Source Decomposition of Inequality Indices*

Vesa-Matti Heikkuri[1] and Matthias Schief[2]

[1]Tampere University, FIT
[2]Economics Department, OECD

Abstract

This paper revisits the longstanding question of how to decompose inequality indices into contributions of factor components, such as labor income, capital income, and taxes and transfers. We start from the axiomatic framework of Shorrocks (1982) and discuss why one of his core axioms, two factor symmetry, imposes an undesirable restriction. We propose an alternative axiom, conditional transfer neutrality, to replace two factor symmetry, and we show that our axiomatic framework determines a unique and index-specific source decomposition for any inequality index. This decomposition coincides with the natural decomposition for many inequality indices, including the Gini coefficient, the coefficient of variation, and top income shares. For other inequality indices, such as the Theil index and the mean log deviation, our axiomatic framework implies novel source decompositions that we argue behave better than the traditional formulas. Furthermore, we show how the source decomposition of the Gini coefficient can be interpreted geometrically.

I Introduction

A longstanding question in the literature on inequality measurement is how to decompose inequality indices into contributions of particular sources, such as labor income, capital income, and taxes and transfers. Several approaches to answer this question have been proposed in the literature, but no consensus has been reached so far.

*Address correspondence to matthias.schief@oecd.org and vesa-matti.heikkuri@tuni.fi. We thank Tommaso Coen for valuable comments and suggestions. We also thank seminar participants at the Finnish Economic Association Annual Meeting 2026, the Institute of Social and Economic Research at Osaka University, the University of Turku, Tampere University, the Labour Institute for Economic Research Labore, the World Inequality Conference 2026, the Second International Conference on Inequalities and Opportunities, and the Normative Economics and Economic Policy webinar. This research benefited from funding provided by the Finnish Centre of Excellence in Tax Systems Research funded by the Research Council of Finland (project 364187).

A prominent paper in the literature is Shorrocks (1982), which introduces a list of axioms that any source decomposition should satisfy and shows that under this framework, all inequality indices should be decomposed according to the variance decomposition formula. Many researchers see this result as unsatisfactory, as the decomposition does not reflect the properties of the inequality index (Fournier, 2001; Sastre and Trannoy, 2002; Paul, 2004; Cowell and Fiorio, 2011; Chantreuil and Trannoy, 2013).

An alternative paradigm is provided by Lerman and Yitzhaki (1985), who define the contribution of a source as the change in total inequality resulting from a marginal proportional increase in the incomes of the source. Decompositions derived using this principle are called *marginal decompositions*. While Lerman and Yitzhaki (1985) derive marginal decompositions for the Gini coefficient and its generalizations, Paul (2004) extends these results to generalized entropy indices. Although marginal decompositions are widely applied, their drawback is that they do not yield relative contributions of sources since the contributions sum to zero rather than to total inequality.

A third approach is based on the Shapley value (Chantreuil and Trannoy, 2013; Shorrocks, 2013). In this decomposition, the contribution of a source equals its average marginal effect over all possible elimination sequences. A disadvantage of the Shapley decomposition is that it does not satisfy the *independence of aggregation* axiom in Shorrocks (1982). For example, the contribution of labor income may change if, instead of considering capital income as a single source, we disaggregate it into interest income, dividend income, etc.

In this paper, we build an axiomatic theory for source decomposition of inequality that yields a unique decomposition formula for all inequality indices satisfying basic regularity conditions. Our starting point is the list of six axioms in Shorrocks (1982). We argue that a key axiom in this list, *two factor symmetry*, imposes an undesirable restriction and leads to counterintuitive source decompositions. We propose an alternative axiom, *conditional transfer neutrality*, to replace two factor symmetry, and we show that our axiomatic framework leads to a unique and well-behaved source decomposition for any differentiable or rank-dependent inequality index. This decomposition coincides with the *natural decomposition* for many inequality indices, including the Gini coefficient, the coefficient of variation, and top income shares.[1] For other inequality indices, such as the Theil index and the mean log deviation, our axiomatic framework implies novel source decompositions that we argue behave better than the traditional formulas.

[1]Shorrocks (1982) coined the term natural decomposition. The natural decomposition for the Gini coefficient has been suggested earlier by Rao (1969) and Fei et al. (1978).

The recent empirical literature on inequality typically measures inequality using top shares (Piketty et al., 2018; Smith et al., 2023; Auten and Splinter, 2024; Alstadsæter et al., 2025). Papers in this literature often decompose top shares by source using ad hoc approaches with intuitive appeal. Our axiomatic framework can also be applied to top shares and delivers unique source decomposition formulas that coincide with the way many papers have decomposed these measures. Hence, our results can also be read as an axiomatic justification for these approaches.

Finally, we show how the natural decomposition of the Gini coefficient can be interpreted geometrically using the concentration curve.

The paper is structured as follows. Section II introduces notation. Section III introduces the axiomatic framework, discusses why the two factor symmetry axiom of Shorrocks (1982) should be rejected, and proposes conditional transfer neutrality as a more desirable axiom. Section IV derives our main result on the unique source decomposition of inequality indices. Section V offers a geometric interpretation for the source decomposition of the Gini coefficient. Section VI concludes.

II Notation and definitions

We consider a population consisting of $n \in \mathbb{N}$ individuals. Let $Y = (y_1, \ldots, y_n)$ denote the vector of total income, where $y_i > 0$ is the income of individual i.[2] Total income consists of $K \in \mathbb{N}$ sources and we use $Y^k = (y_1^k, \ldots, y_n^k)$ to denote the vector of source $k = 1, \ldots, K$ incomes, where $y_i^k \in \mathbb{R}$ is the source-k income of individual i. That is, $y_i = \sum_{k=1}^{K} y_i^k$ for all $i = 1, \ldots, n$.

An inequality index $I\colon \mathbb{R}^n_{++} \to \mathbb{R}_+$ is a function that maps a vector of incomes to a non-negative real number and satisfies the standard axioms of anonymity, scale independence, population independence, normalization, and the Pigou–Dalton principle of transfers.[3] We assume that I is continuously differentiable, and use $I_i(Y) = \partial I(Y)/\partial y_i$ to denote its partial derivatives.[4]

[2]We call Y degenerate if $y_i = y$ for all i. We consider strictly positive incomes because some common inequality indices, such as the Theil index, are not defined for zero or negative incomes. However, our results go through under zero or negative incomes for inequality indices that are defined in this case as long as the mean of total income is positive.

[3]The Pigou–Dalton principle states that a transfer from a richer to a poorer individual that does not reverse their income ranking reduces inequality. We additionally impose a mild regularity condition: for any nondegenerate income distribution, a marginal transfer from the richest to the poorest individual strictly reduces inequality to first order. Equivalently, this requires that $I_n(Y) > I_1(Y)$ for any nondegenerate Y, where individuals are ordered by income.

[4]We relax this assumption in Section IV.1.

A source decomposition of inequality is a collection of source contributions $\{S_k(Y^1,\ldots,Y^K)\}_{k=1}^K$ where $S_k(Y^1,\ldots,Y^K)$ is the contribution of source k to the inequality index $I(Y)$.

III An axiomatic framework

In a seminal paper, Shorrocks (1982) proposes a set of axioms for source decomposition:

Axiom 1 (Symmetry)**.** $S_k(Y^{p_1},\ldots,Y^{p_K}) = S_{p_k}(Y^1,\ldots,Y^K)$ *for any permutation* $(p_1,\ldots,p_K)$ *of* $(1,\ldots,K)$ *and every* $k=1,\ldots,K$*.*

Axiom 2 (Independence of aggregation)**.** $S_1(Y^1,\ldots,Y^K) = S_1(Y^1, Y-Y^1) =: S(Y^1,Y)$*.*

Axiom 3 (Continuity)**.** $S(Y^k,Y)$ *is continuous in* Y^k *for fixed* Y*.*

Axiom 4 (Consistency)**.** $\sum_{k=1}^K S(Y^k,Y) = I(Y)$*.*

Axiom 5 (Normalization)**.** $S(Y^k,Y)=0$ *if* Y^k *is a constant vector.*

These axioms impose basic requirements on the decomposition rule. Symmetry ensures that the labels of sources are irrelevant. Independence of aggregation ensures that the contribution of a particular source is unaffected when the remaining income is split into further subsources. Continuity is a regularity condition requiring that small changes in income distributions lead to small changes in assigned contributions. Consistency is an accounting requirement that the source contributions add up to total inequality, which enables the computation of relative source contributions. Axiom 5 is a normalization convention, requiring that a perfectly equally distributed source contributes neither positively nor negatively to inequality in total income.[5]

Shorrocks (1982) shows that Axioms 1–4 imply that the source decomposition of any inequality index must take the following form,

$$S(Y^k,Y) = \sum_{i=1}^n a_i(Y) y_i^k, \quad \text{where } I(Y) = \sum_{i=1}^n a_i(Y) y_i. \tag{1}$$

That is, when the inequality index is written as a weighted sum of individual incomes, the contribution of source k is equal to the weighted sum of individual source-k incomes,

[5]Note that Axioms 4 and 5 can be satisfied only if the inequality index satisfies a normalization axiom, that is, $I(Y)=0$ if $y_i=\mu$ for all i.

using the same weights. At first sight, this decomposition has some intuitive appeal for inequality measures that are commonly written as a weighted sum of individual incomes, since it is straightforward to substitute $\sum_{k=1}^{K} y_i^k$ for y_i and switch the order of summation to obtain $I(Y) = \sum_{k=1}^{K} \sum_{i=1}^{n} a_i(Y) y_i^k = \sum_{k=1}^{K} S(Y^k, Y)$. This is often referred to as the *natural source decomposition* of inequality indices. However, as Shorrocks (1982) demonstrates, this decomposition is arbitrary, since there are infinitely many ways of writing a given inequality index as a weighted sum of individual incomes where the weights are allowed to depend on the entire income distribution.[6] It is thus clear that additional restrictions need to be placed on the source decomposition to resolve this indeterminacy.

III.1 Revisiting two factor symmetry

To obtain a unique source decomposition, Shorrocks (1982) adds an additional axiom, which he calls *two factor symmetry*.

Two Factor Symmetry. *For all permutation matrices* P, $S(Y^1, Y^1 + Y^1 P) = S(Y^1 P, Y^1 + Y^1 P)$.

Two factor symmetry states that when total income consists of two sources that are permutations of each other, then the two sources contribute equally to total inequality. Suppose Y^2 is a permutation of Y^1. To motivate the axiom, Shorrocks starts by noting that if there are other sources of income, then we would not want to impose that factors 1 and 2 make the same contribution to inequality in total incomes. This is because the arrangement of incomes in vector Y^1 may correspond more closely to the distribution of incomes from other sources, $Y - Y^1$, and hence provide more reinforcement to total inequality, than Y^2 does with $Y - Y^2$. Shorrocks then continues to argue that when there are only two factors that are permutations of each other, everything is symmetric and therefore we should want to impose that both factors make the same contribution to inequality in total incomes. It is this second step where we disagree.

While it is true, as Shorrocks argues, that any symmetric correspondence between Y^1 and $Y - Y^1 = Y^2$ is the same as the correspondence between Y^2 and $Y - Y^2 = Y^1$, source decomposition measures how each source contributes to inequality in total income. For that purpose, the relevant comparison is not between Y^1 and Y^2, but between each

[6] For example, $I(Y) = \sum_{i=1}^{n} (I(Y)/n\mu) y_i$, where μ is the mean income, is a trivial but feasible representation for any index. In addition, for any function f, $a_i(Y) = I(Y)(f(y_i) - \bar{f})/(\sum_{j=1}^{n} (f(y_j) - \bar{f}) y_j)$, where $\bar{f}$ is the arithmetic average of $f(y_j)$, satisfies $I(Y) = \sum_{i=1}^{n} a_i(Y) y_i$ and the corresponding natural decomposition satisfies Axioms 1–5.

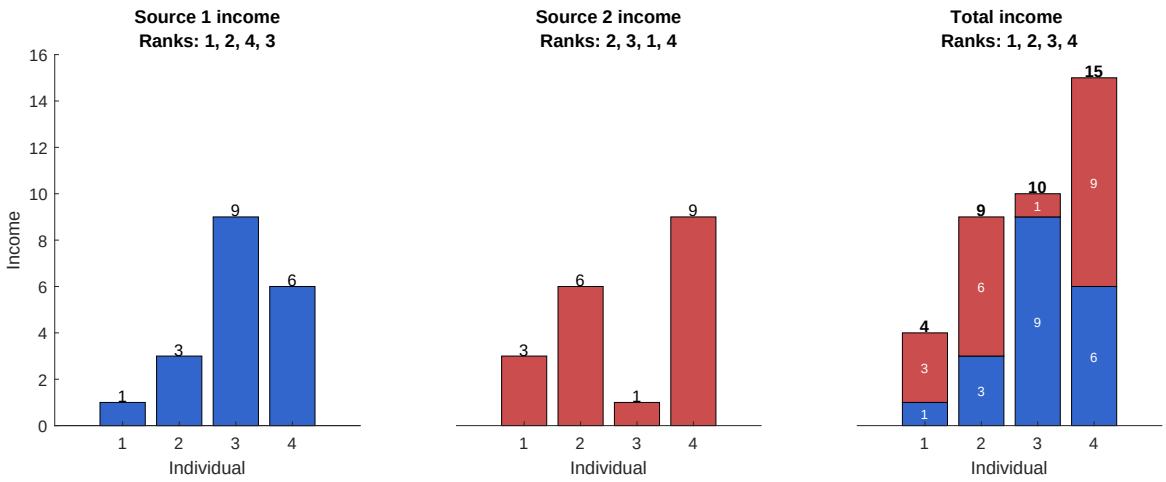


Figure 1: Income distributions of Example 1.

source and $Y = Y^1 + Y^2$. Despite being permutations of each other, Y^1 and Y^2 may not be identical in their relationship to the distribution of total income, and it is therefore not obvious that Y^1 and Y^2 should have to contribute identically to inequality in total incomes.

Example 1 illustrates this point.

Example 1. Consider the setup where the population consists of four individuals and total income consists of two sources with distributions $Y^1 = (1, 3, 9, 6)$ and $Y^2 = (3, 6, 1, 9)$, which imply $Y = (4, 9, 10, 15)$. Figure 1 illustrates these distributions. The distribution of source-1 incomes is a permutation of that for source 2, so the two sources have the same income share and the same level of inequality. Note that if we were to rank people by the distribution of source 1, only individuals 3 and 4 would have to switch positions, but if we rank individuals by source 2, individual 3 would first have to switch positions with individual 2 and then with individual 1. In this sense, the correspondence between source-1 income and total income may be considered "closer" than the correspondence between source-2 income and total income. Indeed, according to the Spearman correlation, source-1 incomes are more correlated with total income than source-2 incomes. This observation shows that symmetry between two permuted source distributions does not imply symmetry in their relationship to total income.

This potential asymmetry matters because, intuitively, the contribution of a source should depend on three factors: the income share of the source, how unequally the source incomes are distributed, and how the source incomes are aligned with total income. The higher the income share of the source, the higher the potential contribution. Similarly, the more unequally the source incomes are distributed, the more potential the

source has to contribute to inequality in total income. Finally, the more the source incomes are concentrated on the rich (poor) individuals, the larger the positive (negative) contribution of the source should be.

Importantly, how we evaluate both inequality and the alignment of the source incomes with total income may depend on which inequality index we use. For example, if we use the variance to measure inequality, then the Pearson correlation is a natural measure of alignment. If, however, we measure inequality with the top 1% income share, then only the allocation of source incomes between the richest one percent and the bottom 99 percent matters, not how correlated the source incomes are with total income within those groups.[7] It is then natural to expect the decomposition to reflect the behavior of the inequality index we use.

Example 1 illustrates this point as well. Consider the top 25% income share. Because the income of the richest 25% consists relatively more of source-2 income compared to the bottom 75%, it is reasonable to say that source 2 contributes more to inequality. Note, however, that we would conclude the opposite if we used the top 50% income share to measure inequality, since the top half earns relatively more source-1 income compared to the bottom half. Since the income share and the level of inequality are equal across sources, these different conclusions can only come from differences in how the sources are aligned with total income. In general, there is no reason to require that both sources should contribute equally to inequality in either case.

Two factor symmetry therefore imposes an undesirable restriction on source decomposition, and it has a strong formal consequence. The main result of Shorrocks (1982) is that requiring two factor symmetry together with Axioms 1–5 implies that any inequality index must be decomposed according to the variance decomposition formula. That is, $S(Y^k, Y)/I(Y) = \mathrm{Cov}(Y^k, Y)/\mathrm{Var}(Y)$ for any inequality index I, which is equal to the relative contribution of source k in the natural source decomposition of the variance. This result is unsatisfactory because different inequality indices evaluate inequality in a given distribution differently, yet the decomposition rule in Shorrocks (1982) does not reflect these differences.

Example 2 offers a concrete example of why the result of Shorrocks (1982) is counterintuitive.

Example 2. Suppose that the population consists of five individuals and total income

[7] We use top income shares in examples even though they do not satisfy the strict Pigou–Dalton principle and the normalization axiom, and are not continuously differentiable everywhere. We claim that these properties are not essential to develop intuition on the behavior of source decomposition. See also Section IV.2 for the treatment of top income shares.

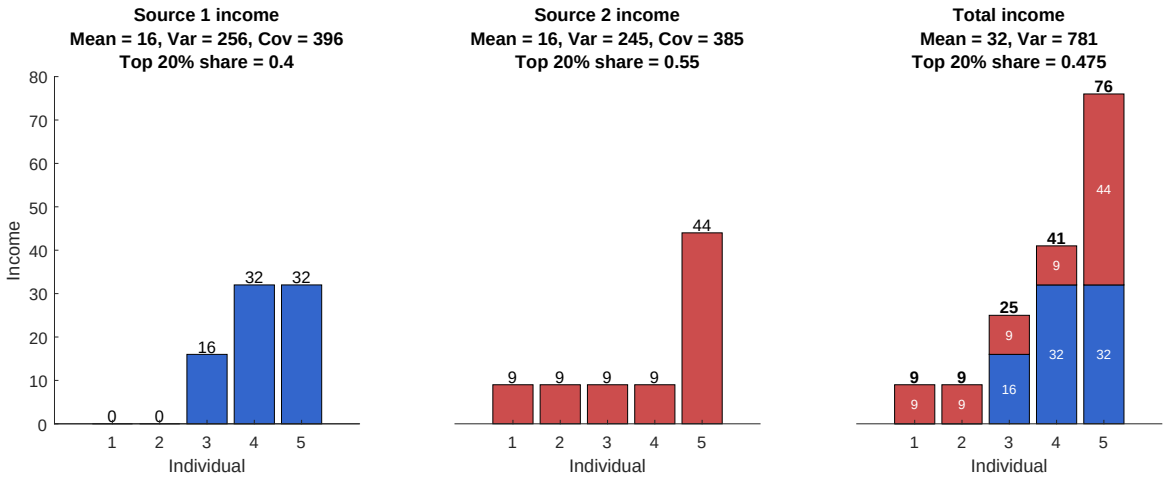


Figure 2: Income distributions of Example 2.

consists of two sources with distributions $Y^1 = (0, 0, 16, 32, 32)$ and $Y^2 = (9, 9, 9, 9, 44)$. The distribution of total income is then $Y = (9, 9, 25, 41, 76)$. These distributions are illustrated in Figure 2. Note first that the income share of each source is 50%. Since both the variance of Y^1 and the covariance between Y^1 and Y are higher than those of Y^2, it is natural to conclude that source 1 should contribute more to inequality in total income than source 2 if we measure inequality using the variance. This is indeed what the natural decomposition of the variance gives us. However, if we measure inequality using the top 20% income share, then the opposite conclusion is reasonable. The richest quintile earns relatively more source-2 income than the bottom 80%, and the bottom 80% earn relatively more source-1 income, so that intuitively source 2 should contribute more to the top 20% share of total income. Hence, more generally, it is reasonable that which source contributes more to inequality should depend on which inequality index we use.

III.2 Conditional transfer neutrality

Without two factor symmetry, Axioms 1–5 do not imply a unique source decomposition. In fact, Shorrocks (1982) demonstrates that for any inequality index, there are infinitely many source decompositions that are consistent with Axioms 1–5. Moreover, the contributions are unbounded. For any given distributions of total income and source incomes, the contribution of a particular source can generally be set to anything between negative and positive infinity by selecting an appropriate decomposition formula. Hence, we need an additional axiom to restrict the behavior of the decomposition.

Shorrocks (1983) recognizes that the indeterminacy in equation (1) means that natural source decompositions need not behave in desirable ways, and he suggests a simple monotonicity property to rule out some of the most severe pathologies.[8] Suppose we transfer a unit of source-k income from individual i to individual j and exactly compensate by transferring an equal amount of source-l income from j to i. If $y_i > y_j$, that is, if individual i is richer than j in total income, then this transfer scheme should decrease the contribution of source k and increase the contribution of source l. The intuition is that while the distribution of total income is unaffected, source-k incomes are more concentrated among the poor and source-l incomes are more concentrated among the rich after the transfers.

Shorrocks (1983) notes that this property is equivalent to requiring that the weight $a_i(Y)$ in the formula for the natural decomposition (equation (1)) is increasing in y_i. It rules out some decomposition formulas consistent with Axioms 1–4, such as the natural decomposition for the mean log deviation[9], but still leaves us with infinitely many possible decompositions. In particular, decomposing all inequality indices according to the natural decomposition of the variance remains feasible.

The reason why we do not get uniqueness under this property is that it is not specific to the inequality index. Any inequality index can be written as a weighted sum of incomes where the weight for individual i is increasing in y_i.[10] Thus, the property does not rule out using the contributions of the natural decomposition of the variance to decompose the Gini coefficient or the top income shares. As we saw in Example 2, this can lead to counterintuitive source contributions.

If we want the decomposition to reflect the behavior of the inequality index, we need an axiom that links the decomposition to the inequality index. A natural approach is to strengthen the property of Shorrocks (1983). Imposing monotonicity with respect to a single progressive transfer, as Shorrocks (1983) does, is not enough because all inequality indices decrease in response to progressive transfers. What distinguishes different

[8] See footnote 10 in Shorrocks (1983).

[9] The mean log deviation is usually written as $MLD(Y) = \frac{1}{n}\sum_{i=1}^{n} \ln(\mu/y_i)$. Written in the form $MLD(Y) = \sum_i a_i(Y) y_i$, this gives $a_i(Y) = \ln(\mu/y_i)/(ny_i)$ and hence the natural source contribution $S(Y^k, Y) = \frac{1}{n}\sum_i (y_i^k/y_i)\ln(\mu/y_i)$. Note that $a_i(Y)$ is largest for low-income individuals and becomes negative for individuals with above-mean income. Thus, for given total income of a particular source, the source's contribution is maximized when all source income is assigned to the poorest individual, which is clearly pathological behavior for a source decomposition. The property suggested in Shorrocks (1983) addresses this particular pathology, but it does not select a unique alternative decomposition for the mean log deviation.

[10] For any function f, $a_i(Y) = I(Y)(f(y_i) - \bar{f})/(\sum_j (f(y_j) - \bar{f})y_j)$, where $\bar{f}$ is the arithmetic average of $f(y_j)$, satisfies $I(Y) = \sum_{i=1}^{n} a_i(Y) y_i$ and the corresponding natural decomposition satisfies Axioms 1–4. If $f(y_i)$ is increasing in y_i, then $a_i(Y)$ is also increasing in y_i.

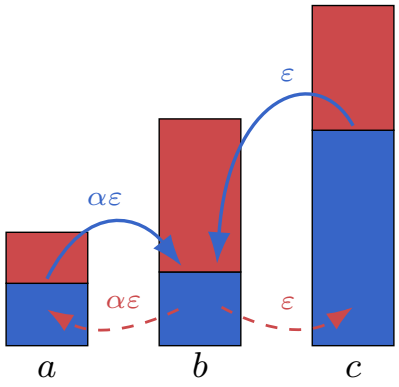


Figure 3: Illustration of compensated neutral transfers.

inequality indices is how they evaluate specific transfers. Thus, a minimal adjustment to Shorrocks' property that takes this evaluation into account is to consider a combination of a progressive transfer and a regressive transfer that leaves inequality unaffected. In particular, suppose that $y_a < y_b < y_c$ and that a combination of a progressive transfer of $\varepsilon > 0$ units from c to b and a regressive transfer of $\alpha\varepsilon \geq 0$ units from a to b leaves the level of inequality intact.[11] Now, if those transfers are implemented with source-k income while compensating opposite transfers are implemented with source-l income, then it is reasonable to require that the contributions of sources k and l do not change because the transfers are inequality-neutral. Figure 3 illustrates such transfers.

We can write this property formally as the following axiom, which we call *conditional transfer neutrality*.

Definition 1. *A transfer* $t \in \mathbb{R}^n$ *is an* inequality-neutral transfer *at* Y *for index* I *if* $\sum_{i=1}^n t_i = 0$ *and*

$$\frac{d}{d\varepsilon} I(Y + \varepsilon t)\Big|_{\varepsilon=0} = 0.$$

Axiom 6 (Conditional transfer neutrality)**.** *For all* k *and any inequality-neutral transfer* t*,*

$$\frac{d}{d\varepsilon} S(Y^k + \varepsilon t, Y)\Big|_{\varepsilon=0} = 0.$$

[11] For a given income distribution, individuals a, b, and c, and inequality index, the value of α is unique unless $I_a(Y) = I_b(Y) = I_c(Y)$. Note also that we allow for $\alpha = 0$, since for inequality indices that satisfy only a weak Pigou–Dalton principle, some progressive transfers do not affect inequality, so no regressive transfer is needed to offset their effect.

IV Results

The main result of this paper is that if we replace *two factor symmetry* with *conditional transfer neutrality*, then there is a unique source decomposition for each inequality index. Moreover, unlike in Shorrocks (1982), the decomposition rule is index-specific, yielding different decompositions for ordinally distinct indices. The result is stated in Theorem 1 for continuously differentiable inequality indices.

Theorem 1. *For any continuously differentiable inequality index I and for every non-degenerate Y, there is a unique decomposition that satisfies Axioms 1–6, where the contribution of source k equals*

$$S(Y^k, Y) = \sum_{i=1}^{n} a_i(Y) y_i^k,$$

with

$$a_i(Y) = \frac{I(Y)\big(I_i(Y) - \bar{I}(Y)\big)}{\sum_{j=1}^{n}\big(I_j(Y) - \bar{I}(Y)\big)y_j}, \tag{2}$$

where $I_i(Y)$ is the partial derivative of $I(Y)$ with respect to y_i and $\bar{I}$ is the arithmetic average of the first-order partial derivatives of I.

Proof. For a given nondegenerate income distribution Y, write $I_i = I_i(Y)$ and $\bar{I} = \frac{1}{n}\sum_{j=1}^{n} I_j$. Let e_i be the i-th standard basis vector of $\mathbb{R}^n$, $\mathbf{1} = (1, \ldots, 1)$, and define

$$D(Y) = \sum_{j=1}^{n}\big(I_j - \bar{I}\big)y_j.$$

Without loss of generality, suppose that $y_1 \leq y_2 \leq \cdots \leq y_n$. Because Y is nondegenerate, we have $y_1 < y_n$ and $I_1 < I_n$, and

$$D(Y) = \sum_{i=1}^{n}\big(I_i - \bar{I}\big)y_i = \frac{1}{2n}\sum_{i=1}^{n}\sum_{j=1}^{n}\big(I_i - I_j\big)\big(y_i - y_j\big) > 0,$$

since every term in the double sum is nonnegative and $(I_n - I_1)(y_n - y_1) > 0$. Hence, (2) is well-defined.

By Theorem 1 in Shorrocks (1982), Axioms 1–4 imply that the contribution of source k is linear in Y^k,

$$S(Y^k, Y) = \sum_{i=1}^{n} a_i(Y)\, y_i^k,$$

for some coefficient vector $a(Y) = \big(a_1(Y), \ldots, a_n(Y)\big)$ that satisfies $\sum_i a_i(Y) y_i = I(Y)$. Note that there are infinitely many such vectors. We proceed to show that equation (2) describes the unique vector such that $S(Y^k, Y)$ satisfies Axioms 1–6.

Axioms 4 and 5 imply two linear restrictions on $a(Y)$. By Axiom 4,

$$\sum_{i=1}^{n} a_i(Y)\, y_i = I(Y). \tag{3}$$

Taking source k to be equally distributed with a positive mean, Axiom 5 gives $0 = S(Y^k, Y) = \mu_k \sum_i a_i(Y)$, where μ_k is the mean of source-k income, i.e.

$$\sum_{i=1}^{n} a_i(Y) = 0. \tag{4}$$

We next show that Axiom 6 implies $n-2$ additional linear restrictions that together with equations (3) and (4) pin down $a(Y)$ uniquely. Consider the space of inequality-neutral transfers

$$W = \Big\{ t \in \mathbb{R}^n : \ \textstyle\sum_{i=1}^n t_i = 0 \ \text{ and } \ \sum_{i=1}^n t_i I_i = 0 \Big\} = \operatorname{span}\big(\mathbf{1}, \nabla I\big)^{\perp},$$

where ∇I is the gradient of I. Because $I_1 < I_n$, the vectors $\mathbf{1}$ and ∇I are linearly independent, so $\dim W = n - 2$.

For $m = 2, \ldots, n-1$ define

$$c_m = \frac{I_n - I_m}{I_n - I_1}, \qquad d_m = \frac{I_m - I_1}{I_n - I_1}, \qquad t^{(m)} = -c_m\, e_1 + e_m - d_m\, e_n.$$

Each $t^{(m)}$ defines a combination of transfers that leave both average income and $I(Y)$ unchanged to first order, so $t^{(m)} \in W$. Restricted to the coordinates $2, \ldots, n-1$ the vector $t^{(m)}$ equals e_m. Hence $t^{(2)}, \ldots, t^{(n-1)}$ are linearly independent and form a basis of W.

Implementing the transfer $t^{(m)}$ with source-k income and offsetting it with other sources leaves Y, and hence $a(Y)$, unchanged, so Axiom 6 gives

$$0 = \frac{d}{d\varepsilon}\, S\big(Y^k + \varepsilon\, t^{(m)},\, Y\big)\Big|_{\varepsilon=0} = \sum_{i=1}^{n} t_i^{(m)}\, a_i(Y), \qquad m = 2, \ldots, n-1.$$

As $\{t^{(m)}\}_{m=2}^{n-1}$ spans W, this yields $\sum_i t_i a_i(Y) = 0$ for all $t \in W$, that is, $a(Y) \perp W$.

Therefore

$$a(Y) \in W^{\perp} = \operatorname{span}\big(\mathbf{1}, \nabla I\big),$$

so $a_i(Y) = \beta + \gamma I_i$ for some scalars β, γ. Now, equation (4) forces $\beta = -\gamma \bar{I}$, i.e., $a_i(Y) = \gamma(I_i - \bar{I})$, and equation (3) forces $\gamma\, D(Y) = I(Y)$. Since $D(Y) > 0$, we obtain

$$a_i(Y) = \frac{I(Y)\big(I_i - \bar{I}\big)}{\sum_{j=1}^{n}\big(I_j - \bar{I}\big)y_j},$$

which is the expression in equation (2).

Finally, the decomposition defined by equation (2) satisfies all six axioms. It is additive with $a_i(Y) = \gamma(Y)(I_i - \bar{I})$, giving Axioms 1–3 because $S(Y^k, Y)$ is linear in y_i^k. Axiom 4 holds because $\sum_{k=1}^{K} S(Y^k, Y) = \sum_{i=1}^{n} a_i(Y) y_i = \gamma(Y) D(Y) = I(Y)$, and Axiom 5 because an equally distributed source gives $S(Y^k, Y) = \mu_k \gamma(Y) \sum_i (I_i - \bar{I}) = 0$. For Axiom 6, any transfer that leaves $I(Y)$ unchanged to first order has direction t with $\sum_i t_i = 0$ and $\sum_i t_i I_i = 0$, so the induced change in the contribution is $\sum_i a_i(Y) t_i = \gamma(Y) \sum_i (I_i - \bar{I}) t_i = 0$. Therefore (2) is the unique decomposition satisfying Axioms 1–6. □

Equivalently, the contribution in Theorem 1 can be written as

$$S(Y^k, Y) = I(Y) \frac{\operatorname{Cov}(Y^k, \nabla I)}{\operatorname{Cov}(Y, \nabla I)}$$

where ∇I is the gradient of I at Y. This expression clarifies that the contribution of source k depends on the alignment between source-k incomes and the inequality index's implicit sensitivity profile, as described by ∇I. A source contributes more when its incomes are concentrated among people whose incomes receive greater weight in the index. In particular, the contribution of source k is maximized (minimized) when source-k incomes are concentrated on the richest (poorest) individuals.[12]

Theorem 1 picks the unique decomposition that satisfies Axiom 6 from the infinite set of decompositions that satisfy Axioms 1–5. It coincides with the standard natural decomposition for the coefficient of variation but yields a new decomposition for the

[12] Moreover, because $a_i(Y)$ is increasing in y_i, the decomposition also satisfies the property in Shorrocks (1983), which requires that a progressive transfer implemented with source-k incomes that is compensated with transfers using other sources decreases the contribution of source k.

Theil index and the mean log deviation. The relative contributions can be written as

$$\frac{\mathrm{Cov}(Y^k, Y)}{\mathrm{Var}(Y)}, \quad \frac{\mathrm{Cov}(Y^k, \ln Y)}{\mathrm{Cov}(Y, \ln Y)}, \quad \frac{\mathrm{Cov}(Y^k, 1/Y)}{\mathrm{Cov}(Y, 1/Y)},$$

for the coefficient of variation, Theil index, and the mean log deviation, respectively.

Finally, Theorem 1 also implies that ordinally equivalent inequality indices yield the same relative source contributions, as the following corollary shows.

Corollary 1 (Ordinal invariance)**.** *Let I and J be two ordinally equivalent inequality indices satisfying the assumptions in Section II, so that $J = h \circ I$ for some strictly increasing, continuously differentiable function h with $h' > 0$. Then, for every nondegenerate Y, the relative contribution of each source k is identical under both indices:*

$$\frac{S^I(Y^k, Y)}{I(Y)} = \frac{S^J(Y^k, Y)}{J(Y)}.$$

Proof. By the chain rule, $\nabla J(Y) = h'(I(Y))\nabla I(Y)$. Hence,

$$\frac{S^J(Y^k, Y)}{J(Y)} = \frac{\mathrm{Cov}(Y^k, \nabla J)}{\mathrm{Cov}(Y, \nabla J)} = \frac{\mathrm{Cov}(Y^k, \nabla I)}{\mathrm{Cov}(Y, \nabla I)} = \frac{S^I(Y^k, Y)}{I(Y)}.$$

□

The intuition for Corollary 1 follows from conditional transfer neutrality. Ordinally equivalent inequality indices identify the same inequality-neutral transfers. Thus, conditional transfer neutrality imposes identical restrictions on the relative source contributions under both indices, while consistency only affects the overall scale of the contributions. Other approaches to source decomposition, such as the Shapley-value approach, generally do not satisfy ordinal invariance.

IV.1 Rank-dependent inequality indices

Many common inequality indices, such as the Gini coefficient and its generalizations, depend on income ranks. These indices are not differentiable at distributions that display income ties, that is, if $y_i = y_j$ for some $i \neq j$. In this section, we show that the same decomposition rule as in Theorem 1 applies for rank-dependent inequality indices when we allow for nondifferentiability at income ties and define the partial derivatives of I appropriately. Moreover, it is the unique decomposition that satisfies Axioms 1–6, when Axiom 6 is slightly modified to accommodate rank-dependent indices.

For this purpose, we define $Y = (y_1, \ldots, y_n)$ as the vector of income where individuals are indexed so that $y_1 \leq y_2 \leq \cdots \leq y_n$. When two or more individuals have the same total income, their given indexing fixes a tie-breaking order. We assume that I is continuously differentiable for any fixed ordering of individuals, and use $I_i(Y) = \partial I(Y)/\partial y_i$ to denote its partial derivatives. When some incomes coincide, $I_i(Y)$ denotes the one-sided derivative consistent with the prevailing indexing of individuals.[13]

We also impose a weaker version of the Pigou–Dalton principle that allows the inequality index to remain unchanged under progressive or regressive transfers. However, for any nondegenerate income distribution, a marginal progressive transfer from the richest to the poorest individual must strictly decrease inequality. Excess top income shares (see Section IV.2) satisfy the weak Pigou–Dalton principle but not the strong one.

Because rank-dependent inequality indices are not differentiable everywhere, we need a slightly different version of Axiom 6 to make sure that the marginal inequality-neutral transfers are admissible. In particular, we restrict Axiom 6 to apply for transfers that do not re-rank individuals.

Definition 2. *Let Y be indexed so that $y_1 \leq \cdots \leq y_n$. We call a transfer $t \in \mathbb{R}^n$ a* rank-preserving transfer *if $\sum_{i=1}^n t_i = 0$ and $t_i \leq t_{i+1}$ whenever $y_i = y_{i+1}$. Moreover, t is* inequality-neutral *at Y for index I if*

$$\frac{d}{d\varepsilon} I\big(Y + \varepsilon t\big)\Big|_{\varepsilon = 0^+} = 0.$$

With this definition, we can state Axiom 6 for rank-dependent inequality indices as follows:

Axiom 6 (Conditional transfer neutrality for rank-dependent indices)**.** *For all k and for any inequality-neutral rank-preserving transfer t,*

$$\frac{d}{d\varepsilon} S\big(Y^k + \varepsilon t, Y\big)\Big|_{\varepsilon = 0^+} = 0.$$

Theorem 2 shows that the same decomposition rule as in Theorem 1 is the unique source decomposition for rank-dependent inequality indices under this setup.

[13] Equivalently, the partial derivatives of I, which exist whenever $y_i \neq y_j$ for all $i \neq j$, extend continuously as incomes approach a vector with ties without changing their ordering, and $I_i(Y)$ is this one-sided limit. The Gini coefficient and its generalizations satisfy this assumption.

Theorem 2. *For any rank-dependent inequality index I and for every nondegenerate Y, there is a unique decomposition that satisfies Axioms 1–6, where the contribution of source k equals*

$$S(Y^k, Y) = \sum_{i=1}^{n} a_i(Y) y_i^k,$$

with

$$a_i(Y) = \frac{I(Y)\big(I_i(Y) - \bar{I}(Y)\big)}{\sum_{j=1}^{n} \big(I_j(Y) - \bar{I}(Y)\big) y_j}, \tag{5}$$

where $I_i(Y)$ is the partial derivative of $I(Y)$ with respect to y_i and $\bar{I}$ is the arithmetic average of the first-order partial derivatives of I.

Proof. When there are no income ties, i.e., when $y_1 < y_2 < \cdots < y_n$, I is continuously differentiable and equation (5) defines the unique source decomposition by Theorem 1.

Let Y now be an arbitrary nondegenerate income vector, indexed so that $y_1 \leq y_2 \leq \cdots \leq y_n$, possibly with ties. Since Theorem 1 in Shorrocks (1982) does not require differentiability, Axioms 1–5 give us

$$S(Y^k, Y) = \sum_{i=1}^{n} a_i(Y)\, y_i^k,$$

for some coefficient vector $a(Y) = \big(a_1(Y), \ldots, a_n(Y)\big)$ that satisfies $\sum_i a_i(Y) y_i = I(Y)$ and $\sum_i a_i(Y) = 0$.

Write $I_i = I_i(Y)$, $\bar{I} = \frac{1}{n}\sum_{i=1}^{n} I_i$, and let

$$W = \left\{ t \in \mathbb{R}^n : \ \textstyle\sum_{i=1}^{n} t_i = 0 \ \text{ and } \ \sum_{i=1}^{n} t_i I_i = 0 \right\}$$

be the space of inequality-neutral transfer schemes, as in the proof of Theorem 1. The Pigou–Dalton principle implies $I_1 \leq I_2 \leq \cdots \leq I_n$, and nondegeneracy gives $I_1 < I_n$. Hence, $\mathbf{1}$ and ∇I are linearly independent and $\dim W = n - 2$.

Let P denote the set of rank-preserving transfers for the given Y. If t is rank-preserving, then $Y + \varepsilon t$ has the same indexing as Y for small $\varepsilon > 0$, so I agrees with its continuously differentiable extension to $y_1 \leq \cdots \leq y_n$ along the path, and

$$\frac{d}{d\varepsilon} I(Y + \varepsilon t)\Big|_{\varepsilon = 0^+} = \sum_{i=1}^{n} t_i I_i.$$

A transfer $t \in W \cap P$ therefore leaves inequality unchanged to first order. Implementing it with source-k income and offsetting it with the other sources leaves Y, and hence

$a(Y)$, unchanged, so $S(Y^k + \varepsilon t, Y)$ is linear in ε and Axiom 6 yields

$$0 = \frac{d}{d\varepsilon} S\big(Y^k + \varepsilon t, Y\big)\Big|_{\varepsilon=0^+} = \sum_{i=1}^{n} t_i\, a_i(Y). \tag{6}$$

Our strategy is to construct $n-2$ linearly independent rank-preserving elements of W to show that $a(Y) \perp W$. The argument then follows as in the proof of Theorem 1. Note that we cannot use the schemes used in the proof of Theorem 1 as they are not guaranteed to be rank-preserving.

Since Y is nondegenerate, there exists $q \in \{1, \ldots, n-1\}$ such that $y_q < y_{q+1}$. For $1 \le k \le n-1$, define

$$\tau^{(k)} = -(n-k) \sum_{i \le k} e_i + k \sum_{i > k} e_i,$$

that is, $\tau^{(k)}$ is a mean-preserving regressive transfer from the poorest k individuals to the richest $n-k$ individuals. For $m = 1, \ldots, q-1, q+1, \ldots, n-1$, let

$$t^{(m)} = w(q)\, \tau^{(m)} - w(m)\, \tau^{(q)},$$

where

$$w(k) = \sum_{i>k} \big(I_i - \bar{I}\big) = \frac{1}{n} \sum_{i \le k} \sum_{j > k} \big(I_j - I_i\big) > 0.$$

Then $t^{(m)}$ combines a regressive transfer $\tau^{(m)}$ with a progressive transfer $-\tau^{(q)}$, weighted so that inequality is unchanged to first order. In particular, $\sum_{i=1}^n \tau_i^{(k)} I_i = \sum_{i=1}^n \tau_i^{(k)}(I_i - \bar{I}) = (n-k)w(k) + kw(k) = nw(k)$, so $\sum_{i=1}^n t_i^{(m)} I_i = nw(q)w(m) - nw(m)w(q) = 0$. Thus, $t^{(m)} \in W$ for all $m \neq q$. Moreover, $t^{(m)}$ is rank-preserving, because $\tau_{i+1}^{(k)} - \tau_i^{(k)} = n\mathbf{1}_{\{i=k\}}$, and so

$$t_{i+1}^{(m)} - t_i^{(m)} = nw(q)\mathbf{1}_{\{i=m\}} - nw(m)\mathbf{1}_{\{i=q\}},$$

which is negative only at $i = q$, where $y_q < y_{q+1}$. Hence, $t^{(m)} \in W \cap P$ for all $m \neq q$.

We next show that the set $\{t^{(m)}\}_{m \neq q}$ consists of $n-2$ linearly independent vectors. Define a linear map $\Delta : \mathbb{R}^n \to \mathbb{R}^{n-1}$ as $\Delta(t) = \big(t_2 - t_1,\ t_3 - t_2,\ \ldots,\ t_n - t_{n-1}\big)$. Now, $\Delta\big(\tau^{(k)}\big)_i = n\mathbf{1}_{\{i=k\}}$. Hence,

$$\Delta\big(t^{(m)}\big)_i = nw(q)\, \mathbf{1}_{\{i=m\}} - nw(m)\, \mathbf{1}_{\{i=q\}}, \qquad i = 1, \ldots, n-1.$$

Let $\lambda_m \in \mathbb{R}$, $m \neq q$, satisfy $\sum_{m \neq q} \lambda_m t^{(m)} = 0$, and fix one index $i \neq q$. By applying Δ

to both sides and reading off position $i \neq q$, we get

$$0 = \Delta\Big(\sum_{m\neq q} \lambda_m t^{(m)}\Big)_i = \sum_{m\neq q} \lambda_m \, \Delta\big(t^{(m)}\big)_i$$
$$= \sum_{m\neq q} \lambda_m \Big(nw(q)\mathbf{1}_{\{i=m\}} - nw(m)\mathbf{1}_{\{i=q\}}\Big) = nw(q)\,\lambda_i.$$

Thus, $\lambda_i = 0$ for all $i \neq q$. Therefore, $\{t^{(m)}\}_{m\neq q}$ is linearly independent and forms a basis of W.

By equation (6), $\sum_{i=1}^n t_i a_i(Y) = 0$ for every element t in the basis of W, hence for every $t \in W$, that is, $a(Y) \in W^\perp = \operatorname{span}(\mathbf{1}, \nabla I)$ and $a_i(Y) = \beta + \gamma I_i$. Equations (3) and (4) then force $\beta = -\gamma \bar{I}$ and $\gamma \sum_j (I_j - \bar{I}) y_j = I(Y)$, which is (5). Conversely, (5) satisfies Axioms 1–5 exactly as in the proof of Theorem 1, and it satisfies Axiom 6 because any rank-preserving inequality-neutral scheme has $\sum_{i=1}^n a_i(Y) t_i = \gamma \sum_{i=1}^n (I_i - \bar{I}) t_i = 0$. □

In the case of the Gini coefficient, the decomposition of Theorem 2 has the contribution

$$S(Y^k, Y) = \frac{\mu_k}{\mu} \bar{G}(Y^k) = \frac{\mu_k}{\mu} \frac{2}{n^2 \mu_k} \sum_{i=1}^n \left(r_i - \frac{n+1}{2} \right) y_i^k, \tag{7}$$

where r_i is the rank of individual i in the distribution of total income and $\bar{G}(Y^k)$ is the pseudo-Gini of source k.[14] In particular, the pseudo-Gini for source k is computed using the formula of the Gini coefficient on source-k incomes but using the ranks of total income instead of those of source k. Alternatively, the relative contribution of source k is

$$\frac{\operatorname{Cov}(Y^k, F)}{\operatorname{Cov}(Y, F)},$$

where F is the cumulative distribution function of total income or, alternatively, the

[14] For the Gini coefficient and other rank-dependent inequality indices, the ranks of individuals with identical total incomes are indeterminate. This does not affect the value of the index, which is invariant to permutations of individuals with equal incomes, but it may affect source contributions when tied individuals have different source incomes. This ambiguity is inherent in source decompositions of rank-dependent inequality indices, including the traditional natural decomposition of the Gini coefficient. For the Gini coefficient, it is commonly resolved by assigning tied individuals their average rank. This convention is not suitable for a general rank-dependent index because, when the index is nonlinear in ranks, replacing the ranks of tied individuals by their average can change the value of the index. A natural extension of the average-rank convention is therefore to compute the source contributions under every possible assignment of ranks within each set of tied individuals and take their average. This is the unique extension that is linear in source incomes, treats tied individuals symmetrically, and agrees with the ordering-specific contributions whenever all admissible orderings yield the same contribution. For the Gini coefficient, this procedure is equivalent to assigning tied individuals their average rank.

vector of income ranks of individuals in terms of total income. The contribution can also be written as

$$S(Y^k, Y) = \frac{\mu_k}{\mu} R_k G(Y^k), \tag{8}$$

where R_k is the Gini correlation that is bounded between -1 and 1 (Lerman and Yitzhaki, 1985). That is, the contribution of source k can be written as the product of the income share of source k, a measure of correlation between source-k income and total income, and the Gini coefficient of source k.

IV.2 Natural decomposition of top income shares

The result of Theorem 2 can also be used to derive a natural decomposition for top income shares. Note, however, that we cannot directly apply Theorem 2, since it requires that the inequality index has a value of zero when incomes are equally distributed. Top income shares do not satisfy this property. For example, the top 10% income share equals 10% when there is no inequality. Instead, we can apply Theorem 2 to derive a natural decomposition for *excess top income shares*, defined below, and use the resulting source contributions to derive a natural decomposition for top income shares by adding the minimum value of the index times the source's income share to each contribution.

The top $\alpha \in (0, 1)$ income share can be written in the form of $I(Y) = \sum_{i=1}^{n} a_i(Y) y_i$ with

$$a_i(Y) = \begin{cases} 0 & \text{for } r_i \leq (1-\alpha)n \\ \frac{1}{n\mu} & \text{for } r_i > (1-\alpha)n, \end{cases} \tag{9}$$

where r_i is the rank of individual i in the distribution of total income, n is the population size, and μ is the mean income.[15] If we want top income shares to satisfy the standard property of an equal distribution having an index of zero, then we need to adjust the weights as follows:

$$a_i(Y) = \begin{cases} -\frac{\alpha}{n\mu} & \text{for } r_i \leq (1-\alpha)n \\ \frac{1}{n\mu} - \frac{\alpha}{n\mu} & \text{for } r_i > (1-\alpha)n. \end{cases} \tag{10}$$

We call the inequality index with such $a_i(Y)$ the *excess top α income share*. If the top 10% income share is equal to 25%, then the excess top 10% income share is equal to 15%, because the top 10% income share exceeds its minimum value by 15 percentage

[15] We require $\alpha n \in \mathbb{N}$ so that the top α share is well-defined.

points.

Applying Theorem 2 to excess top income shares yields the contribution $S(Y^k, Y) = \sum_{i=1}^n a_i(Y) y_i^k$, where $a_i(Y)$ is given by equation (10) as the unique decomposition for excess top shares. If we then add $\alpha\mu_k/\mu$, i.e., the minimum value of the corresponding top income share multiplied by the income share of the source, to each contribution, we get the natural decomposition of the top income shares. In this decomposition, the contribution of source k equals $S_k = \sum_{i=1}^n a_i(Y) y_i^k$, where $a_i(Y)$ is given by (9).[16]

The natural decomposition of top income shares is simple and intuitive. For example, in the case of the top 10% income share, if the sum of labor incomes of the richest 10% (by total income) is equal to 5% of total income in the population, then the contribution of labor income to the top 10% income share is 5 percentage points. The relative contributions are even more intuitive. If labor income accounts for 25% of total income among the richest 10%, then the relative contribution of labor income is 25%. This coincides with the way many studies, such as Piketty et al. (2018) and Alstadsæter et al. (2025), have decomposed top income shares. Our result can be seen as a theoretical justification for this intuitive decomposition.

As shown in equation (8), the contribution of source k in the natural decomposition of the Gini coefficient can be written as the product of source k's income share, the Gini correlation of source k and total incomes, and the Gini coefficient of source k. Similarly, in the natural decomposition of the top α income share, the contribution of source k can be written as

$$S(Y^k, Y) = \frac{\mu_k}{\mu} R_k T_k,$$

where μ_k/μ is the income share of source k, R_k is the ratio of the average source-k income among the top α share by total income to the average source-k income among the top α share by source-k income, and T_k is the top α income share of source k. The middle term R_k can be thought of as a measure of correlation between source-k incomes and total incomes, with the exception that its range is $[0, 1]$ when all incomes are non-negative. The measure equals zero when the richest 100α percent in total income have no source-k income, and the maximum value of 1 is reached when the richest 100α percent in source-k income coincide with the richest α percent in total income.

[16] This decomposition satisfies Axioms 1–4 and 6 but not Axiom 5, which is natural since the top income shares do not satisfy a corresponding normalization axiom either.

IV.3 A counterfactual interpretation

An important question in source decomposition is how to interpret the contributions assigned to individual income sources. Although the contributions in a static decomposition add up to observed inequality for a given income distribution, they do not reveal how these contributions relate to changes in inequality resulting from changes in the underlying income distribution. This distinction has motivated criticism of the interpretation of natural source decompositions, particularly for the Gini coefficient (Podder, 1993; Cancian and Reed, 1998; Urban, 2024). An alternative approach is to define source contributions in terms of their marginal effects on inequality under a specified change in income. In this subsection, we show that the decomposition derived in Theorem 1 admits such a counterfactual interpretation.

A prominent example of this alternative approach is the marginal decomposition proposed by Lerman and Yitzhaki (1985) for the Gini coefficient and extended to other inequality indices by Paul (2004). It considers a proportional increase in income from a particular source, holding all other sources fixed. For a nondegenerate income distribution Y at which I is differentiable, the marginal effect of source k is

$$m_k := \frac{d}{d\varepsilon} I(Y + \varepsilon Y^k)\Big|_{\varepsilon=0} = \sum_{i=1}^{n} I_i(Y) y_i^k. \tag{11}$$

This experiment provides a direct interpretation of how a proportional expansion of source-k income affects inequality. However, the resulting marginal effects do not provide an accounting decomposition of observed inequality. Since increasing all sources by the same proportion only rescales total income, scale independence implies $\sum_{k=1}^{K} m_k = \sum_{i=1}^{n} I_i(Y) y_i = 0$. Thus, conventional marginal effects cannot be interpreted as relative contributions that add up to total inequality.

Our axioms motivate a different marginal experiment. Whereas the conventional approach considers a proportional increase in source income, our normalization axiom requires that an equally distributed source make no contribution to observed inequality, regardless of its income share. Together with linearity, this implies that the contribution of a source depends only on its deviations from the source mean, evaluated relative to the distribution of total income. This suggests examining how inequality responds to a mean-preserving equalization of source income.

Specifically, suppose that a fraction τ of source-k income is replaced by an equal

amount for every individual, holding all other sources fixed.[17] The resulting distribution of source-k income is

$$Y^k(\tau) = (1-\tau)Y^k + \tau\mu_k\mathbf{1}, \tag{12}$$

where μ_k denotes mean income from source k. At $\tau = 0$, the distribution is unchanged, while at $\tau = 1$, source-k income is equally distributed. Importantly, equalizing a source does not necessarily reduce inequality in total income. For example, if source-k income is concentrated among poorer individuals, equalizing it may increase total inequality.

Define the marginal reduction in inequality resulting from this experiment as[18]

$$d_k := -\frac{d}{d\tau} I\big(Y + \tau(\mu_k\mathbf{1} - Y^k)\big)\Big|_{\tau=0} = \sum_{i=1}^{n} I_i(Y)(y_i^k - \mu_k). \tag{13}$$

A positive value of d_k means that equalizing source k reduces inequality to first order, while a negative value means that it increases inequality. A zero value means that the distribution of source income is neutral with respect to the inequality index to first order, even if source income is not equally distributed.

The marginal reduction in inequality from simultaneously equalizing all sources is

$$D(Y) := \sum_{k=1}^{K} d_k = \sum_{i=1}^{n} I_i(Y)(y_i - \mu) > 0, \tag{14}$$

where positivity follows from the proof of Theorem 1. Theorem 1 implies that

$$S(Y^k, Y) = \frac{I(Y)}{D(Y)} d_k, \qquad \frac{S(Y^k, Y)}{I(Y)} = \frac{d_k}{\sum_{\ell=1}^{K} d_\ell}. \tag{15}$$

Equation (15) provides a counterfactual interpretation of our decomposition: the relative contribution of a source to observed inequality equals its share of the marginal inequality reduction from simultaneously equalizing all sources. In particular, sources whose equalization would increase inequality receive negative contributions. The counterfactual interpretation of the natural decomposition of the Gini coefficient was established by Urban (2024) and extended to the half squared coefficient of variation by Urban (2026). Our result establishes this interpretation for the decomposition charac-

[17] This can be achieved by imposing a linear tax on source-k income and distributing the revenue as a lump sum transfer.

[18] For rank-dependent indices that are not differentiable at income ties, the marginal effect of source equalization may depend on how the redistribution changes income ranks. We therefore restrict this interpretation to distributions at which the inequality index is differentiable.

terized by our axioms for any admissible inequality index.

The two marginal experiments are also directly related. Let $\theta_k = \mu_k/\mu$ denote the income share of source k. Scale independence implies $\sum_i I_i(Y)y_i = 0$, and hence $D(Y) = -\mu \sum_i I_i(Y)$. Combining this with equations (11), (13), and (15), we obtain

$$\frac{S(Y^k, Y)}{I(Y)} = \theta_k + \frac{m_k}{D(Y)}. \tag{16}$$

Thus, a proportional increase in source-k income reduces inequality if and only if its relative contribution to observed inequality is smaller than its income share. The two approaches therefore answer different questions: conventional marginal effects measure how inequality responds to proportional changes in source income, whereas our decomposition allocates observed inequality across sources in proportion to their marginal effects under mean-preserving equalization.

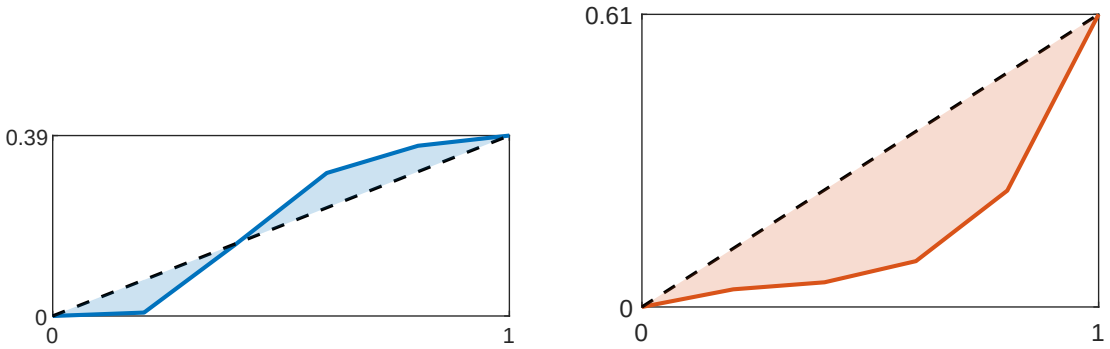
V Geometric interpretation of source decomposition of the Gini coefficient

In this section, we discuss how the unique source decomposition for the Gini coefficient can be interpreted geometrically using the Lorenz and concentration curves.

The Gini coefficient can be interpreted geometrically as twice the area between the Lorenz curve and the 45-degree line. In the natural decomposition of the Gini coefficient, the contribution of a specific source can be interpreted geometrically as twice the area between the concentration curve and the 45-degree line, scaled by the income share of the source.[19] Normalization by income share corresponds to scaling the vertical axis of the Lorenz diagram by the income share.

The concentration curve of source k consists of points (x, y) where y is the cumulative share of source-k income of the poorest fraction x of the population in terms of total income. Thus, the concentration curve of a source corresponds to the Lorenz curve of the source with the exception that the population is ordered by total income instead of the source income. In order to relate the concentration curve to the contribution of the source, we need to scale the y-axis by the income share of the source.

[19] Pyatt et al. (1980) show that the concentration ratio, or the pseudo-Gini coefficient in the terminology of Fei et al. (1978), is equal to one minus twice the area under the concentration curve, which is equal to twice the area between the concentration curve and the 45-degree line.



(a) Source 1 **(b)** Source 2

Figure 4: Normalized concentration curves of sources 1 and 2 in Example 3.

Example 3. Suppose that total income consists of two sources, 1 and 2. The distribution of source-1 incomes is given by $(1, 20, 21, 8, 3)$ and the distribution of source 2 is $(5, 2, 6, 20, 50)$. The distribution of total income is thus $(6, 22, 27, 28, 53)$. Figure 4 depicts the normalized concentration curves in this example. For source 1, the concentration curve crosses the 45-degree line. The area above the 45-degree line counts negatively. Since there is more area above the 45-degree line than below, the contribution of source 1 to total inequality is negative.

Since the normalized concentration curves add up to the Lorenz curve and the normalized 45-degree lines add up to the 45-degree line, we can decompose the area between the Lorenz curve of total income and the 45-degree line into source contributions by adding up the vertical differences between the normalized 45-degree lines and concentration curves for each source. This is illustrated for Example 3 in Figure 5. Note that while the 45-degree line is the reference for source 1 and hence the blue area above it counts negatively for source 1, the reference for source 2 is the concentration curve of source 1. The red area between the blue line and the 45-degree line thus counts positively for source 2, as well as all the red area between the 45-degree line and the Lorenz curve (solid black line).[20]

VI Conclusion

This paper revisits the problem of decomposing inequality indices into contributions of income sources. Starting from the axiomatic framework of Shorrocks (1982), we argue that a key axiom in his framework, *two factor symmetry*, imposes an undesirable restriction because it rules out differences in how inequality indices evaluate the alignment of source income distributions with total income. We propose an alternative

[20]While the order in which the sources are added to the figure matters for how the figure looks, it does not matter for the sizes of the areas, i.e., the contributions of the sources.

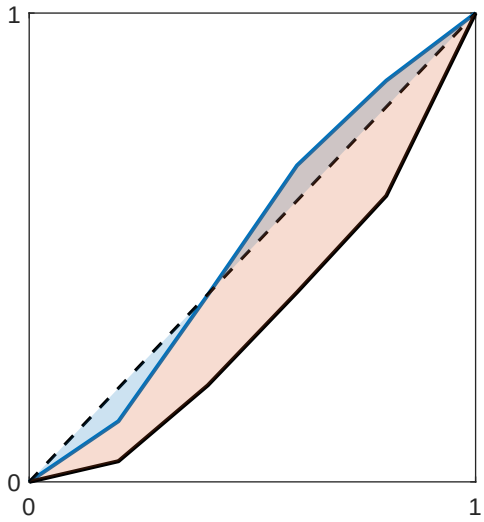


Figure 5: Geometric interpretation of the natural source decomposition of the Gini coefficient when total income consists of two sources. The solid black line depicts the Lorenz curve of total income. The shaded blue area represents the contribution to inequality due to source 1 and the shaded red area depicts the contribution of source 2. The contributions equal -0.024 and 0.318 for source 1 and 2, respectively.

axiom, *conditional transfer neutrality*, which links the decomposition rule to the way the chosen inequality index evaluates income transfers.

The resulting decomposition is unique for any differentiable inequality index. For familiar indices such as the Gini coefficient, the coefficient of variation, and top income shares, it coincides with the standard natural decompositions. For other indices, including the Theil index and the mean log deviation, it yields new source decompositions that reflect the index-specific sensitivity to different parts of the income distribution. The framework therefore provides a general axiomatic foundation for source decomposition that respects how different inequality indices evaluate inequality.

References

Alstadsæter, Annette, Martin Jacob, Wojciech Kopczuk, and Kjetil Telle, "Accounting for business income in measuring top income shares: Integrated accrual approach using individual and firm data from Norway," *Journal of the European Economic Association*, 2025, p. jvaf058.

Auten, Gerald and David Splinter, "Income inequality in the United States: Using

tax data to measure long-term trends," *Journal of Political Economy*, 2024, *132* (7), 2179–2227.

Cancian, Maria and Deborah Reed, "Assessing the Effects of Wives' Earnings on Family Income Inequality," *The Review of Economics and Statistics*, 1998, *80* (1), 73–79.

Chantreuil, Frédéric and Alain Trannoy, "Inequality decomposition values: the trade-off between marginality and efficiency," *The Journal of Economic Inequality*, 2013, *11* (1), 83–98.

Cowell, Frank A and Carlo V Fiorio, "Inequality decompositions—a reconciliation," *The Journal of Economic Inequality*, 2011, *9* (4), 509–528.

Fei, John CH, Gustav Ranis, and Shirley WY Kuo, "Growth and the family distribution of income by factor components," *The Quarterly Journal of Economics*, 1978, *92* (1), 17–53.

Fournier, Martin, "Inequality decomposition by factor component: a "rank-correlation" approach illustrated on the Taiwanese case," *Recherches économiques de Louvain*, 2001, *67* (4), 381–403.

Lerman, Robert I and Shlomo Yitzhaki, "Income Inequality Effects by Income Source: A New Approach and Applications to the United States," *The Review of Economics and Statistics*, 1985, pp. 151–156.

Paul, Satya, "Income sources effects on inequality," *Journal of Development Economics*, 2004, *73* (1), 435–451.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, "Distributional national accounts: methods and estimates for the United States," *The Quarterly Journal of Economics*, 2018, *133* (2), 553–609.

Podder, Nripesh, "The disaggregation of the Gini coefficient by factor components and its applications to Australia," *Review of Income and Wealth*, 1993, *39* (1), 51–61.

Pyatt, Graham, Chau-Nan Chen, and John Fei, "The distribution of income by factor components," *The Quarterly Journal of Economics*, 1980, *95* (3), 451–473.

Rao, Vasala M, "Two decompositions of concentration ratio," *Journal of the Royal Statistical Society. Series A (General)*, 1969, *132* (3), 418–425.

Sastre, Mercedes and Alain Trannoy, "Shapley inequality decomposition by factor components: Some methodological issues," *Journal of Economics*, 2002, *77* (Suppl 1), 51–89.

Shorrocks, Anthony F, "Inequality Decomposition by Factor Components," *Econometrica*, 1982, pp. 193–211.

_ , "The impact of income components on the distribution of family incomes," *The Quarterly Journal of Economics*, 1983, *98* (2), 311–326.

_ , "Decomposition procedures for distributional analysis: a unified framework based on the Shapley value," *Journal of Economic Inequality*, 2013, *11* (1), 99–126.

Smith, Matthew, Owen Zidar, and Eric Zwick, "Top wealth in America: New estimates under heterogeneous returns," *The Quarterly Journal of Economics*, 2023, *138* (1), 515–573.

Urban, Ivica, "Two Classical Decompositions of the Gini Index by Income Source: Interpretation of the Contribution Terms," *Journal of Income Distribution*, Jan. 2024, *32* (3-4).

_ , "Decomposing income inequality by source: a dual-method and dual-index analysis," *The Journal of Economic Inequality*, 2026.