Statistisk sentralbyrå
Statistics Norway

An exact additive decomposition of the weighted arithmetic mean

TALL SOM FORTELLER

Thomas von Brasch, Håkon Grini, Magnus Berglund Johnsen and Trond Christian Vigtel

DISCUSSION PAPERS

944

Discussion Papers No. 944, September 2021
Statistics Norway, Research Department

Thomas von Brasch, Håkon Grini, Magnus Berglund Johnsen and Trond Christian Vigtel

An exact additive decomposition of the weighted arithmetic mean*

Abstract:
The weighted arithmetic mean is used in a wide variety of applications. An infinite number of possible decompositions of the weighted mean are available, and it is therefore an open question which of the possible decompositions should be applied. In this paper, we derive a decomposition of the weighted mean that is rooted in functional analysis. Our proposed decomposition is easy to employ and interpret, and we show that it satisfies the difference counterpart to the index number time reversal test. We illustrate the framework by decomposing aggregate earnings growth from 2020Q1 to 2021Q1 in Norway and compare it with some of the main decompositions proposed in the literature. We find that the wedge between the identified compositional effects from the proposed decomposition and the Bennet decomposition is substantial, and for some industries, the compositional effects are of opposite signs.

Keywords: Index Theory; Weighted Arithmetic Mean; Decomposition

JEL classification: C43; J21

Acknowledgements: We thank Pål Boug, W. Erwin Diewert and Terje Skjerpen for valuable comments. The usual disclaimer applies.

***An earlier version of this paper circulated as "An Exact Decomposition of the Unit Value Index"**

Address: Thomas von Brasch, Statistics Norway, Research Department. E-mail: thomas.vonbrasch@ssb.no.

Håkon Grini, Statistics Norway, Division for Labour Market and Wage Statistics. E-mail: knut.grini@ssb.no.

Magnus Berglund Johnsen, Statistics Norway, Division for Labour Market and Wage Statistics. E-mail: magnus.johnsen@ssb.no.

Trond Christian Vigtel, Statistics Norway, Research Department. E-mail: trond.vigtel@ssb.no.

Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

The Discussion Papers series presents results from ongoing research projects and other research and analysis by SSB staff. The views and conclusions in this document are those of the authors

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no/en/forskning/discussion-papers
http://ideas.repec.org/s/ssb/dispap.html

ISSN 1892-753X (electronic)

Sammendrag

Vektede gjennomsnitt brukes på flere statistiske områder, blant annet for å beskrive prisutvikling. Slike gjennomsnitt gir et godt mål på prisutviklingen når produktene man sammenligner er like (homogene). Vektede gjennomsnitt gir derimot ikke et godt bilde på prisutviklingen når man sammenligner produkter som er av forskjellig typer. En endring i gjennomsnittet vil da ikke bare kunne oppstå som følge av at prisene på produktene endres, men også som følge av at sammensetningen av produkter endrer seg. I litteraturen er det utviklet flere ulike mål på slike sammensetningseffekter, men disse målene kan vise sammensetningseffekter også for produkter med uendret volum. I denne artikkelen utledes en formel som eksakt dekomponerer endringen i det vektede gjennomsnittet som en sum av en priskomponent og en sammensetningskomponent. Metoden identifiserer bidraget til den samlede sammensetningskomponenten fra hvert enkelt produkt og, i motsetning til rammeverkene brukt i litteraturen, gir metoden kun sammensetningseffekter for produkter med endret volum.

Et godt eksempel på bruken av vektede gjennomsnitt er i produksjon av kvartalsvis statistikk for lønn. Vi anvender metoden for å dekomponere endringen i det vektede gjennomsnittet for gjennomsnittlig avtalt månedslønn fra 1. kvartal 2020 til 1. kvartal 2021. Dekomponeringen gjøres med hensyn på næringer, og viser at vår dekomponering avviker fra de vanlige dekomponeringene brukt i litteraturen.

1. Introduction

What are the driving forces underlying aggregate productivity growth? Why has the labour force participation rate changed during the last two decades? What has driven the change in annual earnings over the last year and why have import prices changed? All these questions have a common feature in that statistics for productivity, the labour force participation rate, earnings and import prices are often constructed using a weighted arithmetic mean formula.

A natural starting point for answering these questions is to decompose the change in the weighted mean. A frequently used decomposition is the Bennet (1920) decomposition, often also referred to as shift-share analysis. This decomposition enables within-group growth effects to be distinguished from between-group compositional effects. For example, when examining productivity dynamics in U.S. manufacturing plants between 1972 and 1987, Baily et al. (1992) find a positive contribution to growth due to increasing output shares among high-productivity plants and decreasing output shares among low-productivity plants. Daly & Hobijn (2017) show that compositional effects due to labour market status flows are important in explaining aggregate real wage growth in the U.S. Analysing the fall in the U.S. labour force participation rate, Krueger (2017) finds that the population composition has shifted toward groups with lower participation rates, and that this accounts for well over half of the decline in the labour force participation rate between 1997 and 2017. Moreover, a large body of literature has identified the deflationary effects of international trade resulting from increased import shares from low-price countries, such as China; see e.g., Kamin et al. (2006), Thomas and Marquez (2009) and Benedictow and Boug (2017, 2021).

Although the Bennet decomposition is useful for identifying the overall contribution from compositional effects, it does not identify how much of the change in overall compositional effects can be attributed to a particular group or subset. To overcome this shortcoming, the Bennet decomposition is often rewritten by subtracting a scalar A from each group in the between effect, where the scalar A typically represents some measure of the weighted mean. Huerga (2010) labelled this decomposition, when A represents the average of the weighted means between two consecutive time periods, the Marshall-Edgeworth-type decomposition with extended weight effect. Foster et al. (2001) analyzes productivity developments and measures the between effect as the product of changes in the plant-level output share and the deviation of average plant-level productivity from the overall industry average. If the composition of firms changes such that the output share of a low-productivity plant increases, this would lower the aggregate weighted mean productivity level and thus contribute negatively to the compositional effect. Note that in these decompositions a plant may contribute negatively to the compositional effect even if there is no change in the output of that plant. The reason is that it is the output share, and not the output of the plant, that enters the decomposition, and the output share of a given plant may change because the output of all the other plants changes. Moreover, as pointed out by Balk (2003), the choice of scalar A is arbitrary. Since any scalar may be subtracted from the Bennet decomposition, an infinite number of possible decompositions are available, and it is therefore an open question which of the possible decompositions should be applied.

In this paper, we derive a decomposition of the weighted mean that is rooted in functional analysis. Following the lines of the index number literature, our key idea is to apply the quadratic approximation lemma (QAL) to the weighted mean. The QAL provides an exact decomposition of a quadratic function, where each component represents the contribution of a change in a single independent variable to the overall change in the dependent variable. This lemma dates at least back to Theil (1967, p. 222) and has been used extensively to decompose price and volume indices; see e.g. Diewert (2002) and references therein.

In our proposed decomposition, the weighted mean is regarded as a two-stage function: first, as a function of weights and indicators and second, as a function of weights being non-linear in the underlying volume variable. For example, in terms of the productivity decomposition referred to above, the weight is the output share of a given plant and the underlying volume variable is the output of that plant. The weighted mean productivity level is thus a composite function of plant-specific output (volume variable) and productivity level (indicator). Applying QAL to both stages of the weighted mean yields our proposed decomposition. In this decomposition, all the terms related to the within effects are identical to those in the Bennet decomposition. Also, the overall between effect, or compositional effect, is identical to the overall between effect in the Bennet decomposition. The group-specific between effects, however, differ from those in the Bennet decomposition. In our proposed decomposition, the group-specific between effect is a function of the change in the underlying volume variable. The decomposition captures the intuitive property that the weighted mean increases if a group whose volume variable is growing has a level that is above the weighted mean level. There are two ways in which compositional effects for a group will be zero: either the group-specific indicator equals the weighted mean level, and/or there is no change in the volume variable of that group. The proposed decomposition is easy to employ and interpret and furthermore gives a better platform for comparing groups.

Moreover, we show that the decomposition is invariant with respect to treatment of time and that it therefore satisfies the difference counterpart to the index number time reversal test; see ILO et al. (2004, p. 411).

Closely related to our analysis is the literature studying the unit value bias; see e.g. Párniczky (1974), Silver (2009) and Diewert and Lippe (2010). This literature decomposes the *ratio* between the unit value index, which is based on the weighted mean formula, and some well-known price indices, such as the Laspeyres, Paasche, and Fisher indices. In contrast to our proposed decomposition, which identifies the group-specific contribution to the overall compositional effect, this literature is concerned with identifying the overall compositional effect, referred to as the "unit value bias". These decompositions are moreover *multiplicative*, i.e. the unit value index is written as the product of a standard price index and the derived bias term for compositional effects. Our proposed decomposition is additive.

To illustrate our proposed decomposition, we use data on aggregate earnings growth in Norway between 2020Q1 and 2021Q1. We find that the wedge between the identified compositional effects from our decomposition and the Bennet decomposition is substantial, and for some industries, the compositional effects are of opposite signs.

The paper proceeds as follows: Section 2 outlines the weighted mean formula, some of the most standard decompositions applied in the literature and our proposed decomposition. Section 3 contrasts and compares empirically our proposed decomposition with those used in the literature, using the case of earnings growth in Norway. Section 4 provides a conclusion.

2. Decomposing the weighted mean

Our point of departure is the weighted mean of indicators P_{it} across units i at time t of the form:

$$P_t = \sum_{i=1}^{N} S_{it} P_{it}, \tag{1}$$

with weights $S_{it} = \frac{X_{it}}{\sum_{j=1}^{N} X_{jt}}$, where the volume variable $X_{it} \geq 0$ and $\sum_{j=1}^{N} X_{jt} > 0$. Note that the weights sum to unity. The weighted mean in Equation (1) has numerous applications within the fields of economics and measurement theory. Although the weighted mean has been applied in a variety of fields, and the indicator and volume variables may refer to *inter alia* wages, hours worked, productivity, output prices etc., we will henceforth refer to P_{it} and X_{it} as representing prices and quantities, respectively, and unit i as product i. In the following we are concerned with identifying the contribution to the change in the weighted mean of a change in both prices and quantities. Before we present our proposed decomposition, we start by recapitulating the most widely utilized decompositions in the literature.

The Bennet decomposition

Bennet (1920) provided a decomposition of the nominal value change into the sum of a price change and a quantity change. This decomposition stands in contrast to traditional index theory, which focuses on decomposing a value ratio into the product of a price index and a quantity index. Diewert (2005) analyzed the axiomatic and economic properties of the Bennet decomposition. The Bennet decomposition applied to Equation (1) yields:

$$\Delta P = \sum_{i=1}^{N} \bar{S}_i \Delta P_i + \sum_{i=1}^{N} \bar{P}_i \Delta S_i, \tag{2}$$

where Δ is the difference operator and a bar over a variable represents the moving average operator between time t and v, i.e. $\Delta x = x_t - x_v$ and $\bar{x} = 1/2(x_t + x_v)$, and the time subscript is dropped when it is superfluous, for notational convenience. The Bennet decomposition is standard in productivity and shift-share analysis, see e.g. Baily et al. (1992) and OECD (2018). The terms $\bar{S}_i \Delta P_i$ and $\bar{P}_i \Delta S_i$ represent the contribution to the change in the weighted mean of a change in the price of product i, and the quantity of product i, respectively. In contrast to the case which Bennet studied, where S_{it} represented quantities, the variable S_{it} in Equation (1) is a weight and thus a non-linear function of quantities. This distinction is crucial for the interpretation of the contribution to the change in the weighted mean of a change in quantities. For example, when the quantity of a low-price product increases, one would intuitively think that this would lead to a lowering of the weighted mean price level, as shown in the case of the unit value bias; see e.g. Diewert and Lippe (2010). However, Equation (2) does not identify such an effect, since the term which shows the contribution of the increased share of product i, $\bar{P}_i \Delta S_i$, is always positive, regardless of the price level of product i.

The Marshall-Edgeworth decomposition with extended weight effect

To capture the fact that the weighted mean decreases when the quantity of a low-priced product increases, the decomposition in Equation (2) can be changed accordingly; see also Balk (2003). Since the weights sum to unity, we may subtract the term $\sum_{i=1}^{N}(A\, \Delta S_i)$, for any given scalar A, such that:

$$\Delta P = \sum_{i=1}^{N} \bar{S}_i \Delta P_i + \sum_{i=1}^{N} (\bar{P}_i - A)\, \Delta S_i. \tag{3}$$

In this case, the contribution to the change in the weighted mean of a change in the quantity of product i is given by the term $(\bar{P}_i - A)\,\Delta S_i$. This term captures the fact that the weighted mean price level increases if products whose quantity shares are growing have an average price level for product i between time v and t that is larger than the scalar A. Note that the quantity share of product i (S_{it}) may change even if there is no change in the quantity of product i (X_{it}), i.e. if there is a change in the sum of all the other products. That said, by choosing the scalar A to represent some measure of the mean price level, the above framework provides the same qualitative contribution from compositional effects as identified in the case of the unit value bias; see e.g. Diewert & Lippe (2010). For example, the choice $A = \bar{P}$, where $\bar{P} = 1/2(P_t + P_v)$, yields the decomposition Huerga (2010) labelled the Marshall-Edgeworth-type decomposition with extended weight effect. The contribution to the change in the weighted mean of a change in the share of unit i is then given by $(\bar{P}_i - \bar{P})\,\Delta S_i$. The weighted mean price level will thus increase if products whose quantity shares are growing have a price level that is higher than the weighted mean price level.[1] Conversely, the weighted mean price level will decrease if products whose quantity shares are growing have a price level that is lower than the weighted mean price level. Although such a choice of scalar A yields a decomposition property that fits qualitatively well with the corresponding results for the unit value bias, the choice of scalar A is completely arbitrary, as argued by Balk (2003).

The Bennet decomposition and the quadratic approximation lemma

A different way to interpret the Bennet decomposition in Equation (2) is through functional analysis. Consider a function $y_t = F(\mathbf{x}_t, \mathbf{z}_t)$, where $\mathbf{x}_t = (x_{1t}, x_{2t}, \dots, x_{Nt})$ and $\boldsymbol{z}_t = (z_{1t}, z_{2t}, \dots, z_{Nt})$. In the following we are concerned with identifying the effect on the change in y due to changes in $\mathbf{x}_t$ and $\boldsymbol{z}_t$. To this end, we apply the quadratic approximation lemma (QAL). The QAL provides a second-order approximation to a non-linear function F, and according to Theil (1975), the quadratic approximation lemma provides an "approximation which is as simple as the linear approximation and as accurate as the quadratic approximation" (p. 38).[2] Diewert (1976) showed that the QAL holds exactly for any two points $(\mathbf{x}_t, \mathbf{z}_t)$ and $(\mathbf{x}_v, \mathbf{z}_v)$ if, and only if, the function F is quadratic. Let $\boldsymbol{F}_{x,t}$ denote the vector of first-order partial derivatives with respect to $\mathbf{x}$ evaluated at $(\mathbf{x}_t, \mathbf{z}_t)$, i.e. $\boldsymbol{F}_{x,t} = (\partial F/\partial x_{1t}, \dots, \partial F/\partial x_{Nt})$, and $\boldsymbol{F}_{z,t}$ denote the vector of first-order partial derivatives with respect to $\mathbf{z}$ evaluated at $(\mathbf{x}_t, \mathbf{z}_t)$, i.e. $\boldsymbol{F}_{z,t} = (\partial F/\partial z_{1t}, \dots, \partial F/\partial z_{Nt})$. Furthermore, let

---

[1] Note that when choosing $\bar{P} = A$, the framework above is invariant with respect to treatment of time, i.e. it satisfies the difference counterpart to the index number time reversal test, see Diewert and Fox (2010).

[2] Where the term "quadratic approximation" refers to a second-order Taylor approximation.

$\overline{\boldsymbol{F}_x} = 1/2[\boldsymbol{F}_{x,t} + \boldsymbol{F}_{x,v}]$ and $\overline{\boldsymbol{F}_z} = 1/2[\boldsymbol{F}_{z,t} + \boldsymbol{F}_{z,v}]$. If and only if the function F is quadratic, the following identity holds:

$$\Delta y = \overline{\boldsymbol{F}_x}\,\Delta\mathbf{x} + \overline{\boldsymbol{F}_z}\,\Delta\mathbf{z}. \tag{4}$$

Henceforth we refer to the quadratic identity in Equation (4) as the quadratic approximation lemma (QAL). The QAL provides a decomposition of a quadratic function in which each component represents the contribution of a change in a single independent variable to the overall change in y_t.

Now consider the weighted mean formula in Equation (1), which is quadratic in the variables $\mathbf{S}_t = (S_{1t}, S_{2t}, \dots, S_{Nt})$ and $\mathbf{P}_t = (P_{1t}, P_{2t}, \dots, P_{Nt})$. Applying QAL to Equation (1) yields Equation (2). We may thus think of the Bennet decomposition as being the result of applying QAL to the quadratic function $P = F(\mathbf{S}_t, \mathbf{P}_t)$. However, since the variables in $\mathbf{S}_t$ in Equation (1) are weights and thus a non-linear function of quantities, the case of the weighted mean in Equation (1) is somewhat more complex than the quadratic function $P = F(\mathbf{S}_t, \mathbf{P}_t)$.

The weighted mean and the QAL

Since Equation (4) holds as an identity for quadratic functions, it has been used extensively to decompose price and volume indices; see e.g. Diewert (2002) and references therein. Following the lines of the index number literature, the key idea in this paper is to apply the QAL to the weighted mean to identify the contributions from both price and quantity changes. As mentioned in the introduction, we represent the weighted mean as a two-stage function. First, the weighted mean is a function of prices and quantity shares, $P_t = F(\mathbf{S}_t, \mathbf{P}_t)$. Second, the quantity shares are functions of quantities, and of the sum of quantities, $\mathbf{S}_t = \mathbf{G}(\mathbf{X}_t, Q(\mathbf{X}_t)) = (G_1(X_{1t}, Q(\mathbf{X}_t)), \dots, G_N(X_{Nt}, Q(\mathbf{X}_t)))$, where $\mathbf{X}_t = (X_{1t}, X_{2t}, \dots, X_{Nt})$ represents the vector of quantities, the function Q represents the sum of quanties, i.e. $Q(\mathbf{X}_t) = \sum_i X_{it}$, and the function G_i is the quantity share of product i, i.e. $G_i = X_i/Q_t$. The weighted mean is therefore a composite function of both quantities ($\mathbf{X}_t$) and prices ($\mathbf{P}_t$), i.e. $P_t = F(\boldsymbol{G}(\mathbf{X}_t, Q(\mathbf{X}_t)), \mathbf{P}_t)$. This two-stage setup shows what happens to the weighted mean when quantities change, giving rise to a "chain reaction" in two stages. First, the weights, $\mathbf{S}_t = \mathbf{G}(\mathbf{X}_t, Q(\mathbf{X}_t))$, react directly to the change in quantities, owing both to changes in the quantity of product i and of the aggregate Q. Second, the weighted mean P_t reacts to the change in weights.

To analytically decompose both steps of this "chain reaction", we first consider the function of quantity shares for product i, $S_{it} = G_i(X_{it}, Q(\mathbf{X}_t)) = X_{it}/Q(\mathbf{X}_t)$. We are interested in identifying how much of the change in the share ΔS_i can be attributed to the change in X_{it}, and how much can be attributed to the change in the sum of quantities Q. This question may be answered by considering the inverse function, i.e. $X_{it} = S_{it}Q$. Note that this inverse function (G_i^{-1}) conveys the same information as the G_i-function, since these functions represent a one-to-one relationship for the set of all non-negative numbers. Instead of considering how much of the change in S_i can be attributed to changes in X_i and Q, we can therefore first consider the inverse function and decompose the change in X_i that can be attributed to S_i and Q, and then back out how much X_i and Q contribute to the change in S_i. Applying the QAL to the quadratic function $X_i = S_iQ$ yields the exact decomposition:

$$\Delta X_i = \bar{Q}\Delta S_i + \bar{S_i}\Delta Q. \tag{5}$$

The terms $\bar{Q}\Delta S_i$ and $\bar{S_i}\Delta Q$ capture how much the variables S_i and Q , respectively, contribute to the change in X_i. Note that Equation (5) shows the possible discrepancy between the change in the quantity variable ΔX_i and the change in the weight ΔS_i. In particular, and as we will return to below, the sign of ΔX_i may be the opposite to that of ΔS_i, depending on how much the aggregate quantity (Q) changes. Further, it follows from Equation (5) that the change in the share S_i can be exactly decomposed as:

$$\Delta S_i = \left(\frac{1}{\bar{Q}}\right)\Delta X_i - \left(\frac{\bar{S_i}}{\bar{Q}}\right)\Delta Q. \tag{6}$$

The first term after the equality sign, $\left(\frac{1}{\bar{Q}}\right)\Delta X_i$, captures how much of the change in S_i can be attributed to the change in X_i, while the last term, $-\left(\frac{\bar{S_i}}{\bar{Q}}\right)\Delta Q$, captures how much can be attributed to the change in Q. Since $\Delta Q = \sum_{i=1}^{N} \Delta X_i$, we have:

$$\Delta S_i = \left(\frac{1-\bar{S_i}}{\bar{Q}}\right)\Delta X_i - \left(\frac{\bar{S_i}}{\bar{Q}}\right)\sum_{\substack{j=1 \\ j\neq i}}^{N} \Delta X_j. \tag{7}$$

Equation (7) represents the first part of the “chain reaction”. It shows that the share S_i changes both because the quantity of product i changes (the first term after the equality sign) and because the quantity of the other products ($j \neq i$) changes (the second term after the equality sign). The second part of the “chain reaction” is given by Equation (2), which shows how the weighted mean P_t reacts to the change in weights as a result of applying the QAL to Equation (1). Inserting Equation (7) into Equation (2) and collecting terms yields the following exact decomposition of the weighted mean:

Proposition 1 (Exact additive decomposition of the weighted mean)

Consider the weighted mean across units i at time t of the form: $P_t = \sum_{i=1}^{N} S_{it} P_{it}$, with weights $S_{it} = \frac{X_{it}}{\sum_{j=1}^{N} X_{jt}}$, where $X_{it} \geq 0$ and $Q_t = \sum_{j=1}^{N} X_{jt} > 0$. The change in the weighted mean between time t and v can be exactly decomposed as

$$\Delta P = \sum_{i=1}^{N} \bar{S}_i \Delta P_i + \sum_{i=1}^{N} \left(\frac{1}{\bar{Q}}\right) \left(\bar{P}_i - \bar{\bar{P}}\right) \Delta X_i \tag{8}$$

where $\bar{\bar{P}} = \sum_{i=1}^{N} \bar{S}_i \bar{P}_i$, Δ is the difference operator and a bar over a variable represents the moving average operator between time t and v, i.e. $\Delta x = x_t - x_v$ and $\bar{x} = 1/2(x_t + x_v)$.

PROOF. See the Appendix.

Several features of the decomposition in Proposition 1 merit attention. First, the term that shows the contribution to the change in the weighted mean of the change in the price of product i, $\bar{S}_{it} \Delta P_i$, is identical to the term in the Bennet decomposition shown in Equation (2). Second, the aggregate term $\sum_{i=1}^{N} \left(\frac{1}{\bar{Q}}\right) \left(\bar{P}_i - \bar{\bar{P}}\right) \Delta X_i$ that shows the total compositional effect, or the contribution to the change in the weighted mean of the sum of all quantity changes, is identical to the term for the compositional effect in the Bennet decomposition. Third, the term that shows the contribution to the change in the weighted mean of the change in the quantity of product i, is given by

$$\left(\frac{1}{\bar{Q}}\right) \left(\bar{P}_i - \bar{\bar{P}}\right) \Delta X_i.$$

This term differs from that in the Bennet decomposition. It has a natural interpretation and captures the intuitive property that the weighted mean price level increases if products that are growing in quantity have price levels that are higher than the mean price level. $\bar{P}_i - \bar{\bar{P}}$ compares the price level of product i with a measure of the weighted mean price level $\bar{\bar{P}} = \sum_{i=1}^{N} \bar{S}_i \bar{P}_i$. There are thus two ways in which the compositional effects of product i can equal zero: the price of product i equals the weighted average price level, and/or there is no change in the quantity of product i.

A fourth distinctive feature of the decomposition in Proposition 1 is that it does not hold a time subscript. In other words, the framework is invariant with respect to treatment of time and it therefore satisfies the difference counterpart to the index number time reversal test. The time reversal test for indices states that if the data for the two time periods are interchanged, then the resulting formula should equal the reciprocal of the original index, see e.g. ILO et al. (2004, p. 295). This test can be rephrased in the case where the formula is in the form of differences, such as the decomposition in Proposition 1: if the data for the two time periods are interchanged, then the resulting formula should equal the negative of the original formula. To illustrate this analytically, let the function $H(\mathbf{P}_t, \mathbf{P}_v, \mathbf{X}_t, \mathbf{X}_v)$ represent the formula for decomposing the change in the weighted mean. The function H passes the time reversal test if and only if $H(\mathbf{P}_t, \mathbf{P}_v, \mathbf{X}_t, \mathbf{X}_v) = -H(\mathbf{P}_v, \mathbf{P}_t, \mathbf{X}_v, \mathbf{X}_t)$. The proposed decomposition in Proposition 1 satisfies this counterpart to the time reversal test.

We commented above on the practice in the literature of choosing a scalar A when decomposing the weighted mean, see Equation (3). Although the choice of A is arbitrary, it is nevertheless interesting to see whether it is possible to derive a value for A that is consistent with the decomposition in Proposition 1. From Equation (3), the contribution to the change in the weighted mean from a change in the quantity share of product i is given by the term $(\bar{P}_{it} - A)\Delta S_{it}$. In Proposition 1, the contribution to the change in the weighted mean of a change in the quantity of product i is given by the term $\left(\frac{1}{\bar{Q}}\right)\left(\bar{P}_{it} - \bar{\bar{P}}_t\right)\Delta X_i$. For these terms to be equal, the scalar A must be given by (see the Appendix):

$$A_i = \bar{\bar{P}} - \left(\frac{\Delta Q/\bar{Q}}{\Delta S_i/\bar{S}_i}\right)\left(\bar{P}_{it} - \bar{\bar{P}}\right) \quad (9)$$

The derived value of A_i depends on i. This feature stands in contrast to Equation (3), where the property that A is a scalar and independent of i is central to deriving Equation (3) from Equation (2). In the case where the aggregate quantity is unchanged, i.e. $\Delta Q = 0$, Equation (9) reduces to $A = \bar{\bar{P}}$, which is independent of i. Moreover, in this case the value of A is close to the choices commonly used in the literature. Several values for the scalar A have been applied, most frequently P_t, P_v and the average of the two, which are all close to the average measure $\bar{\bar{P}}_t$. However, when the aggregate quantity changes, $\Delta Q \neq 0$, the factor $\left(\frac{\Delta Q/\bar{Q}}{\Delta S_i/\bar{S}_i}\right)$ may differ from zero, possibly leaving a sizable discrepancy between the decomposition in Proposition 1 and the most common decompositions applied in the literature. In particular, and as can be seen from Equation (5), the sign of ΔX_i may be the opposite of the sign of ΔS_i, depending on how much aggregate quantity (Q) changes. As a result, the measured contributions from compositional effects in Equation (3) and Proposition 1 may have opposite signs. In the empirical section, we examine in depth how large the discrepancy between the two decompositions may be in practice when aggregate earnings growth in Norway is decomposed.

3. Empirical application

The data used in the empirical application are obtained through the "a-ordning", which is a collaborative digital system shared by Statistics Norway, the Norwegian Tax Administration and the Norwegian Labour and Welfare Administration (NAV). It provides information about employment, remuneration in cash and in kind and taxes. Data for all industries and individuals are collected and compiled monthly, and this is the main source Statistics Norway utilizes for producing statistics on earnings and the labor market.

We focus on the change in monthly basic earnings per full-time equivalent as the price variable from 2020Q1 to 2021Q1 and allow for compositional effects across industries using the number of jobs in each industry as the volume variable. Table 1 shows the mean monthly basic earnings and the number of jobs in each industry and in the aggregate for 2020Q1 and 2021Q1.

Table 1 Monthly basic earnings per full-time equivalent and number of jobs, 2020Q1 and 2021Q1

	2020Q1		2021Q1	
	Monthly basic earnings (NOK)	Number of jobs	Monthly basic earnings (NOK)	Number of jobs
All industries	44,982	2,892,481	46,258	2,808,076
Agriculture, forestry and fishing	39,630	31,375	41,720	31,730
Mining and quarrying	64,060	62,326	65,080	61,218
Manufacturing	45,930	220,366	47,140	213,042
Electricity, water supply, sewerage, waste management	51,950	33,669	53,380	34,129
Construction	43,440	239,733	44,790	237,298
Wholesale and retail trade; repair of motor vehicles and motorcycles	40,980	367,772	42,390	364,197
Transportation and storage	44,040	142,947	45,270	128,863
Accommodation and food service activities	32,010	114,097	33,450	73,881
Information and communication	58,540	100,137	60,390	102,262
Financial and insurance activities	62,410	47,532	64,100	48,408
Real estate, professional, scientific and technical activities	56,470	174,818	58,090	171,674
Administrative and support service activities	39,270	159,059	40,220	144,862
Public administration and defence; compulsory social security	49,690	184,955	50,210	187,848
Education	46,410	254,153	46,820	252,616
Human health and social work activities	41,580	641,702	42,110	647,022
Other service activities	42,060	117,840	43,750	109,026

Source: Statbank Table 11654, Statistics Norway.

Table 2 shows the results from using the Bennet decomposition in Equation (2), the Marshall-Edgeworth type decomposition with extended weight effect in Equation (3), and our proposed decomposition in Proposition 1.[3] As expected, the contribution to the change in the weighted mean from the change in earnings of each industry (and the aggregate) is identical across the three decompositions, as is the total compositional effect. We find that the wedge between the identified compositional effects from (i) our decomposition and (ii) the Bennet decomposition and the Marshall-Edgeworth decomposition is considerable, and for some industries such as mining and quarrying, construction and wholesale and retail trade, the compositional effects are of opposite signs. The compositional effects from each industry from the three different decompositions are illustrated in Figure 1.

Figure 2 illustrates that the discrepancies between the measured contributions from compositional effects are due to the changes in both the share and the volume variable that are, for some industries, of opposite signs, see e.g. mining and quarrying, construction and wholesale and retail trade. As discussed earlier and shown in Equation (9), this leads to a discrepancy between the decompositions.

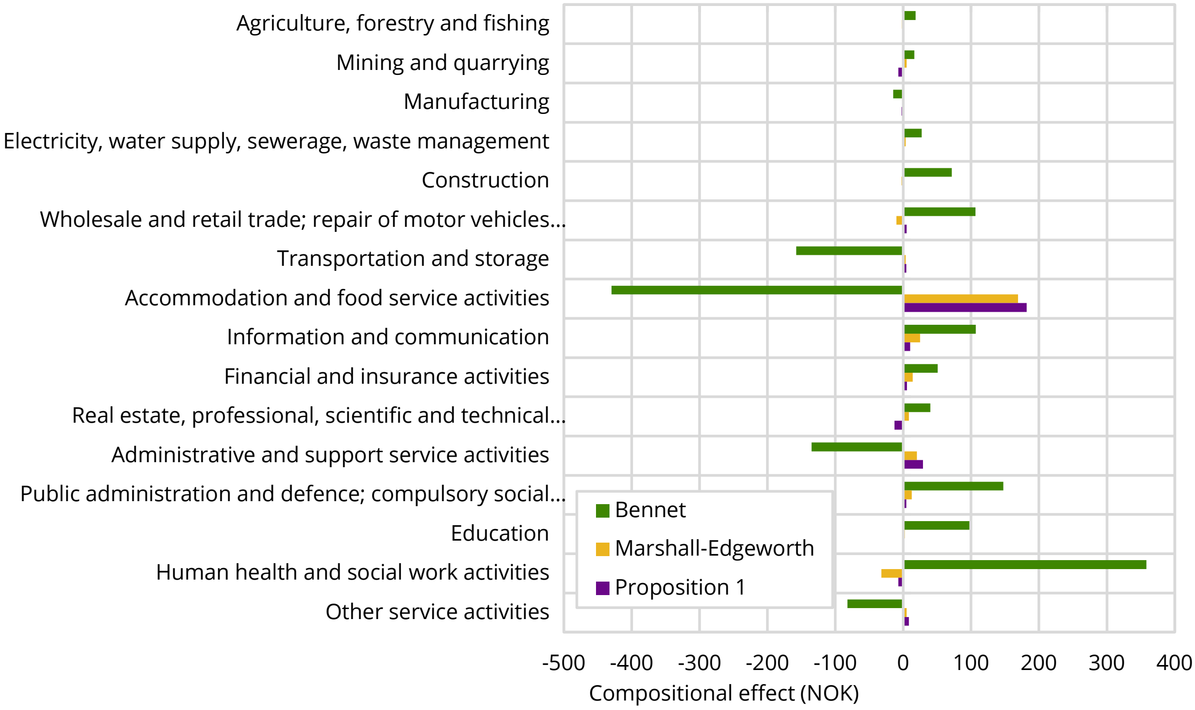
[3] An add-in to undertake this decomposition in EViews, and a Stata replication code to generate the results in Table 2 and Figure 1, are available from the authors upon request.

Table 2 Decomposition of change in monthly basic earnings, from 2020Q1 to 2021Q1

	Bennet decomposition			Marshall-Edgeworth decomposition			Decomposition in Proposition 1		
	Earnings contribution	Compositional effect	Total	Earnings contribution	Compositional effect	Total	Earnings contribution	Compositional effect	Total
All industries	1,055	221	1,276	1,055	221	1,276	1,055	221	1,276
Agriculture, forestry and fishing	23	18	42	23	-2	21	23	-1	23
Mining and quarrying	22	16	38	22	5	27	22	-7	15
Manufacturing	92	-15	77	92	0	92	92	-2	90
Electricity, water supply, sewerage, waste management	17	27	44	17	4	21	17	1	18
Construction	113	72	185	113	-2	111	113	1	114
Wholesale and retail trade; repair of motor vehicles and motorcycles	181	106	287	181	-10	171	181	5	186
Transportation and storage	59	-158	-99	59	3	62	59	5	63
Accommodation and food service activities	47	-430	-383	47	169	217	47	182	229
Information and communication	66	107	173	66	25	91	66	10	76
Financial and insurance activities	28	51	79	28	14	43	28	5	34
Real estate, professional, scientific and technical activities	98	40	138	98	8	107	98	-13	86
Administrative and support service activities	51	-135	-85	51	20	71	51	29	80
Public administration and defence; compulsory social security	34	147	181	34	13	47	34	4	38
Education	36	98	134	36	2	39	36	-1	36
Human health and social work activities	120	358	478	120	-32	88	120	-7	113
Other service activities	67	-82	-15	67	5	72	67	8	76

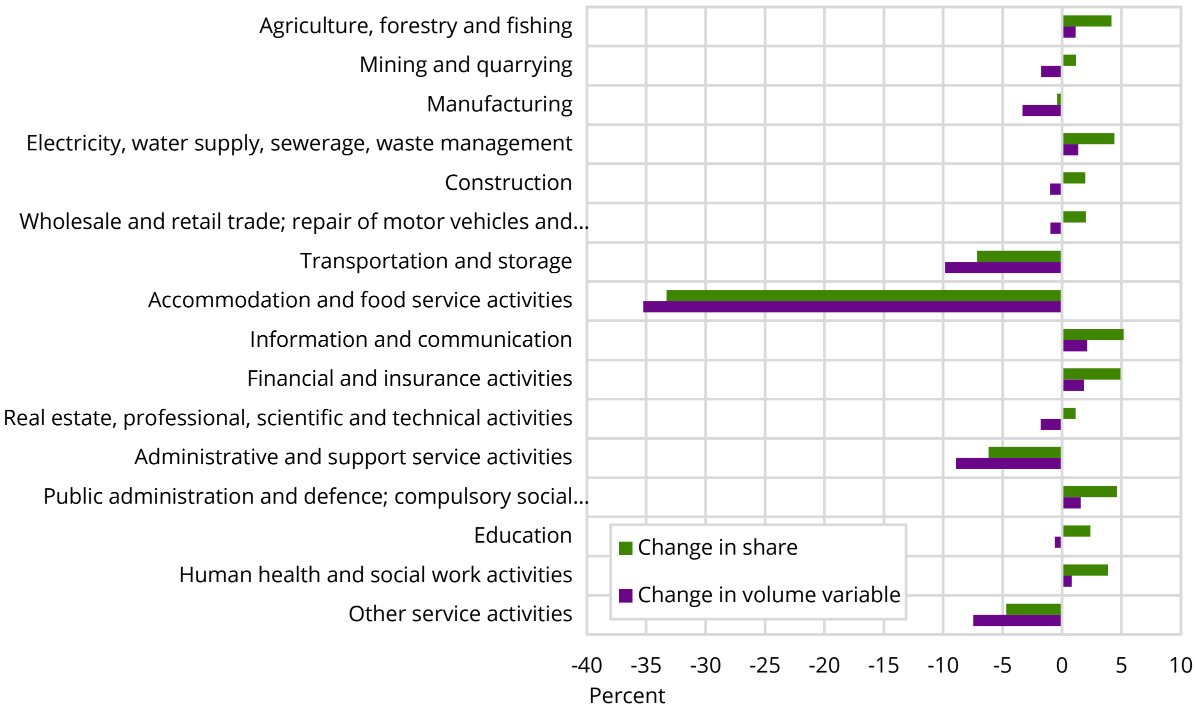
Source: Authors' own calculations using data from Statistics Norway.

Figure 1 Compositional effects across decompositions[1]



[1]See Table 2 for precise magnitudes of compositional effects for each industry and decomposition method.
Source: Authors' own calculations using data from Statistics Norway.

Figure 2 Change in share and volume variable, from 2020Q1 to 2021Q1[1]



[1]Change in share and volume variable for each industry from 2020Q1 to 2021Q1, measured in percent.
Source: Authors' own calculations using data from Statistics Norway.

4. Conclusion

In this paper, we have derived an exact additive decomposition of the weighted mean that is rooted in functional analysis. Our proposed decomposition is easy to employ and interpret. We also show that it satisfies the difference counterpart to the index number time reversal test. The fundamental difference between our proposed decomposition and many of the applied decompositions used in the literature is that our measure of the contribution to compositional changes of a given product is based on the change in the quantity of that product: If there is no change in the quantity of a product, then that product does not contribute to a compositional change in the weighted mean. In contrast, in the decompositions employed in the literature, the measure of the contribution to compositional changes of a given product is based on the change in the quantity share of that product. Since the quantity share of a product may change because the quantities of other products change, this may lead to compositional changes stemming from a product whose quantity level is unchanged. When comparing our proposed decomposition to the standard decomposition applied in the literature in the case of aggregate earnings growth in Norway from 2020Q1 to 2021Q1, we find that the wedge between the identified compositional effects is substantial, and for some industries the compositional effects are of opposite signs.

5. References

Baily, M. N., Hulten, C., Campbell, D., Bresnehan, T., & Caves, R. E. (1992). Productivity dynamics in manufacturing plants. *Brookings Papers on Economic Activity: Microeconomics*, 187–267. https://doi.org/10.2307/2534764

Balk, B. M. (2003). The Residual: On Monitoring and Benchmarking Firms, Industries, and Economies with Respect to Productivity. *Journal of Productivity Analysis*, *20*(1), 5–47. https://doi.org/10.1023/A:1024817024364

Benedictow, A., & Boug, P. (2017). Calculating the real return on a sovereign wealth fund. *Canadian Journal of Economics*, *50*(2), 571–594. https://doi.org/10.1111/caje.12270

Benedictow, A., & Boug, P. (2021). Exact and inexact decompositions of trade price indices. *Empirical Economics*, (January 2018). https://doi.org/10.1007/s00181-021-02078-4

Bennet, T. L. (1920). The Measurement of Changes in the Cost of Living. *Journal of the Royal Statistical Society*, *83*(3), 455–462. https://doi.org/10.2307/2340777

Daly, M. C., & Hobijn, B. (2017). Composition and Aggregate Real Wage Growth. *American Economic Review*, *107*(5), 349–352. https://doi.org/10.1257/aer.p20171075

Diewert, W. E. (1976). Exact and superlative index numbers. *Journal of Econometrics*, *4*, 115–145.

Diewert, W. E. (2002). The quadratic approximation lemma and decompositions of superlative indexes. *Journal of Economic and Social Measurement*, *28*, 63–88.

Diewert, W. E. (2005). Index number theory using differences rather than ratios. *American Journal of Economics and Sociology*, *64*(1), 311–360. https://doi.org/10.1111/j.1536-7150.2005.00365.x

Diewert, W. E., & Fox, K. J. (2010). On measuring the contribution of entering and exiting firms to aggregate productivity growth. In W. E. Diewert, B. . Balk, D. Fixler, K. J. Fox, & A. Nakamura (Eds.), *Price and Productivity Measurement* (Vol. 6, pp. 41–66). Trafford Press.

Diewert, W. E., & Lippe, P. von der. (2010). Notes on Unit Value Index Bias. *Jahrbücher Für Nationalökonomie & Statistik*, *230*(6), 690–708.

Foster, L., Haltiwanger, J. C., & Krizan, C. J. (2001). Aggregate Productivity Growth: Lessons from Microeconomic Evidence. In C. R. Hulten, E. R. Dean, & M. J. Harper (Eds.), *New Developments in Productivity Analysis* (pp. 303–372). University of Chicago Press.

Huerga, J. (2010). An Application of Index Numbers Theory to Interest Rates. In L. Biggeri & G. Ferrari (Eds.), *Price Indexes in Time and Space: Methods and Practice* (pp. 239–248). https://doi.org/10.1007/978-3-7908-2140-6_13

ILO, IMF, OECD, Eurostat, United Nations, & World Bank. (2004). *Consumer price index manual - Theory and practice*. (I. L. Office, Ed.). Geneva.

Kamin, S. B., Marazzi, M., & Schindler, J. W. (2006). The Impact of Chinese Exports on Global Import Prices. *Review of International Economics*, *14*(2), 179–201. https://doi.org/10.1111/j.1467-9396.2006.00569.x

Krueger, A. B. (2017). Where have all the workers gone? An inquiry into the decline of the U.S. Labor force participation rate. *Brookings Papers on Economic Activity*, *2017*(Fall), 1–87.

https://doi.org/10.1353/eca.2017.0012

OECD. (2018). *OECD Compendium of Productivity Indicators 2018*. OECD. https://doi.org/10.1787/pdtvy-2018-en

Párniczky, G. (1974). Some Problems of Price Measurement in External Trade Statistics. *Acta Oeconomica*, *12*(2), 229–240.

Silver, M. (2009). An index number formula problem: The aggregation of broadly comparable items. *Journal of Official Statistics*, *27*(4), 553–567.

Theil, H. (1967). *Economics and information theory*. North-Holland (Amsterdam).

Theil, H. (1975). *Theory and measurement of consumer demand*. Vol(1), North-Holland (Amsterdam).

Thomas, C. P., & Marquez, J. (2009). Measurement matters for modelling US import prices. *International Journal of Finance and Economics*, *14*(2), 120–138. https://doi.org/10.1002/ijfe.370

6. Appendix

Proof of Proposition 1

Inserting $\Delta Q = \sum_{i=1}^{N} \Delta X_i$ into Equation (6) yields:

$$\Delta S_i = a_i \Delta X_i - b_i \sum_{\substack{j=1 \\ j \neq i}}^{N} \Delta X_j,$$

where $a_i = \frac{1-\bar{S}_i}{\bar{Q}}$ and $b_i = \frac{\bar{S}_i}{\bar{Q}}$. We thus have:

$$\sum_{i=1}^{N} \bar{P}_i \Delta S_i = \sum_{i=1}^{N} \bar{P}_i \left(a_i \Delta X_i - b_i \sum_{\substack{j=1 \\ j \neq i}}^{N} \Delta X_j \right).$$

This can be written as:

$$\begin{gathered} \bar{P}_1 a_1 \Delta X_1 - \bar{P}_1 b_1 \Delta X_2 - \bar{P}_1 b_1 \Delta X_3 - \cdots - \bar{P}_1 b_1 \Delta X_N \\ +\bar{P}_2 a_2 \Delta X_2 - \bar{P}_2 b_2 \Delta X_1 - \bar{P}_2 b_2 \Delta X_3 - \cdots - \bar{P}_2 b_2 \Delta X_N + \cdots \\ +\bar{P}_N a_N \Delta X_N - \bar{P}_N b_N \Delta X_1 - \bar{P}_N b_N \Delta X_2 - \cdots - \bar{P}_{N-1} b_{N-1} \Delta X_{N-1} \end{gathered}$$

When collecting terms, this can be written as:

$$\begin{gathered} \bar{P}_1 a_1 \Delta X_1 - \bar{P}_2 b_2 \Delta X_1 - \bar{P}_3 b_3 \Delta X_1 - \cdots - \bar{P}_{Nt} b_N \Delta X_1 \\ +\bar{P}_2 a_2 \Delta X_2 - \bar{P}_1 b_1 \Delta X_2 - \bar{P}_3 b_3 \Delta X_2 - \cdots - \bar{P}_N b_N \Delta X_2 \\ +\bar{P}_3 a_3 \Delta X_3 - \bar{P}_1 b_1 \Delta X_3 - \bar{P}_2 b_3 \Delta X_3 - \cdots - \bar{P}_N b_N \Delta X_3 \\ + \cdots \end{gathered}$$

This in turn can be written more compactly as:

$$\sum_{i=1}^{N} \left(\bar{P}_i a_i - \sum_{\substack{j=1 \\ j \neq i}}^{N} b_j \bar{P}_j \right) \Delta X_i.$$

By inserting $a_i = \frac{1-\bar{S}_i}{\bar{Q}}$ and $b_i = \frac{\bar{S}_i}{\bar{Q}}$, we get:

$$\sum_{i=1}^{N} \left(\frac{1}{\bar{Q}} \right) \left(\bar{P}_i - \sum_{j=1}^{N} \bar{S}_j \bar{P}_j \right) \Delta X_i,$$

which equals the second term after the equals sign in Proposition 1:

$$\sum_{i=1}^{N} \left(\frac{1}{\bar{Q}} \right) \left(\bar{P}_i - \bar{\bar{P}} \right) \Delta X_i,$$

where $\bar{\bar{P}} = \sum_{i=1}^{N} \bar{S}_i \bar{P}_i$.

Derivation of the scalar A

From Equation (3), the contribution to the change in the weighted mean from a change in the quantity of product i is given by the term $(\bar{P}_i - A_i)\Delta S_i$. Setting this term equal to the term $\left(\frac{1}{\bar{Q}}\right)\left(\bar{P}_i - \bar{\bar{P}}\right)\Delta X_i$ yields:

$$(\bar{P}_i - A_i)\Delta S_i = \left(\frac{1}{\bar{Q}}\right)\left(\bar{P}_i - \bar{\bar{P}}\right)\Delta X_i.$$

Solving for A yields:

$$A_i = \bar{\bar{P}} - \left(\frac{\Delta X_i}{\bar{Q}\Delta S_i}\right)\left(\bar{P}_i - \bar{\bar{P}}\right).$$