous command. *-initsprevious-* specifies that the parameter estimates from the previous model are used as the initial values. *-mcmc(cc on)-* specifies a cross-classified model and fits it using default MCMC options. It is important to note that *-level4-* is used for country effects because the *-runmlwin-* command does not allow indicating two Level 3s. Switching cross-classification option on allows accounting for the actual data structure. In general, while one can put the levels in any order, it is recommended putting them in the partly nested way as MLwiN will run IGLS first for starting values and this assumes nesting (Browne, 2017; Leckie & Charlton, 2012; Rasbash et al., 2009).
runmlwin ROA cons year_dummies, level4 (country: cons) level3(industry: cons) level2(firm: cons) level1(year: cons)	A command to fit Model 5 using a naive unconditional hierarchical linear model. *-year_dummies-* includes all year dummy variables from the sample (need to be manually generated).
runmlwin ROA cons year_dummies, level4 (country: cons) level3(industry: cons) level2(firm: cons) level1(year: cons) mcmc(cc on) initsprevious	A command to fit Model 5 with a Bayesian MCMC estimation using the IGLS estimates from the previous command.

To obtain the final results, we need to adjust residual, firm, and industry effects estimated in Model 1 by year effects and the respective parts of home country effects. As indicated in the bottom panel of Table 3, residual effects constitute 36.33%, firm effects are equal to 56.63%, and industry effects are 2.74%. Firm effects align with prior variance decomposition research (Guo, 2017; McGahan & Porter, 1997; Misangyi et al., 2006; Sharapov, Kattuman, Rodriguez, & Velazquez, 2021), explaining the largest proportion of total variance in B2B firm ROA. Some scholars interpret this finding to support the resource-based perspective (McGahan & Porter, 1997, 2002), emphasizing that idiosyncratic historical factors give rise to firm performance differences (Barney, 1991; Peteraf, 1993; Peteraf & Barney, 2003; Wernerfelt, 1984). Under this view, firm resources and capabilities play a major role in gaining and sustaining competitive advantages, while industry structures are less important for competitive advantage. Indeed, industry and home country effects are 20.67 and 13.55 times smaller than firm effects.

We then fit Models 4 and 5 using a Bayesian MCMC estimation (Browne, 2017). Table 5 presents the results of fitting Models 4 and 5. The relative magnitudes of the effects in Table 5 are similar to Table 4. Standard errors for the parameter estimates are calculated by using post estimation commands of *"runmlwin"*: we first save the MCMC parameter chains from the *"runmlwin"* estimations using the *"mcmcsum, getchains"* command and then calculate MCMC summary statistics using the *"mcmcsum, variables"* command.

Table 4
Results from maximum likelihood estimation of Models 1, 2, and 3.

	Parameter estimate
Model 1, unconditional model	
Level 1 variance (residual), σ_e^2	48.95319
Level 2 variance (between firms), σ_u^2	81.05917
Level 3 variance (between industries), σ_v^2	4.298835
% of total residual variance	36.45
% of total variance explained by firm effects	60.35
% of total variance explained by industry effects	3.20
Model 2, incorporating year effects at Level 1	
Level 1 variance (residual), σ_e^2	48.78532
Level 2 variance (between firms), σ_u^2	81.22981
Level 3 variance (between industries), σ_v^2	4.318855
% of total variance explained by year effects	0.12
Model 3, incorporating year effects at Level 1 and home country effects at Level 2	
Level 1 variance (residual), σ_e^2	48.80061
Level 2 variance (between firms), σ_u^2	76.22929
Level 3 variance (between industries), σ_v^2	3.69721
% of total variance explained by home country effects	4.18
Final results	
% of total variance	
residual	36.33
explained by year effects	0.12
explained by firm effects	56.63
explained by industry effects	2.74
explained by home country effects	4.18

Table 5
Results from Bayesian MCMC estimation of Models 4 and 5.

	Parameter estimate	Standard error
Model 4, unconditional model		
Level 1 variance (residual), σ_e^2	48.96681	0.3302873
Level 2 variance (between firms), σ_u^2	76.56591	1.485538
Level 3 variance (between industries), σ_v^2	3.7189	0.7561023
Level 3 variance (between home countries), σ_k^2	5.352018	2.010238
% of total variance		
residual	36.38	0.71
explained by firm effects	56.88	1.05
explained by industry effects	2.76	0.55
explained by home country effects	3.98	1.41
Model 5, incorporating year effects at Level 1		
Level 1 variance (residual), σ_e^2	48.80599	0.3267018
% of total variance		
explained by year effects	0.12	0.34
Final results		
% of total variance		
residual	36.26	0.72
explained by year effects	0.12	0.34
explained by firm effects	56.88	1.05
explained by industry effects	2.76	0.55
explained by home country effects	3.98	1.41

5. Conclusion

The variance decomposition method is a step forward in enhancing our understanding of how much effects at various levels of analysis contribute to the variation in an outcome variable, like firm performance. However, it has not been widely used in marketing yet. This study aims to show the relevance of variance decomposition study for a marketing audience and equip researchers with a step-by-step guide on how to conduct variance decomposition analysis to study research questions relevant to marketing. We hope that the present study will help marketing scholars and practitioners better understand the importance of the variance decomposition approach in explaining the relative contribution of firm, industry, country, year, and residual effects on the variation of firm performance in the B2B context. The results of our study suggest which factors should be devoted the most attention to in future research and managerial practice.

Data availability

The data that has been used is confidential.

References

Andersson, U., Cuervo-Cazurra, A., & Nielsen, B. B. (2014). From the editors: Explaining interaction effects within and across levels of analysis. *Journal of International Business Studies, 45*(9), 1063–1071.

Bamiatzi, V., Bozos, K., Cavusgil, S. T., & Hult, G. T. M. (2016). Revisiting the firm, industry, and country effects on profitability under recessionary and expansion periods: A multilevel analysis. *Strategic Management Journal, 37*(7), 1448–1471.

Barney, J. B. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Bowman, E. H., & Helfat, C. E. (2001). Does corporate strategy matter? *Strategic Management Journal, 22*(1), 1–23.

Browne, W. J. (2017). *MCMC estimation in MLwiN version 3.00*. Bristol, UK: Centre for Multilevel Modelling, University of Bristol.

Browne, W. J., Goldstein, H., & Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling, 1*(2), 103–124.

Brush, T. H., & Bromiley, P. (1997). What does a small corporate effect mean? A variance components simulation of corporate and business effects. *Strategic Management Journal, 18*(10), 825–835.

Castellaneta, F., & Gottschalg, O. (2016). Does ownership matter in private equity? The sources of variance in buyouts' performance. *Strategic Management Journal, 37*(2), 330–348.

Chan, C. M., Isobe, T., & Makino, S. (2008). Which country matters? Institutional development and foreign affiliate performance. *Strategic Management Journal, 29*(11), 1179–1205.

Chang, S. J., & Hong, J. (2002). How much does the business group matter in Korea? *Strategic Management Journal, 23*(3), 265–274.

Davvetas, V., Diamantopoulos, A., Zaefarian, G., & Sichtmann, C. (2020). Ten basic questions about structural equations modeling you should know the answers to–but perhaps you don't. *Industrial Marketing Management, 90*, 252–263.

Delgado, M., & Mills, K. G. (2020). The supply chain economy: A new industry categorization for understanding innovation in services. *Research Policy, 49*(8), Article 104039.

Garson, G. D. (2012). *Variance components analysis*. Asheboro, NC: Statistical Associates Publishers.

Goldstein, H. (2011). *Multilevel statistical models* (4th ed.). Hoboken, NJ: Wiley.

Guo, G. R. (2017). Demystifying variance in performance: A longitudinal multilevel perspective. *Strategic Management Journal, 38*(6), 1327–1342.

Hedeker, D. R., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Hoboken, NJ: Wiley.

Henderson, R., & Mitchell, W. (1997). The interactions of organizational and competitive influences on strategy and performance. *Strategic Management Journal, 18*(S1), 5–14.

Hoffman, L. (2015). *Longitudinal analysis: Modeling within-person fluctuation and change*. New York: Routledge.

Hofmann, D. A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management, 23*(6), 723–744.

Hough, J. R. (2006). Business segment performance redux: A multilevel approach. *Strategic Management Journal, 27*(1), 45–61.

Karniouchina, E. V., Carson, S. J., Short, J. C., & Ketchen, D. J. (2013). Extending the firm vs. industry debate: Does industry life cycle stage matter? *Strategic Management Journal, 34*(8), 1010–1018.

Leckie, G., & Charlton, C. (2012). Runmlwin: A program to run the MLwin multilevel modeling software from within Stata. *Journal of Statistical Software, 52*(11), 1–40.

Ma, X., Tong, T. W., & Fitza, M. A. (2013). How much does subnational region matter to foreign subsidiary performance? Evidence from fortune global 500 Corporations' investment in China. *Journal of International Business Studies, 44*(1), 66–87.

Makino, S., Isobe, T., & Chan, C. M. (2004). Does country matter? *Strategic Management Journal, 25*(10), 1027–1043.

Mathieu, J. E., & Chen, G. (2011). The etiology of the multilevel paradigm in management research. *Journal of Management, 37*(2), 610–641.

McGahan, A. M., & Porter, M. E. (1997). How much does industry matter, really? *Strategic Management Journal, 18*(S1), 15–30.

McGahan, A. M., & Porter, M. E. (2002). What do we know about variance in accounting profitability? *Management Science, 48*(7), 834–851.

McGahan, A. M., & Victer, R. (2010). How much does home country matter to corporate profitability? *Journal of International Business Studies, 41*(1), 142–165.

Meyer-Doyle, P., Lee, S., & Helfat, C. E. (2019). Disentangling the microfoundations of acquisition behavior and performance. *Strategic Management Journal, 40*(11), 1733–1756.

Misangyi, V. F., Elms, H., Greckhamer, T., & Lepine, J. A. (2006). A new perspective on a fundamental debate: A multilevel approach to industry, corporate, and business unit effects. *Strategic Management Journal, 27*(6), 571–590.

Oliver, C. (1997). Sustainable competitive advantage: Combining institutional and resource-based views. *Strategic Management Journal, 18*(9), 697–713.

Peteraf, M. A. (1993). The cornerstones of competitive advantage: A resource-based view. *Strategic Management Journal, 14*(3), 179–191.

Peteraf, M. A., & Barney, J. B. (2003). Unraveling the resource-based tangle. *Managerial and Decision Economics, 24*(4), 309–323.

Peterson, M. F., Arregle, J. L., & Martin, X. (2012). Multilevel models in international business research. *Journal of International Business Studies, 43*(5), 451–457.

Porter, M. E. (1980). *Competitive strategy*. New York: Free Press.

Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and longitudinal modeling using Stata* (3rd ed.). College Station, TX: Stata Press Publication.

Rasbash, J., Charlton, C., Browne, W. J., Healy, M., & Cameron, B. (2009). *MLwiN version 2.1*. Bristol, UK: Centre for Multilevel Modelling, University of Bristol.

Rasbash, J., & Goldstein, H. (1994). Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. *Journal of Educational and Behavioral Statistics, 19*(4), 337–350.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Rumelt, R. P. (1991). How much does industry matter? *Strategic Management Journal, 12*(3), 167–185.

Schmalensee, R. (1985). Do markets differ much? *American Economic Review, 75*(3), 341–351.

Sharapov, D., Kattuman, P., Rodriguez, D., & Velazquez, F. J. (2021). Using the SHAPLEY value approach to variance decomposition in strategy research: Diversification, internationalization, and corporate group effects on affiliate profitability. *Strategic Management Journal, 42*(3), 608–623.

Tarziján, J., & Ramirez, C. (2010). Firm, industry and corporation effects revisited: A mixed multilevel analysis for Chilean companies. *Applied Economics Letters, 18*(1), 95–100.

Ullah, S., Zaefarian, G., Ahmed, R., & Kimani, D. (2021). How to apply the event study methodology in STATA: An overview and a step-by-step guide for authors. *Industrial Marketing Management, 99*, A1–A12.

Ullah, S., Zaefarian, G., & Ullah, F. (2021). How to use instrumental variables in addressing endogeneity? A step-by-step procedure for non-specialists. *Industrial Marketing Management, 96*, A1–A6.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180.

Zhang, Y., Hult, G. T. M., Ketchen, D. J., & Calantone, R. J. (2020). Effects of firm-, industry-, and country-level innovation on firm performance. *Marketing Letters, 31*(2), 231–245.