Subgroup Decomposition of the Gini Coefficient: A New Solution to an Old Problem[*]

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

[1]Department of Economics, Brown University
[2]Department of Economics, Brown University

[Link to most recent version]

Abstract

We study inequality decomposition by population subgroups. We define properties of a satisfactory decomposition and derive the implied decomposition formulas for well-known inequality indices. We find that the Gini coefficient, the generalized entropy indices, and the Foster-Shneyerov indices all admit satisfactory decomposition formulas derived from a common set of axioms. While our axiomatic approach recovers the established decomposition formulas for the generalized entropy and the Foster-Shneyerov indices, it leads us to a novel decomposition formula for the Gini coefficient. The decomposition of the Gini coefficient is unique given our axioms, easy to compute, and has both a geometric and an arithmetic interpretation.

I Introduction

Empirical analyses of inequality often aim to quantify the contribution of within-group inequality and between-group differences to aggregate inequality. This classic objective is known as inequality decomposition by population subgroups. Examples include analyses of inequality within and between countries (Sala-i Martin, 2006), demographic subgroups (Cowell and Jenkins, 1995), and firms (Song et al., 2019).

The Gini coefficient is the most popular measure of inequality, yet it lacks a universally accepted decomposition formula.[1] Existing approaches to decompose the Gini coefficient are criticised for producing within-group or between-group inequality terms that behave counter-intuitively. Further, these approaches often rely on introducing a third term that describes neither within-group nor between-group inequality. This lack of a satisfactory subgroup decomposition formula is arguably the most significant drawback of the Gini coefficient, which is otherwise valued for its intuitive arithmetic definition and geometric relation to the Lorenz curve.

Whether an inequality measure admits a satisfactory decomposition depends on the definition of decomposability. In a series of seminal papers, Bourguignon (1979), Shorrocks (1980, 1984), Cowell (1980), and Cowell and Kuga (1981) define decomposability in terms of a strict *aggregativity* requirement. An inequality index is decomposable if aggregate inequality is a function of subgroup inequality indices, average incomes, and population shares. The main result from this literature is that any inequality measure which satisfies this aggregativity requirement is ordinally equivalent to a generalized entropy index. These indices include the Theil index, the mean log deviation, and half the squared coefficient of variation, but not the Gini coefficient.

Foster and Shneyerov (2000) explore a different definition of decomposability which they call *path independence*. They define within-group inequality and between-group inequality in terms of representative incomes. Within-group inequality is the level of inequality when all subgroup distributions are re-scaled to have a common representative income level. Between-group inequality is the level of inequality when each individual's income is replaced with their subgroup's representative income. An inequality index is decomposable if it can be written as the sum of such a within and between-group inequality term. Imposing fairly weak constraints on what constitutes a representative income function, Foster and Shneyerov (2000) derive a family of decomposable inequality indices that is uniquely determined (up to a scalar multiple) by their path

---

[1]The Gini coefficient is also one of the earliest measures of inequality. It was introduced by Corrado Gini in 1914 and constitutes an algebraic equivalent of the geometric measure of inequality devised by Max Otto Lorenz in 1905. Suggested decompositions for the Gini coefficient can be found in Soltow (1960), Bhattacharya and Mahalanobis (1967), Rao (1969), Mehran (1975), Piesch (1975), Mangahas (1975), Paglin (1975), Pyatt (1976), Murray (1979), Dagum (1980), Nygård and Sandström (1981), Kakwani (1984), Silber (1989), Yitzhaki and Lerman (1991), Lambert and Aronson (1993), Yitzhaki (1994), Dagum (1997), Aaberge (2000), Frosini (1989, 1990, 2003), Aaberge et al. (2005), Lambert and Decoster (2005), Deutsch and Silber (2007), Monti (2007), Costa (2008), Okamoto (2009), Shorrocks (2013), Mussard and Mornet (2019), and Attili (2021). See Giorgi (2011) for a history of different suggestions to decompose the Gini coefficient. The most influential decomposition of the Gini coefficient is due to Bhattacharya and Mahalanobis (1967).

independence requirement. This family includes the mean log deviation and half the variance of logs, but not the Gini coefficient.[2]

In this paper, we propose a definition of decomposability that requires neither aggregativity (violated by the Foster-Shneyerov indices) nor path independence (violated by the generalized entropy indices). Instead, our approach is to specify desirable behavior for the within-group and the between-group inequality terms. We call an inequality index decomposable if we can write it as the sum of two terms where one term satisfies the axioms for within-group inequality, and the other term satisfies the axioms for between-group inequality. We apply our axiomatic framework to study the subgroup decomposition of well-known inequality indices.

We show that our axioms uniquely determine the decomposition formulas for the generalized entropy and the Foster-Shneyerov families. Moreover, we show that the Gini coefficient is also decomposable under our framework and derive its unique decomposition formula. While our axiomatic approach recovers the known decomposition formulas for the generalized entropy and the Foster-Shneyerov indices, it leads us to a novel decomposition formula for the Gini coefficient. The decomposition of the Gini coefficient is easy to compute, and it has both a geometric and an arithmetic intuition. The decomposition formula for the Gini coefficient is given in Theorem 2.

Within-group inequality summarizes how inequality within subgroups contributes to aggregate inequality. An important result in this paper concerns the possible functional form that within-group inequality can take. We show that imposing intuitive restrictions implies that the within-group inequality term must take the form of a weighted quasi-arithmetic mean of subgroup inequalities with weights that depend on the subgroups' population and income shares. This result follows from axioms similar to those used in the characterization of means and aggregation functions.[3] Our result can be viewed as a characterization of quasi-arithmetic aggregation functions that allow for two-dimensional weights and satisfy weak reflexivity.

Between-group inequality summarizes how differences between subgroup income distributions contribute to aggregate inequality. In our framework, between-group inequality inherits the aggregativity properties of the inequality measure. For given subgroup inequality levels and population shares, generalized entropy indices depend on differences in average incomes between subgroups. Therefore, the between-group inequality

---

[2]Note that the variance of logs is not, strictly speaking, an inequality measure as it violates the Pigou-Dalton principle, see Foster and Ok (1999).

[3]For summaries of these literatures, see the introductions in Matkowski and Páles (2015) and Burai et al. (2021), and Grabisch et al. (2011).

term for the generalized entropy measures depends only on differences in average incomes. Foster-Shneyerov indices depend on differences in power means, and between-group inequality is likewise a function of differences in power means. We show that the Gini coefficient generally depends on each subgroup's entire income distribution. Consequently, between-group inequality in the Gini decomposition depends on differences between subgroups in all moments of the income distribution.

We demonstrate our decomposition for the Gini coefficient by analyzing the evolution of within-group and between-group income inequality in the United States over the past fifty years. We find that observed characteristics like the age, gender, education, and race of household heads explain a lower share of household income inequality today than in the past. We also find that the convergence of male and female earnings distributions has significantly slowed the increase in overall earnings inequality. The documented decrease in between-group inequality is mostly driven by the convergence in average earnings between men and women while differences in the shape of the earnings distributions continue to contribute to overall inequality.

Throughout this paper, we make use of the fact that the Lorenz curve of an aggregate population can be related to the Lorenz curves of the subgroups via the geometric operation of Minkowski addition. This geometric representation of the aggregation problem unlocks the Brunn-Minkowski theorem, a powerful theorem from convex geometry. We use these insights to discuss the extent of the subgroup inconsistency problem of the Gini coefficient.[4] The Brunn-Minkowski theorem is also at the core of the arithmetic and geometric interpretation of the decomposition of the Gini coefficient.

The rest of the paper is structured as follows. Section II introduces the notation and defines the inequality indices discussed in this paper. In section III, we introduce the axiomatic framework for subgroup decomposition of inequality indices. In section IV, we derive the decompositions of the generalized entropy and the Foster-Shneyerov indices from our axioms and show that a decomposition for Atkinson indices that satisfies the same list of axioms does not exist. In section V, we discuss the aggregativity properties of the Gini coefficient and derive the unique decomposition of the Gini coefficient. Section VI offers geometric and arithmetic interpretations of our Gini decomposition. We apply this decomposition to the US income distribution in section VII. Section VIII concludes. All proofs are presented in appendix A.

---

[4]The subgroup inconsistency problem has been discussed by Mookherjee and Shorrocks (1982), Cowell (1988), and Sen and Foster (1997).

II Notation and Definitions

We use Y to denote the income distribution of the population and Y_k to denote the distribution of incomes of subgroup k. Subgroup k's population size and aggregate income are denoted by n_k and $\bar{y}_k$, respectively, and population and income shares are denoted by π_k and θ_k, respectively.[5] We use $\mathbb{R}_+$ to denote the interval $[0, \infty)$ and $\mathbb{R}_{++}$ to denote the interval $(0, \infty)$.

An inequality index $I \colon \mathcal{D} \to \mathbb{R}_+$ is a function that maps a space of distributions $\mathcal{D}$ to the non-negative real numbers and satisfies the five standard axioms of anonymity, scale independence, population independence, normalization, and the Pigou-Dalton principle of transfers. For different inequality indices, we will impose different restrictions on $\mathcal{D}$. In this paper, we discuss four different classes of inequality indices: The Gini coefficient, the generalized entropy indices, the Foster-Shneyerov indices, and the Atkinson indices.

The Gini coefficient is defined for any distribution given by a generalized Lorenz curve, that is, a convex function $L \colon [0, 1] \to \mathbb{R}$, such that $L(0) = 0$.[6] Specifically, the Gini coefficient is given by

$$G(Y) = 2 \int_0^1 p - \frac{L(p)}{\mu} dp,$$

where μ is the mean income.

The other inequality indices require a more restricted domain as they are only defined for positive incomes and distributions for which a density function or a probability mass function exists. Thus, we take the union of the space of absolutely continuous probability distributions supported on the positive real line and the space of discrete distributions supported on the positive real line as the domain for these indices. Gen-

---

[5]Throughout this paper, we frame our discussion in terms of income distributions, but income can be substituted with any non-negative real valued attribute.

[6]A generalized Lorenz curve is equal to a standard Lorenz curve scaled by the mean income. Unlike standard Lorenz curves, generalized Lorenz curves preserve information on the mean of the distribution. Moreover, generalized Lorenz curves are a more general way to represent distributions than cumulative distribution functions as they allow for distributions with perfect inequality, that is, distributions where a zero measure set of people earn all income. Note also that we allow for negative incomes by not requiring the generalized Lorenz curve to be non-decreasing.

eralized entropy indices are a family of inequality indices defined as

$$GE_\alpha(Y) = \begin{cases} \frac{1}{\alpha(\alpha-1)} \int f(y)\left(\left(\frac{y}{\mu}\right)^\alpha - 1\right)dy & \text{for } \alpha \neq 0, 1 \\ \int f(y) \ln \frac{\mu}{y} dy & \text{for } \alpha = 0 \\ \int f(y)\frac{y}{\mu} \ln \frac{y}{\mu} dy & \text{for } \alpha = 1, \end{cases}$$

where f is the probability density function of Y, μ is the mean income, and $\alpha \in \mathbb{R}$ is a parameter. The Foster-Shneyerov family, introduced in Foster and Shneyerov (2000), is defined as

$$FS_q(Y) = \begin{cases} \frac{1}{q} \ln \frac{\mu_q(Y)}{g(Y)} & \text{for } q \neq 0 \\ \frac{1}{2}\text{Var}(\ln Y) & \text{for } q = 0, \end{cases}$$

where $\mu_q(Y) = \left(\int y^q f(y)dy\right)^{\frac{1}{q}}$ is the power mean of order q of Y and $g(Y) = e^{\int \ln y f(y)dy}$ is the geometric mean of Y.[7] Finally, the family of Atkinson indices, introduced in Atkinson (1970), is defined as

$$A_\varepsilon(Y) = 1 - \frac{\mu_{1-\varepsilon}(Y)}{\mu(Y)} \qquad \text{for } \varepsilon < 1,$$

where $\mu_{1-\varepsilon}(Y)$ is the power mean of order $1-\varepsilon$ of Y and $\mu_0(Y) := g(Y)$ is the geometric mean of Y.

III Axiomatic Framework

We define a subgroup decomposition of an inequality index I as the sum of a within-group inequality term W and a between-group inequality term B that add up to aggregate inequality. We propose a list of axioms that must be satisfied by any within-group or between-group inequality term. We call an inequality measure decomposable if it admits a decomposition that satisfies these axioms.

III.1 Within-group Inequality

Let I denote an inequality index. We require that within-group inequality depends only on subgroup inequality levels and aggregate characteristics, that is, total population

---

[7]Note that, strictly speaking, only Foster-Shneyerov indices with parameter values $q \geq 1$ are inequality indices. For other parameter values, the Pigou-Dalton transfer principle is violated.

and total income. Let $W_K \colon \mathbb{R}^{3K}_+ \to \mathbb{R}_+$ denote within-group inequality for a population consisting of K subgroups. Furthermore, let $I_k, n_k, \bar{y}_k$ denote subgroup k's level of inequality, total population, and total income, respectively. We posit the following axioms:

1. **Regularity.** W_K is continuous, and strictly increasing in I_k if $n_k, \bar{y}_k > 0$.

2. **Symmetry.** $W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = W_K\Big(P\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big)\Big)$
   for any permutation P and for all $(I_k, n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}^{3K}_+$.

3. **Scale and population independence.**

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = W_K\big((I_1, an_1, b\bar{y}_1), \ldots, (I_K, an_K, b\bar{y}_K)\big)$$

   for all $a, b > 0$ and for all $(I_k, n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}^{3K}_+$.

4. **Normalization.** $W_K\big((0, n_1, \bar{y}_1), \ldots, (0, n_K, \bar{y}_K)\big) = 0$ for all $(n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}^{2K}_+$.

5. **Weak reflexivity.** $W_K\big((I, a_1 n, a_1 \bar{y}), \ldots, (I, a_K n, a_K \bar{y})\big) = I$ for all $(I, n, \bar{y})_{k=1}^K \in \mathbb{R}^{3K}_+$ and for all $(a_k)_{k=1}^K \in \mathbb{R}^{K}_{++}$.

6. **Replacement.**

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big)$$
$$= W_{K-m+1}\Big(\big(\tilde{I}, \textstyle\sum_{k=1}^m n_k, \sum_{k=1}^m \bar{y}_k\big), (I_{m+1}, n_{m+1}, \bar{y}_{m+1}), \ldots, (I_K, n_K, \bar{y}_K)\Big)$$

   for all $(I_k, n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}^{3K}_+$ and $m \leq K$, where $\tilde{I} = W_m\big((I_1, n_1, \bar{y}_1), \ldots, (I_m, n_m, \bar{y}_m)\big)$.

Symmetry ensures that within-group inequality is independent of the labels given to each subgroup and is sometimes also called anonymity. Scale and population independence ensures that the decomposition is independent of the size of the population and the unit of measurement for income. Normalization ensures that within-group inequality is zero when there is no inequality within any subgroup. Weak reflexivity ensures that within-group inequality equals aggregate inequality when the population consists of only one group and that within-group inequality does not change if that group is subdivided into subgroups with identical income distributions. Finally, replacement ensures that the level of within-group inequality does not change if we merge any number of subgroups into one composite subgroup and redistribute incomes in that group so that the level of inequality equals the level of within-group inequality among its constituent subgroups.

The first important result is that the six axioms stated above imply strong restrictions on the functional form that within-group inequality can have. This result is stated in theorem 1.

Theorem 1. *Let* $(W_K)_{K=1}^{\infty}$ *be a sequence of functions* $W_K \colon \mathbb{R}_+^{3K} \to \mathbb{R}_+$. *Then,* (W_K) *satisfies axioms 1-6 if and only if*

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = f^{-1}\Big(\sum_{k=1}^{K} \pi_k^{1-\alpha} \theta_k^{\alpha} f(I_k) \Big), \tag{1}$$

for all $K \geq 1$ *and for all* $(I_k, n_k, \bar{y}_k)_{k=1}^{K} \in \mathbb{R}_+^{3K}$, *for some continuous strictly increasing function* f *such that* $f(0) = 0$ *and for some* $\alpha \in \mathbb{R}$, *where* $\pi_k = \frac{n_k}{\sum_{k=1}^{K} n_k}$ *and* $\theta_k = \frac{\bar{y}_k}{\sum_{k=1}^{K} \bar{y}_k}$.

Theorem 1 holds that within-group inequality must take the form of a quasi-arithmetic mean of subgroup inequalities with weights that do not necessarily sum up to one. Moreover, each subgroup's weight must be a weighted geometric average of its population and income share. The proof proceeds by showing that the axioms imply that for given aggregate characteristics W_K satisfies bisymmetry. Bisymmetry has been used by Aczél (1948) and Münnich et al. (2000) to characterize quasi-arithmetic means. The rest of the proof shows that the generating function f and the weight function do not depend on either the number of subgroups or their aggregate characteristics. The full proof can be found in appendix A.

The functional form in theorem 1 nests the expressions for within-group inequality in the standard decomposition formulas for the generalized entropy (Bourguignon, 1979; Shorrocks, 1980; Cowell, 1980) and the Foster-Shneyerov indices (Foster and Shneyerov, 2000), but is more general than the functional form assumed in these papers. Interestingly, however, the literature has also suggested expressions for within-group inequality that are ruled out by this theorem. For example, Shorrocks (2013) proposes an algorithm that results in within-group inequality terms that do not have the form of a quasi-arithmetic mean. Moreover, the most common decomposition formula for the Gini coefficient due to Bhattacharya and Mahalanobis (1967) is also ruled out by theorem 1. In particular, it violates weak reflexivity as within-group inequality does not equal aggregate inequality even if all subgroups are identical.

III.2   Between-group Inequality

Between-group inequality results from differences in income distributions across subgroups. What differences in subgroup income distributions give rise to between-group

inequality depends on the choice of inequality index and its aggregativity properties.[8] Axioms for between-group inequality of a given inequality index are therefore stated in terms of the aggregativity properties of the inequality index. For this purpose, we define the concept of Ω-aggregativity.

Definition 1. *An inequality measure $I: \mathcal{D} \to \mathbb{R}_+$ is Ω-aggregative if there exists a set of moments Ω and a function $F: \mathbb{R}_+^K \times \mathbb{R}_+^{MK} \times \mathbb{R}_+^K \to \mathbb{R}_+$ such that*

$$I(Y) = F\big(I(Y_1), \ldots, I(Y_K), \Omega_1, \ldots, \Omega_K; n_1, \ldots, n_K\big),$$

where Ω_k are subgroup k's values for the moments contained in Ω, Ω is the minimal set of moments such that the above formula holds, and M is the cardinality of Ω, If there does not exist an Ω such that F exists, then I is Ω-nonaggregative.

Generalized entropy indices and Atkinson indices are inequality measures that can be written as functions of subgroup inequality levels, average incomes, and population sizes. Hence, these measures are Ω-aggregative with $\Omega = \{\mathbb{E}[Y]\}$. The Foster-Shneyerov index of order $q \neq 0$ is aggregative with qth moments, that is $\Omega = \{\mathbb{E}[Y^q]\}$. In proposition 4 in section V, we show that the aggregate Gini coefficient can generally not be computed from subgroup Gini coefficients, population sizes, and any number of subgroup moments. Thus, the Gini coefficient is Ω-nonaggregative.

For a given inequality index, it is useful to introduce an equivalence relation between distributions, which we call Ω-equivalence.

Definition 2. *For a given inequality index, we call two distributions Ω-equivalent if they are identical in all moments contained in Ω. If Ω does not exist, then we call two distributions Ω-equivalent if they are the same distribution.*

We define between-group inequality to be a function of subgroup income distributions, population sizes, and total incomes, that is, $B_K: (\mathcal{D} \times \mathbb{R}_+ \times \mathbb{R}_+)^K \to \mathbb{R}_+$. We posit the following axioms:

7. **Normalization.** If all subgroup income distributions are Ω-equivalent, then B_K is equal to zero.

---

[8]To see this, suppose all subgroups have the same mean income and the same level of inequality. Weak reflexivity then implies that within-group inequality must equal the level of inequality within the subgroups. Depending on the inequality index, however, aggregate inequality can strictly exceed the level of within-group inequality. For example, Foster-Shneyerov indices for the aggregate population will generally be larger than the level of inequality in the subgroups. This excess inequality at the aggregate level is attributed to between-group inequality. On the other hand, generalized entropy indices do not produce between-group inequality in this case.

8. **Conditional distribution independence.** For given aggregate characteristics $n_1, \ldots, n_K, \bar{y}_1, \ldots, \bar{y}_K$, if there exists a function $F \colon \mathbb{R}_+^K \to \mathbb{R}_+$ such that

$$I(Y) = F\big(I(Y_1), \ldots, I(Y_K)\big),$$

then B_K does not depend on the distribution of income within subgroups.

Normalization ensures that between-group inequality is equal to zero when subgroup income distributions are identical in all moments that can affect aggregate inequality conditional on subgroup inequality levels. Conditional distribution independence ensures that whenever aggregate inequality can be computed from subgroup inequality levels and aggregate characteristics, between-group inequality only depends on aggregate characteristics of the subgroups. Equivalently, if redistributing incomes within subgroups cannot change aggregate inequality without changing subgroup inequality levels, then between-group inequality is independent of the distribution of income within subgroups. Since axioms 7 and 8 depend on the aggregativity properties of a given inequality index, they do not imply a general functional form for between-group inequality. However, as we show below, axioms 1 through 8 are restrictive enough to uniquely determine the decomposition formulas of the generalized entropy indices, the Foster-Shneyerov indices, and the Gini coefficient.

IV Decomposition of the Generalized Entropy, Foster-Shneyerov, and Atkinson Indices

The axiomatic framework introduced in the previous section leads to the standard decomposition formulas for the generalized entropy and the Foster-Shneyerov indices. Moreover, these are the only decomposition formulas that satisfy all our axioms for these indices. We also show that for Atkinson indices there does not exist a decomposition that satisfies all axioms.

Proposition 1. *For all $\alpha \in \mathbb{R}$, the decomposition of generalized entropy index*

$$GE_\alpha(Y) = \sum_{k=1}^K \pi_k^{1-\alpha} \theta_k^\alpha GE_\alpha(Y_k) + GE_\alpha(\bar{Y}), \tag{2}$$

where $\bar{Y}$ is derived from Y by replacing each individual's income by the relevant subgroup mean, is the unique decomposition that satisfies axioms 1-8.

Proposition 2. *For all $q \in \mathbb{R}$, the decomposition of Foster-Shneyerov index*

$$FS_q(Y) = \sum_{k=1}^{K} \pi_k FS_q(Y_k) + FS_q(\tilde{Y}),\tag{3}$$

where $\tilde{Y}$ is derived from Y by replacing each individual's income by the relevant subgroup power mean of order q, is the unique decomposition that satisfies axioms 1-8.

The proofs of propositions 1 and 2 are in appendix A.

We next turn to the Atkinson indices. These indices are aggregative with means and therefore ordinally equivalent to generalized entropy indices (Shorrocks, 1984). However, there is no universally accepted way to decompose Atkinson indices into within-group and between-group inequality terms that sum to aggregate inequality.[9] We show that there cannot exist an additive decomposition of Atkinson indices that satisfies all of our axioms.

Proposition 3. *For any $\varepsilon < 1$, the Atkinson index A_ε cannot be decomposed into a within-group inequality term and a between-group inequality term that satisfy axioms 1-8 and sum to the aggregate Atkinson index.*

The proof is in appendix A.

V   Decomposition of the Gini Coefficient

In this section, we derive a novel decomposition for the Gini coefficient from the axiomatic framework introduced in section III and show that it is unique. We also show that the decomposition for the Gini coefficient satisfies additional useful properties that are not imposed as axioms in our framework.

In order to derive a decomposition for the Gini coefficient from our axiomatic framework, we need to know the aggregativity properties of the Gini coefficient to understand the implication of axiom 7. Proposition 4 states that the Gini is Ω-nonaggregative. That is, the Gini coefficient generally depends on each subgroup's entire income distribution. Hence, axiom 7 requires that between-group inequality is zero if all subgroup distributions are identical.

---

[9]The most common decomposition of the Atkinson index was derived by Blackorby et al. (1981) and consists of a within-group inequality term W and a between-group inequality term B that combine into the aggregate Atkinson index as $A = W + B - WB$.

Proposition 4. *The Gini coefficient is Ω-nonaggregative. That is, there does not exist Ω, such that $G(Y) = F\big(G(Y_1), \ldots, G(Y_K), \Omega_1, \ldots, \Omega_K; n_1, \ldots, n_K\big)$.*

We show that there exists a unique decomposition of the Gini coefficient into a within-group and a between-group term that satisfies all the axioms introduced in section III.

Theorem 2. *A decomposition for the Gini coefficient satisfies axioms 1-8 if and only if the within-group inequality term is*

$$G^W(Y) = \bigg(\sum_{k=1}^{K} \sqrt{\pi_k \theta_k G(Y_k)} \bigg)^2, \tag{4}$$

where π_k and θ_k are the population and income share of subgroup k, respectively. Between-group inequality is the difference between the aggregate Gini coefficient and within-group inequality.

Within-group inequality in (4) is a weighted power-mean of subgroup inequalities where each subgroup is weighted by the geometric mean of their income and population share. This decomposition of the Gini coefficient is the only decomposition that satisfies all our axioms on the behavior of within-group and between-group inequality. In section VI, we offer geometric and arithmetic interpretations of this decomposition.

The decomposition of the Gini coefficient has a number of desirable properties that are not imposed as axioms. First, within-group inequality is homogeneous[10]. Any redistribution of incomes within subgroups that reduces subgroup Gini coefficients by some given factor will also reduce within-group inequality by the same factor. Note, however, that as a power mean the within-group inequality term is not linear in the subgroup Gini coefficients.

The between-group inequality term for the Gini coefficient should summarize the contribution to aggregate inequality stemming from all differences across subgroup income distributions. It is therefore desirable that for any given distribution of income, between-group inequality decreases if subgroup distributions become more similar. This is indeed the case. Randomly permuting the group affiliation of a sample of individuals results in subgroup distributions that are more similar to each other[11]. Proposition 5

---

[10]Homogeneity is in fact the only additional property that power means have over quasi-arithmetic means implied by theorem 1.

[11]Equivalently, you may consider replacing the incomes of a fixed fraction of randomly sampled individuals in each subgroup with a random draw from the aggregate income distribution. Equivalency between these two operations is a direct consequence of the anonymity property of inequality indices.

states that this operation reduces between-group inequality.

Proposition 5. *In an infinite population, randomly permuting the group affiliations for a random subset of individuals weakly reduces between-group inequality (while keeping the aggregate Gini coefficient constant). Between-group inequality remains constant if and only if all groups have the same distribution of income.*

We can also consider operations that should reduce both within-group inequality and between-group inequality. For example, redistributing incomes so that the difference between any individual's income and the average income is reduced by a fixed percentage clearly reduces aggregate inequality. Moreover, as such a redistribution at the same time compresses subgroup income distributions and brings them closer to each other, one may expect it to also decrease within-group and between-group inequality by similar proportions. Proposition 6 shows that this is indeed the case[12]. This property implies, for example, that a universal basic income financed by a flat tax that ignores group affiliations would not affect the share of aggregate inequality that is attributed to group differences.

Proposition 6. *Let $y(i)$ denote the income of quantile i, and let μ denote the average income in the population. For some $\alpha \in [0, 100]$, replacing every income $y(i)$ by $\tilde{y}(i) = y(i) - \frac{\alpha}{100}(y(i) - \mu)$ reduces within-group inequality and between-group inequality for the Gini coefficient by α percent.*

It is often of interest to know by how much aggregate inequality would be reduced if all subgroup means were equalized while preserving within-group inequality. For many inequality indices, such as the Theil index, this question is not easily answered. As was pointed out already by Shorrocks (1980), there is no obvious operation that would eliminate differences in average incomes between subgroups while keeping within-group inequality for any given generalized entropy index fixed[13]. Proposition 7 shows that the decomposition of the Gini coefficient admits an operation that eliminates differences between subgroups in average incomes while keeping within-group inequality fixed. This property can be used to further decompose between-group inequality into a first part that reflects differences in means and a second part that reflects differences in the shape of the distribution. We implement such an exercise in the final section of this paper.

---

[12]Note that generalized entropy indices lack this property.

[13]For example, scaling the incomes in each subgroup so that the subgroup means are equalized affects the within-group inequality term in the decomposition of the Theil index by changing the subgroup income shares. An exception is the mean log deviation, which belongs to both the generalized entropy and the Foster-Shneyerov families, and has the property that scaling the incomes in each subgroup leaves within-group inequality fixed. In fact, all Foster-Shneyerov indices have this property by design.

Proposition 7. *Lump-sum transfers between subgroups can only affect between-group inequality, but not within-group inequality for the Gini coefficient.*

Finally, we note that scaling or translating all incomes does not affect the share of aggregate inequality that is attributed to within-group or between-group inequality in the Gini decomposition. While the scale independence of the decomposition follows directly from the scale independence of the Gini coefficient and therefore applies to the decomposition of all scale-independent inequality indices, translation independence is a special feature of the Gini decomposition. This property is convenient, for example, if one is interested in decomposing inequality of income above the subsistence level, but lacks a good estimate of the level of subsistence expenses.

Proposition 8. *The Gini decomposition is both scale and translation invariant. Specifically, changing incomes at each quantile i from $y(i)$ to $\tilde{y}(i) = ay(i) + b$ does not affect the relative magnitude of within-group and between-group inequality for any $a > 0$ and $b \in \mathbb{R}$ such that all subgroups have positive average income.*

VI   Interpretation of the Gini decomposition

Just like the Gini coefficient can be interpreted geometrically and arithmetically, the decomposition of the Gini coefficient has both a geometric and an arithmetic interpretation. Both interpretations view within-group inequality as the solution to a minimization problem.

The interpretation of the Gini decomposition stems from a geometric representation of aggregating subgroup Lorenz curves. In particular, the Lorenz curve of an aggregate population can be related to the Lorenz curves of the subgroups via the geometric operation of Minkowski addition. This geometric representation of the aggregation problem unlocks a powerful theorem from convex geometry, the Brunn-Minkowski theorem, which is at the core of the arithmetic and geometric interpretations of the decomposition of the Gini coefficient.

The Gini coefficient is traditionally defined as twice the area between the Lorenz curve and the line of perfect equality. For our purposes, it is useful to consider the region within the unit square that is bounded by the Lorenz curve and its centrally reflected counterpart. Let us call this region the *Lorenz region*. Clearly, the Gini coefficient is equal to the area of the Lorenz region.

For a given partition of the population into subgroups, we can also define subgroup
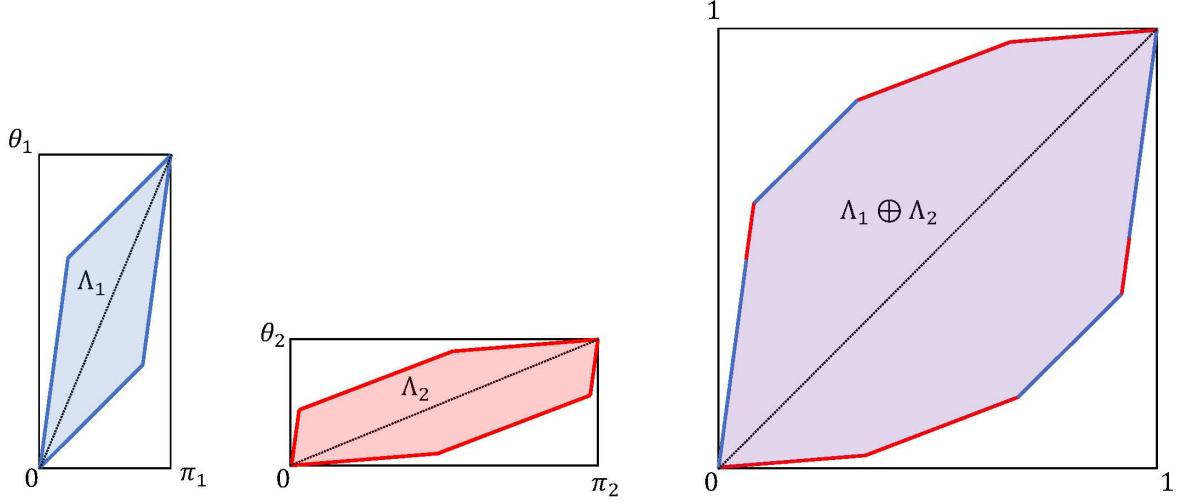
Figure 1: Aggregation of subgroup Lorenz regions. In this example, there are two different levels of income in subgroup 1 and three different levels of income in subgroup 2. In the aggregate Lorenz curve, individuals from both groups are ordered by their income. Geometrically, this corresponds to arranging the linear segments and the subgroup Lorenz curves in ascending order by their slopes. The resulting aggregate Lorenz region coincides with the Minkowski sum of the subgroup Lorenz regions.

Lorenz regions by scaling each subgroup's Lorenz region with a vector (π_k, θ_k) where π_k and θ_k are subgroup k's population and income shares, respectively.

Definition 3. *The Lorenz region Λ_k of subgroup k with population share π_k and income share θ_k is defined as the region bounded by the scaled Lorenz curve and its centrally reflected counterpart in a rectangle with side lengths π_k and θ_k.*

Proposition 9 states that the Lorenz region of the aggregate population is the Minkowski sum of the subgroup Lorenz regions. The Minkowski addition of two subgroup Lorenz regions in the case of discrete income distributions is illustrated in figure 1.

Proposition 9. *A population consisting of $k \geq 2$ subgroups with Lorenz regions $\Lambda_1, \Lambda_2, \ldots, \Lambda_K$ has an aggregate Lorenz region*

$$\Lambda = \Lambda_1 \oplus \Lambda_2 \oplus \ldots \oplus \Lambda_K,$$

where $\oplus$ denotes the Minkowski sum of sets.

The proposition is proven as part of theorem 1 in Zagier (1983).

A useful implication of representing the aggregate Gini coefficient as the area of the Minkowski sum of subgroup Lorenz regions is that we can make use of an important

15

inequality relating the areas of compact sets: the Brunn-Minkowski theorem. Applied to the case of subgroup Lorenz regions, the Brunn-Minkowski theorem provides a lower bound for the aggregate Gini coefficient in terms of subgroup Gini coefficients and aggregate characteristics.

Proposition 10 (Brunn-Minkowski theorem). *For a population with income distribution Y and Lorenz region Λ consisting of K subgroups with income distributions $Y_1, Y_2, \ldots, Y_K$, Lorenz regions $\Lambda_1, \Lambda_2, \ldots, \Lambda_K$, population shares $\pi_1, \pi_2, \ldots, \pi_K$, and income shares $\theta_1, \theta_2, \ldots, \theta_K$, we have*

$$G(Y) = |\Lambda| \geq \Big(\sum_{k=1}^{K} \sqrt{|\Lambda_k|} \Big)^2 = \Big(\sum_{k=1}^{K} \sqrt{\pi_k \theta_k G(Y_k)} \Big)^2$$

where $|\cdot|$ is the Lebesgue measure. The inequality holds as equality holds if and only if $\Lambda_1, \Lambda_2, \ldots, \Lambda_K$ are homothetic.

The subgroup Lorenz regions are homothetic if and only if the distribution of income is identical across subgroups. Note that the right-hand side of the inequality in proposition 10 coincides with the within-group inequality term in our Gini decomposition.

The geometric interpretation hence views within-group inequality as the minimal area of the population Lorenz region for given areas of the subgroup Lorenz regions. The population Lorenz region is the Minkowski sum of the subgroup Lorenz regions and its area therefore depends on both the areas of the subgroup Lorenz regions and how similar the shapes of the subgroup Lorenz regions are. The constrained minimum is achieved when the subgroup income distributions are all identical so that the subgroup Lorenz regions are homothetic. As a consequence, between-group inequality measures the excess area in the population Lorenz region resulting from non-homotheticity of the subgroup Lorenz regions. In other words, between-group inequality measures how differences in subgroup distributions contribute to aggregate inequality[14].

Within-group inequality is also the solution to an arithmetic minimization problem. For illustration, we consider a finite population. The Gini coefficient can then be defined as the sum of absolute income differences between all possible pairs of individuals in a population P, normalized by twice the product of the average income and the total

---

[14]In convex geometry, the square root of the excess area of the Minkowski sum over the Brunn-Minkowski lower bound is sometimes called the Brunn-Minkowski deficit, which has been shown to relate to the relative asymmetry of the added sets, see for example Figalli et al. (2009).

number of such pairs:

$$G = \frac{\sum_{i,j \in P} |y_i - y_j|}{2\mu N^2}. \tag{5}$$

What is the minimal sum of absolute differences in the population for given sums of absolute difference in the subgroups? Proposition 11 states the solution to this minimization problem and relates it to the within-group inequality term in our Gini decomposition. Specifically, within-group inequality in the Gini decomposition is the minimal sum of absolute differences in the population for given total absolute differences in each subgroup, normalized by twice the product of the average income and the total number of possible pairs in the population. The proof of the proposition re-states the arithmetic minimization problem in a geometric form and relates it to the Brunn-Minkowski inequality.

Proposition 11. *Let S_m be a collection of real numbers for each $m \in \{1, \ldots, M\}$. Then, the following inequality holds:*

$$\sum_{x,y \in \bigcup_{m=1}^{M} S_m} |x - y| \geq \left(\sum_{m=1}^{M} \sqrt{\sum_{x,y \in S_m} |x - y|} \right)^2. \tag{6}$$

Equality holds if and only if the distribution of numbers in each collection S_m is identical. Moreover, we have

$$G^W = \frac{1}{2\mu N^2} \left(\sum_{m=1}^{M} \sqrt{\sum_{x,y \in S_m} |x - y|} \right)^2,$$

where μ and N are the mean and the cardinality of the union of collections S_m.

The arithmetic interpretation of the Gini decomposition allows for an insightful comparison between our decomposition formula and previous formulas suggested in the literature. The formula in Bhattacharya and Mahalanobis (1967), for example, defines within-group inequality as the sum of absolute differences between all pairs of observations from the same group. This definition of within-group inequality has the undesirable property that within-group inequality is always smaller than aggregate inequality, even if all subgroups are identical. Our within-group inequality term is also related to the sums of absolute differences between all pairs of observations from the same group. Unlike the formula in Bhattacharya and Mahalanobis (1967), however, it takes into account that the sums of absolute differences in the subgroups imply a lower bound for the sum of absolute differences in the population.

Finally, the geometric representation of the aggregation of subgroup Lorenz curves also provides helpful insights into a notorious behavior of the Gini coefficient that is sometimes called *subgroup inconsistency*; an increase in subgroup inequality, while keeping subgroup means and population sizes constant, can lead to a decrease in aggregate inequality. Different examples of specific distributions and transfers that produce this behavior have been discussed in the literature (see, for example, Mookherjee and Shorrocks (1982) or Cowell (1988)). However, the conditions for the emergence of this behavior have not been stated explicitly.

Proposition 12 states that the Gini coefficient satisfies a weak subgroup consistency property. That is, any transfers within subgroups that increase subgroup inequalities according to the Lorenz criterion must increase the aggregate Gini coefficient. It follows that if transfers that increase inequality within subgroups reduce the aggregate Gini coefficient, it must be the case that there exists at least one subgroup whose new Lorenz curve intersects with the old Lorenz curve. The proof of proposition 12 makes use of the fact that outward shifts of the subgroup Lorenz curves must result in an outward shift of the population Lorenz curve.

Proposition 12 (Weak subgroup consistency of the Gini coefficient). *Implementing transfers within one or more subgroups that increase the level of subgroup inequality according to the Lorenz criterion must also increase the aggregate Gini coefficient.*

VII   Empirical Application

We demonstrate our decomposition for the Gini coefficient with two empirical applications. In particular, we study the following two questions: To what extent does household income inequality in the US population reflect inequality within "similar" households versus differences between demographic subgroups? And how do differences in male versus female earnings distributions contribute to overall earnings inequality in the United States?

It is well known that income inequality in the United States has increased substantially since the early 1980s. Some of this increase may be driven by globalization or skill-biased technical change that affects the relative demand for high-skilled and low-skilled labor. Interestingly, however, it has been shown that inequality is also increasing
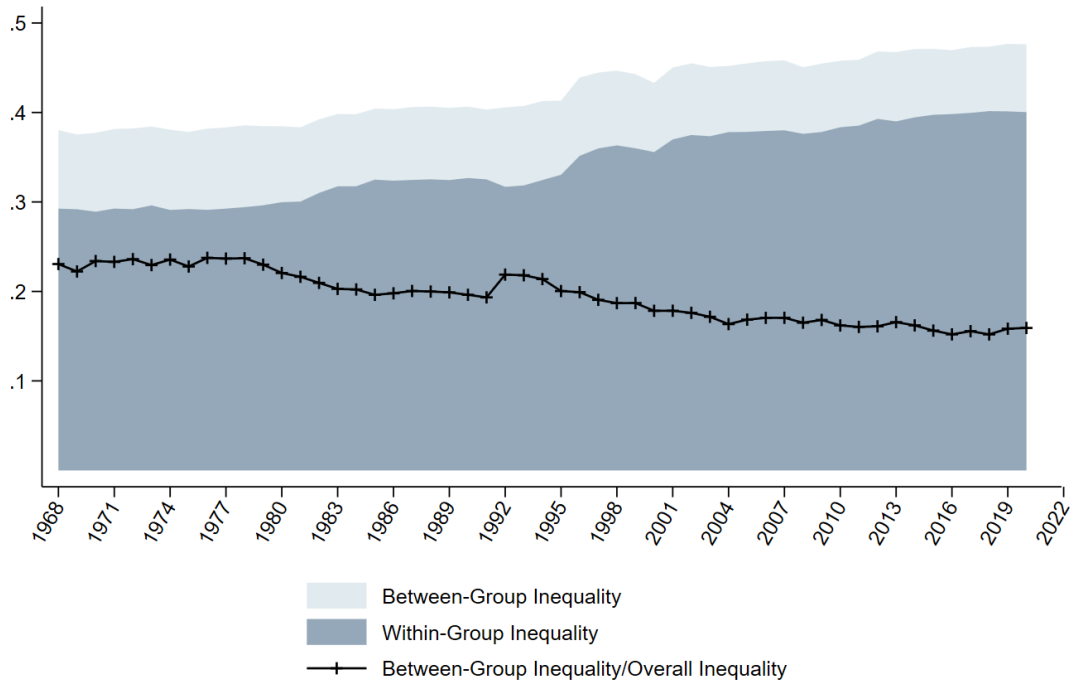
Figure 2: Household-level income inequality within and between demographic subgroups. Households are grouped by the age (≤ 35, 36–45, 46-55, 56-65,> 65), education (no college, some college, Bachelor's degree, more than Bachelor's degree), sex (male, female), and race (Black, White, other) of the household head.

within narrowly defined demographic groups.[15] It is then natural to ask what share of the aggregate inequality can be accounted for by demographic characteristics and whether that has changed over time. The results from section V allow us to study this question using the Gini coefficient as our measure of inequality.[16]

We use data on household-level income in the United States from the Current Population Surveys (CPS). To define demographic subgroups, we group households by the age, education, sex, and race of the household head. Figure 2 shows the evolution of within-group and between-group inequality over the time period 1968–2020. Over this time period, the aggregate Gini coefficient increased from 0.38 to 0.48. The rise in aggregate inequality is entirely driven by a rise in within-group inequality, which increases from 0.29 to 0.40. Between-group inequality, on the other hand, decreases slightly from 0.09 to 0.08. As a consequence, the share of aggregate inequality that can be attributed

---

[15]In an influential paper, Juhn et al. (1993) document rising wage inequality among men who are otherwise similar in terms of education and labor market experience. Since then, a large literature in labor economics devoted to explaining the rise in "residual inequality" has emerged.

[16]Cowell and Jenkins (1995) study this question using different Atkinson indices.

Figure 3: Individual-level income inequality within and between sexes. The shaded region shows between-group inequality when sex differences in average income are removed by implementing suitable lump-sump transfers.

to demographic characteristics decreases from 23 to 16 percent.

Moving from household-level income to individual-level earnings, we next ask how differences in male versus female earnings distributions contribute to overall earnings inequality in the United States. We use CPS data on individual-level earnings for people aged 18-64, working full-time, and earning more than 10% of the median, and we group individuals by sex. Figure 3 shows the evolution of within-group and between-group inequality since the beginning of our sample period. We document a strong decline in the share of the aggregate Gini coefficient that can be attributed to differences between the earnings distributions of men and women (from 13% in 1968 to 1% in 2022). Making use of the fact that lump-sum transfers between subgroups do not affect within-group inequality, we can isolate the part of between-group inequality that remains after differences in average earnings between men and women are removed. While differences in average earnings contributed significantly to the level of aggregate inequality early in the sample period, we find that removing differences in average

earnings via lump-sum transfers hardly reduces the Gini coefficient in recent years.[17] Hence, between-group inequality in the year 2022 reflects almost entirely differences in higher-order moments between the male and female earnings distributions.

VIII   Concluding Remarks

The Gini coefficient is the most prominent measure of inequality. Yet, there has been much disagreement regarding its decomposition by population subgroups. As a consequence, many researchers are relying on other inequality indices when assessing the contribution of within-group inequality and between-group differences to aggregate inequality, even if they otherwise prefer to work with the Gini coefficient.

In this paper, we have shown that the Gini coefficient, the generalized entropy indices, and the Foster-Shneyerov indices all admit satisfactory decomposition formulas derived from a common set of axioms. While our axiomatic approach recovers the established decomposition formulas for the generalized entropy and the Foster-Shneyerov indices, it has led us to a novel decomposition formula for the Gini coefficient. The decomposition of the Gini coefficient is unique given our axioms and can be interpreted both geometrically and arithmetically. Given its ease of computation, we hope that the decomposition formula for the Gini coefficient will prove to be a valuable tool for empirical research.

---

[17]This finding does not mean that the average earnings of men and women have converged. In fact, average earnings for women in 2022 remains at only 79% of that of men. These differences in means, however, are very small compared to differences in earnings among men and women.

References

Aaberge, Rolf, "Characterizations of Lorenz curves and income distributions," *Social Choice and Welfare*, 2000, *17* (4), 639–653.

_ , **Steinar Bjerve, and Kjell Doksum**, "Decomposition of rank-dependent measures of inequality by subgroups," Technical Report, Discussion Papers 2005.

Aczél, János, "On mean values," *Bulletin of the American Mathematical Society*, 1948, *54* (4), 392–400.

_ , *Lectures on functional equations and their applications*, Academic press, 1966.

_ **and J Dhombres**, *Functional equations in several variables*, Cambridge university press, 1989.

Atkinson, Anthony B, "On the measurement of inequality," *Journal of Economic Theory*, 1970, *2* (3), 244–263.

Attili, Federico, "Within-Between Decomposition of the Gini Index: A Novel Proposal," November 2021.

Bhattacharya, Nath and B Mahalanobis, "Regional disparities in household consumption in India," *Journal of the American Statistical Association*, 1967, *62* (317), 143–161.

Blackorby, Charles, David Donaldson, and Maria Auersperg, "A New Procedure for the Measurement of Inequality within and among Population Subgroups," *Canadian Journal of Economics*, 1981, *14* (4), 665–685.

Bourguignon, Francois, "Decomposable income inequality measures," *Econometrica: Journal of the Econometric Society*, 1979, pp. 901–920.

Bullen, Peter S, *Handbook of means and their inequalities*, Vol. 560, Springer Science & Business Media, 2013.

Burai, Pál, Gergely Kiss, and Patricia Szokol, "Characterization of quasi-arithmetic means without regularity condition," *Acta Mathematica Hungarica*, 2021, *165* (2), 474–485.

Costa, Michele, "Gini Index Decomposition for the Case of Two Subgroups," *Communications in Statistics - Simulation and Computation*, February 2008, *37* (4), 631–644.

Cowell, Frank A, "On the structure of additive inequality measures," *The Review of Economic Studies*, 1980, *47* (3), 521–531.

_ , "Inequality decomposition: Three bad measures," *Bulletin of Economic Research*, April 1988, *40* (4), 309–312.

_ **and Kiyoshi Kuga**, "Additivity and the entropy concept: An axiomatic approach to inequality measurement," *Journal of Economic Theory*, 1981, *25* (1), 131–143.

_ **and Stephen P Jenkins**, "How much inequality can we explain? A methodology and an application to the United States," *The Economic Journal*, 1995, *105* (429), 421–430.

Dagum, Camilo, "Inequality Measures between Income Distributions with Applications," *Econometrica*, 1980, *48* (7), 1791–1803.

_ , "Decomposition and interpretation of Gini and the generalized entropy inequality measures," *Statistica*, 1997, *57* (3), 295–308.

Deutsch, Joseph and Jacques Silber, "Decomposing income inequality by population subgroups: a generalization," in "Inequality and Poverty," Emerald Group Publishing Limited, 2007.

Figalli, Alessio, Francesco Maggi, and Aldo Pratelli, "A refined Brunn–Minkowski inequality for convex sets," in "Annales de l'Institut Henri Poincaré C, Analyse non linéaire," Vol. 26 Elsevier 2009, pp. 2511–2519.

Foster, James E and Artyom A Shneyerov, "Path independent inequality measures," *Journal of Economic Theory*, 2000, *91* (2), 199–222.

_ **and Efe A Ok**, "Lorenz Dominance and the Variance of Logarithms," *Econometrica*, 1999, *67* (4), 901–907.

Frosini, "Aggregate units, within-group inequality and the decomposition of inequality measures," *Statistica*, 1989.

_ , "Decomposition of inequality measures based on aggregate units," *Estadastica*, 2003.

Frosini, Benito V, "Ordinal Decomposition Of Inequality Measures In Case Of Dagum Distributions," in "Income and Wealth Distribution, Inequality and Poverty" Springer Berlin Heidelberg 1990, pp. 215–227.

Giorgi, Giovanni M, "The Gini inequality index decomposition. An evolutionary study," *The measurement of individual well-being and group inequalities: Essays in memory of ZM Berrebi*, 2011, pp. 185–218.

Grabisch, Michel, Jean-Luc Marichal, Radko Mesiar, and Endre Pap, "Aggregation functions: means," *Information Sciences*, 2011, *181* (1), 1–22.

i Martin, Xavier Sala, "The world distribution of income: falling poverty and... convergence, period," *The Quarterly Journal of Economics*, 2006, *121* (2), 351–397.

Juhn, Chinhui, Kevin M Murphy, and Brooks Pierce, "Wage inequality and the rise in returns to skill," *Journal of Political Economy*, 1993, *101* (3), 410–442.

Kakwani, Nanak, "The relative deprivation curve and its applications," *Journal of Business & Economic Statistics*, 1984, *2* (4), 384–394.

Lambert, Peter J and André Decoster, "The Gini coefficient reveals more," *Available at SSRN 809004*, 2005.

_ **and J Richard Aronson**, "Inequality decomposition analysis and the Gini coefficient revisited," *The Economic Journal*, 1993, *103* (420), 1221–1227.

Mangahas, M, "Income Inequality in the Philippines: A Decomposition Analysis," *Income Distribution, Employment, and Economic Development in Southeast and East Asia. The Japan Economic Research Center and Council for Asian Manpower Studies. Manila*, 1975.

Matkowski, Janusz and Zsolt Páles, "Characterization of generalized quasi-arithmetic means," *Acta Scientiarum Mathematicarum*, 2015, *81* (3), 447–456.

Mehran, Farhad, "A Statistical Analysis of Income Inequality Based on a Decomposition of the Gini Index," *Bulletin of the International Statistical Institute*, 1975, *40th Session*, 145–150.

Monti, Maria Giovanna, "A Note on the Dagum Decomposition of the Gini Inequality Index," May 2007.

Mookherjee, Dilip and Anthony Shorrocks, "A decomposition analysis of the trend in UK income inequality," *Econ. J. Nepal*, 1982, *92* (368), 886–902.

Münnich, Ákos, Gyula Maksa, and Robert J Mokken, "n-variable bisection," *Journal of Mathematical Psychology*, 2000, *44* (4), 569–581.

Murray, David, "Changes in income inequality: Australia, 1968–69 to 1973–74," *Review of Income and Wealth*, 1979, *25* (3), 309–325.

Mussard, Stéphane and Pauline Mornet, "A Note on α-Gini Measures," *Review of Income and Wealth*, 2019, *65* (3), 675–682.

Nygård, Fredrik and Arne Sandström, *Measuring income inequality*, Almqvist & Wiksell International, 1981.

Okamoto, Masato, "Decomposition of gini and multivariate gini indices," *Journal of Economic Inequality*, June 2009, *7* (2), 153–177.

Paglin, Morton, "The measurement and trend of inequality: A basic revision," *The American Economic Review*, 1975, *65* (4), 598–609.

Piesch, Walter, *Statistische Konzentrationsmaße: Formale Eigenschaften und verteilungstheoretische Zusammenhänge*, Mohr, 1975.

Pyatt, Graham, "On the Interpretation and Disaggregation of Gini Coefficients," *The Economic Journal*, 1976, *86* (342), 243–255.

Rao, V M, "Two Decompositions of Concentration Ratio," *J. R. Stat. Soc. Ser. A*, 1969, *132* (3), 418–425.

Schmüdgen, Konrad, *The moment problem*, Vol. 9, Springer, 2017.

Sen, Amartya and James Foster, *On economic inequality*, Oxford university press, 1997.

Shorrocks, Anthony F, "The class of additively decomposable inequality measures," *Econometrica: Journal of the Econometric Society*, 1980, pp. 613–625.

_ , "Inequality decomposition by population subgroups," *Econometrica: Journal of the Econometric Society*, 1984, pp. 1369–1385.

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

Silber, Jacques, "Factor Components, Population Subgroups and the Computation of the Gini Index of Inequality," *Review of Economics and Statistics*, 1989, *71* (1), 107–115.

Soltow, Lee, "The Distribution of Income Related to Changes in the Distributions of Education, Age, and Occupation," *Review of Economics and Statistics*, 1960, *42* (4), 450–453.

Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter, "Firming up inequality," *The Quarterly journal of economics*, 2019, *134* (1), 1–50.

Yitzhaki, Shlomo, "Economic distance and overlapping of distributions," *Journal of Econometrics*, March 1994, *61* (1), 147–159.

_ **and Robert I Lerman**, "Income stratification and income inequality," *Review of Income and Wealth*, September 1991, *37* (3), 313–329.

Zagier, Don, "Inequalities for the Gini coefficient of composite populations," *Journal of Mathematical Economics*, 1983, *12* (2), 103–118.

Appendix

A Proofs

A.1 Proof of Theorem 1

It is easy to verify that (1) satisfies axioms 1-6. The remainder of the proof shows that if (W_K) satisfies axioms 1-6, then (1) holds.

We first show that for given $K \geq 2$ and aggregate characteristics $(n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}_{++}^{2K}$, W_K must have a weaker functional form.

Lemma 3. *Suppose (W_K) satisfies axioms 1-6. Then, for given $K \geq 2$ and $(n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}_{++}^{2K}$, there exist a continuous increasing function $f \colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(0) = 0$ and $a_k > 0$ for all $k = 1, \ldots, K$, such that*

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = f^{-1}\bigg(\sum_{k=1}^K a_k f(I_k)\bigg) \tag{7}$$

for all $k = 1, \ldots, K$ and for all $(I_k)_{k=1}^K \in \mathbb{R}_+^K$.

Proof. Suppose $K \geq 2$ and $(n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}_{++}^{2K}$ are given and define $B \colon \mathbb{R}_+^K \to \mathbb{R}_+$ as $B(x_1, \ldots, x_K) = W_K\big((x_1, n_1, \bar{y}_1), \ldots, (x_K, n_K, \bar{y}_K)\big)$. Now, B satisfies the bisymmetry equation

$$B\big(B(x_{11}, \ldots, x_{1K}), \ldots, B(x_{K1}, \ldots, x_{KK})\big) = B\big(B(x_{11}, \ldots, x_{K1}), \ldots, B(x_{1K}, \ldots, x_{KK})\big)$$

for all $(x_{ij})_{i,j=1}^K \in \mathbb{R}_+^{K \cdot K}$. To show this, we use replacement, symmetry, and scale and population independence:

$$
\begin{aligned}
&B(B(x_{11}, \ldots, x_{1K}), \ldots, B(x_{K1}, \ldots, x_{KK})) \\
=&W_K((W_K((x_{11}, n_1, \bar{y}_1), \ldots, (x_{1K}, n_K, \bar{y}_K)), n_1, \bar{y}_1), \ldots, (W_K((x_{K1}, n_1, \bar{y}_1), \ldots, (x_{KK}, n_K, \bar{y}_K)), n_K, \bar{y}_K)) \\
=&W_K((W_K((x_{11}, \tfrac{n_1}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_1}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1), \ldots, (x_{1K}, \tfrac{n_1}{\sum_{k=1}^K n_k} n_K, \tfrac{\bar{y}_1}{\sum_{k=1}^K \bar{y}_k} \bar{y}_K)), n_1, \bar{y}_1), \\
&\ldots, (W_K((x_{K1}, \tfrac{n_K}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_K}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1), \ldots, (x_{KK}, \tfrac{n_K}{\sum_{k=1}^K n_k} n_K, \tfrac{\bar{y}_K}{\sum_{k=1}^K \bar{y}_k} \bar{y}_K)), n_K, \bar{y}_K)) \\
=&W_{K^2}((x_{11}, \tfrac{n_1}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_1}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1), \ldots, (x_{1K}, \tfrac{n_1}{\sum_{k=1}^K n_k} n_K, \tfrac{\bar{y}_1}{\sum_{k=1}^K \bar{y}_k} \bar{y}_K), \\
&\ldots, (x_{K1}, \tfrac{n_K}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_K}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1), \ldots, (x_{KK}, \tfrac{n_K}{\sum_{k=1}^K n_k} n_K, \tfrac{\bar{y}_K}{\sum_{k=1}^K \bar{y}_k} \bar{y}_K)) \\
=&W_{K^2}((x_{11}, \tfrac{n_1}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_1}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1), \ldots, (x_{K1}, \tfrac{n_K}{\sum_{k=1}^K n_k} n_1, \tfrac{\bar{y}_K}{\sum_{k=1}^K \bar{y}_k} \bar{y}_1),
\end{aligned}
$$

$$\ldots, \left(x_{1K}, \tfrac{n_1}{\sum_{k=1}^{K} n_k} n_K, \tfrac{\bar{y}_1}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_K\right), \ldots, \left(x_{KK}, \tfrac{n_K}{\sum_{k=1}^{K} n_k} n_K, \tfrac{\bar{y}_K}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_K\right)$$

$$= W_K\Big(W_K\big((x_{11}, \tfrac{n_1}{\sum_{k=1}^{K} n_k} n_1, \tfrac{\bar{y}_1}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_1), \ldots, (x_{K1}, \tfrac{n_K}{\sum_{k=1}^{K} n_k} n_1, \tfrac{\bar{y}_K}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_1), n_1, \bar{y}_1\big),$$

$$\ldots, W_K\big((x_{1K}, \tfrac{n_1}{\sum_{k=1}^{K} n_k} n_K, \tfrac{\bar{y}_1}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_K), \ldots, (x_{KK}, \tfrac{n_K}{\sum_{k=1}^{K} n_k} n_K, \tfrac{\bar{y}_K}{\sum_{k=1}^{K} \bar{y}_k} \bar{y}_K), n_K, \bar{y}_K\big)$$

$$= W_K\Big(\big(W_K((x_{11}, n_1, \bar{y}_1), \ldots, (x_{K1}, n_K, \bar{y}_K), n_1, \bar{y}_1),$$

$$\ldots, w_K((x_{1K}, n_1, \bar{y}_1), \ldots, (x_{KK}, n_K, \bar{y}_K), n_K, \bar{y}_K\big)$$

$$= B(B(x_{11}, \ldots, x_{K1}), \ldots, B(x_{1K}, \ldots, x_{KK})).$$

Thus, B is a continuous function that is strictly increasing in each of its arguments, symmetric, and satisfies bisymmetry. Following Aczél (1948), we show that we can use B to construct a function $M \colon \mathbb{R}_+^K \to \mathbb{R}_+$ that has the same aforementioned properties as B but is also reflexive. Define $F(z) \coloneqq B(z, \ldots, z)$ and $F^2(z) \coloneqq B(B(z, \ldots, z), \ldots, B(z, \ldots, z))$ for all $z \in \mathbb{R}_+$. Now, $F \colon \mathbb{R}_+ \to \mathbb{R}_+$ and $F^2 \colon \mathbb{R}_+ \to \mathbb{R}_+$ are continuous and strictly increasing functions. Thus, $F^{-1} \colon \mathbb{R}_+ \to \mathbb{R}_+$ exists and is also continuous and strictly increasing. Define $M(x_1, \ldots, x_K) = F^{-1}(B(x_1, \ldots, x_K))$. It follows that M is continuous and strictly monotonic in each x_k. Since

$$M(x, \ldots, x) = F^{-1}(B(x, \ldots, x)) = F^{-1}(F(x)) = x,$$

for any $x \in \mathbb{R}_+$, we see that M is reflexive. To show bisymmetry, let $(x_{ij})_{i,j=1}^{K} \in \mathbb{R}_+^{K \cdot K}$ and define

$$z_1 \coloneqq M(x_{11}, \ldots, x_{1K}), \text{ i.e., } B(x_{11}, \ldots, x_{1K}) = B(z_1, \ldots, z_1)$$

$$\bar{z}_1 \coloneqq M(x_{11}, \ldots, x_{K1}), \text{ i.e., } B(x_{11}, \ldots, x_{K1}) = B(\bar{z}_1, \ldots, \bar{z}_1)$$

$$\vdots$$

$$z_K \coloneqq M(x_{K1}, \ldots, x_{KK}), \text{ i.e., } B(x_{K1}, \ldots, x_{KK}) = B(z_K, \ldots, z_K)$$

$$\bar{z}_1 \coloneqq M(x_{1K}, \ldots, x_{KK}), \text{ i.e., } B(x_{1K}, \ldots, x_{KK}) = B(\bar{z}_K, \ldots, \bar{z}_K)$$

and

$$z \coloneqq M(z_1, \ldots, z_K), \text{ i.e., } B(z_1, \ldots, z_K) = B(z, \ldots, z)$$

$$\bar{z} \coloneqq M(\bar{z}_1, \ldots, \bar{z}_K), \text{ i.e., } B(\bar{z}_1, \ldots, \bar{z}_K) = B(\bar{z}, \ldots, \bar{z}).$$

By the bisymmetry of B, we get

$$
\begin{aligned}
F^2(z) = & B\big(B(z,\ldots,z),\ldots,B(z,\ldots,z)\big) \\
& B\big(B(z_1,\ldots,z_K),\ldots,B(z_1,\ldots,z_K)\big) \\
& B\big(B(z_1,\ldots,z_1),\ldots,B(z_K,\ldots,z_K)\big) \\
& B\big(B(x_{11},\ldots,x_{1K}),\ldots,B(x_{K1},\ldots,x_{KK})\big) \\
& B\big(B(x_{11},\ldots,x_{K1}),\ldots,B(x_{1K},\ldots,x_{KK})\big) \\
& B\big(B(\bar{z}_1,\ldots,\bar{z}_1),\ldots,B(\bar{z}_K,\ldots,\bar{z}_K)\big) \\
& B\big(B(\bar{z}_1,\ldots,\bar{z}_K),\ldots,B(\bar{z}_1,\ldots,\bar{z}_K)\big) \\
& B\big(B(\bar{z},\ldots,\bar{z}),\ldots,B(\bar{z},\ldots,\bar{z})\big) = F^2(\bar{z}),
\end{aligned}
$$

and hence, we get $z = \bar{z}$ which is equivalent to

$$
M\big(M(x_{11},\ldots,x_{1K}),\ldots,M(x_{K1},\ldots,x_{KK})\big) = M\big(M(x_{11},\ldots,x_{K1}),\ldots,M(x_{1K},\ldots,x_{KK})\big).
$$

Thus, M satisfies bisymmetry. We have shown that M satisfies the conditions of Theorem 2 in Münnich et al. (2000), which gives us

$$
M(x_1,\ldots,x_K) = f^{-1}\bigg(\sum_{k=1}^{K} b_k f(I_k)\bigg),
$$

for some continuous increasing function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$ such that $f(0) = 0$ and for some $b_k > 0$ such that $\sum_{k=1}^{K} b_k = 1$. Thus, $B(x_1,\ldots,x_K) = F\big(f^{-1}\big(\sum_{k=1}^{K} b_k f(I_k)\big)\big) = \psi\big(\sum_{k=1}^{K} b_k f(I_k)\big)$, where $\psi := F \circ f^{-1}$ is a strictly monotonic function. Using bisymmetry of B, we have

$$
B(B(x_1,\ldots,x_K),\ldots,B(x_1,\ldots,x_K)) = B(B(x_1,\ldots,x_1),\ldots,(x_K,\ldots,x_K))
$$

for all $(x_k)_{k=1}^{K} \in \mathbb{R}_+^{K}$, and thus,

$$
\psi\bigg(f \circ \psi\bigg(\sum_{k=1}^{K} b_k f(x_k)\bigg)\bigg) = \psi\bigg(\sum_{k=1}^{K} b_k f \circ \psi\big(f(x_k)\big)\bigg)
$$

for all $(x_k)_{k=1}^K \in \mathbb{R}_+^K$. Let $\phi := f \circ \psi$, and $z_k = f(x_k)$ for all k. Then, we have

$$\phi\left(\sum_{k=1}^K b_k z_k\right) = \sum_{k=1}^K b_k \phi(z_k)$$

for all $(z_k)_{k=1}^K \in \mathbb{R}_+^K$. This is Jensen's equation which has the unique solution $\phi(x) = ax + b$ for some $a \neq 0$ and $b \in \mathbb{R}$. Thus,

$$f\big(B(x_1,\ldots,x_K)\big) = \phi\left(\sum_{k=1}^K b_k f(x_k)\right) = \sum_{k=1}^K a_k f(x_k) + b,$$

where $a_k = ab_k$, and hence

$$B(x_1,\ldots,x_K) = f^{-1}\left(\sum_{k=1}^K a_k f(x_k) + b\right).$$

Now, since $f(0) = 0$, we have by the zero inequality axiom,

$$B(0,\ldots,0) = f^{-1}(b) = 0,$$

which implies $b = 0$. Thus, we get

$$W_K\big((I_1, n_1, \bar{y}_1),\ldots,(I_K, n_K, \bar{y}_K)\big) = f^{-1}\left(\sum_{k=1}^K a_k f(I_k)\right)$$

for given $(n_k, \bar{y}_k) \in \mathbb{R}_{++}^{2K}$ and for all $(I_k)_{k=1}^K \in \mathbb{R}_+^K$. $\qquad\square$

In lemma 3, the constants a_k and the generating function f can depend on K and $(n_k, \bar{y}_k)_{k=1}^K$. It follows from the symmetry of W_K that (a_k) and f are invariant under any permutation of $(I_k, n_k, \bar{y}_k)$.

Note that f and cf generate the same W_K for any constant $c > 0$. Thus, if $g(x) = cf(x)$ for all x, we call f and g the same generating function. We next show that the generating function is independent of K and $(n_k, \bar{y}_k)_{k=1}^K$. First, we show that the generating function does not change if we add an arbitrary subgroup to the existing population.

Lemma 4. *Suppose* $(W_K)_{K=2}^\infty$ *satisfies axioms 1-6. Let* $K \geq 2$. *Then,* $W_K\big((I_1, n_1, \bar{y}_1),\ldots,(I_K, n_K, \bar{y}_K)\big)$ *and* $W_{K+1}\big((I_1, n_1, \bar{y}_1),\ldots,(I_{K+1}, n_{K+1}, \bar{y}_{K+1})\big)$ *have the same generating function for all* $(I_k, n_k, \bar{y}_k)_{k=1}^{K+1} \in \mathbb{R}_+^{3(K+1)}$.

Proof. By replacement, we have

$$W_{K+1}\big((I_1, n_1, \bar{y}_1), \ldots, (I_{K+1}, n_{K+1}, \bar{y}_{K+1})\big)$$

$$= W_2\left(\left(W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big), \sum_{k=1}^{K} n_k, \sum_{k=1}^{K} \bar{y}_k\right), \big(I_{K+1}, n_{K+1}, \bar{y}_{K+1}\big)\right)$$

for all $(I_k, n_k, \bar{y}_k)_{k=1}^{K+1} \in \mathbb{R}_+^{3K}$. Suppose $(n_k, \bar{y}_k)_{k=1}^{K} \in \mathbb{R}_{++}^{2(K+1)}$ are given. Using lemma 3, we get

$$f^{-1}\left(\sum_{k=1}^{K+1} a_k f(I_k)\right) = g^{-1}\left(b_1 g\left(h^{-1}\left(\sum_{k=1}^{K} c_k h(I_k)\right)\right) + b_2 g(I_{K+1})\right)$$

or

$$\sum_{k=1}^{K+1} a_k f(I_k) = f \circ g^{-1}\left(b_1 g \circ h^{-1}\left(\sum_{k=1}^{K} c_k h(I_k)\right) + b_2 g(I_{K+1})\right)$$

for all $(I_k, n_k, \bar{y}_k)_{k=1}^{K+1} \in \mathbb{R}_+^{3(K+1)}$, where f, g, h are continuous and strictly increasing functions and a_k, b_k, c_k are some constants for all $k = 1, \ldots, K+1$. Set $x_k = f(I_k)$ for all $i = 1, \ldots, K+1$. Then we have

$$\sum_{k=1}^{K+1} a_k x_k = f \circ g^{-1}\left(b_1 g \circ h^{-1}\left(\sum_{k=1}^{K} c_k h \circ f^{-1}(x_k)\right) + b_2 g \circ f^{-1}(x_{K+1})\right). \qquad (8)$$

Since $f \circ g^{-1}$ and $g \circ f^{-1}$ are differentiable almost everywhere, we get

$$(f \circ g^{-1})'\left(b_1 g \circ h^{-1}\left(\sum_{k=1}^{K} c_k h \circ f^{-1}(x_k)\right) + b_2 g \circ f^{-1}(x_{K+1})\right) b_2 (g \circ f^{-1})'(x_{K+1}) = a_{K+1}$$

for almost every $(x_k)_{k=1}^{K} \in \mathbb{R}_+^K$. Now, either $(f \circ g^{-1})'$ and $(g \circ f^{-1})'$ are constants almost everywhere, or $b_1 g \circ h^{-1}\big(\sum_{k=1}^{K} c_k h \circ f^{-1}(x_k)\big) + b_2 g \circ f^{-1}(x_{K+1}) = \phi(x_{K+1})$ for some function ϕ. Because

$$b_1 g \circ h^{-1}\left(\sum_{k=1}^{K} c_k h \circ f^{-1}(x_k)\right) + b_2 g \circ f^{-1}(x_{K+1}) = b_1 g(W_K((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K))),$$

where the right-hand side is strictly increasing in each I_k and thus the left-hand side is strictly increasing in each x_k, the latter cannot be true. Thus, we get

$$(f \circ g^{-1})'(x) = c \quad a.e.$$

31

$$\Longleftrightarrow f \circ g^{-1}(x) = \int_0^x cdt = cx \quad a.e.$$

$$\Longleftrightarrow g^{-1}(x) = f^{-1}(cx) \quad a.e.$$

$$\Longleftrightarrow f(x) = cg(x) \quad a.e.$$

for some constant $c > 0$. Since f and g are continuous functions, we get $f(x) = cg(x)$ for all $x \in \mathbb{R}_+$. Plugging this result into (8) and differentiating with respect to x_k for some $k = 2, \ldots, K$, we get

$$b_1(f \circ h^{-1})'\left(\sum_{k=1}^K c_k h \circ f^{-1}(x_k)\right) c_k(h \circ f^{-1})'(x_k) = a_k$$

almost everywhere. Again, we get that either $(f \circ h^{-1})'$ and $(h \circ f^{-1})'$ are constants almost everywhere, or $b_1(f \circ h^{-1})'\left(\sum_{k=1}^K c_k h \circ f^{-1}(x_k)\right) = \phi(x_k)$ for some function ϕ. Since

$$b_1(f \circ h^{-1})'\left(\sum_{k=1}^K c_k h \circ f^{-1}(x_k)\right) = b_1 f\Big(W_K((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K))\Big),$$

where the right-hand side is strictly increasing in each I_k and thus the left-hand side is strictly increasing in each x_k, the latter cannot be true. Hence, we get that $(f \circ h^{-1})'$ is constant almost everywhere which implies $f(x) = ch(x)$ for all $x \in \mathbb{R}_+$. Therefore, $W_n((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K))$ and $W_{K+1}((I_1, n_1, \bar{y}_1), \ldots, (I_{K+1}, n_{K+1}, \bar{y}_{K+1}))$ have the same generating function for all $K \geq 2$ and for all $(I_k, n_k, \bar{y}_k)_{k=1}^{K+1}$.

$\square$

Now, it follows that the generating function is independent of K and $(n_k, \bar{y}_k)_{k=1}^K$. First, for any $I_1, n_1, \bar{y}_1, I_2, n_2, \bar{y}_2$ and $I_1', n_1', \bar{y}_1', I_2', n_2', \bar{y}_2'$, $W_2\big((I_1, n_1, \bar{y}_1), (I_2, n_2, \bar{y}_2)\big)$ and $W_2\big((I_1', n_1', \bar{y}_1'), (I_2', n_2', \bar{y}_2')\big)$ have the same generating function. To see why this is true, note that $W_2\big((I_1, n_1, \bar{y}_1), (I_2, n_2, \bar{y}_2)\big)$ and $W_4\big((I_1, n_1, \bar{y}_1), (I_2, n_2, \bar{y}_2), (I_1', n_1', \bar{y}_1'), (I_2', n_2', \bar{y}_2')\big)$ have the same generating function by lemma 4, and $W_2\big((I_1', n_1', \bar{y}_1'), (I_2', n_2', \bar{y}_2')\big)$ and $W_4\big((I_1, n_1, \bar{y}_1), (I_2, n_2, \bar{y}_2), (I_1', n_1', \bar{y}_1'), (I_2', n_2', \bar{y}_2')\big)$ have the same generating function again by lemma 4 and the symmetry of W_4. Finally, W_K has the same generating function for all $K \geq 2$ and for all $(n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}_{++}^{2K}$ since W_2 has the same generating function for all $(n_k, \bar{y}_k)_{k=1}^K$, and we can construct W_K from W_2 without changing the generating function by using lemma 4 iteratively.

Finally, we show that the constants a_k in

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = f^{-1}\left(\sum_{k=1}^{K} a_k f(I_k)\right)$$

have the form $a_k = \pi_k^{1-\alpha}\theta_k^{\alpha}$.

Lemma 5. *In equation (7), $a_k = \pi_k^{1-\alpha}\theta_k^{\alpha}$ for all $(I_k, n_k, \bar{y}_k)_{k=1}^{K+1} \in \mathbb{R}_+^{3K}$.*

Proof. Using replacement and lemmas 3 and 4, we get

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = f^{-1}\left(\sum_{k=1}^{K} a_k f(I_k)\right) = f^{-1}\left(b_1 \sum_{k=1}^{K-1} c_k f(I_k) + b_2 f(I_K)\right)$$

$$(9)$$

for all $K \geq 2$ and $(I_k, n_k, \bar{y}_k)_{k=1}^{K} \in \mathbb{R}_+^{3K}$ where $a_k = a(n_k, \bar{y}_k, \ldots, n_K, \bar{y}_K)$, $c_k = c(n_k, \bar{y}_k, \ldots, n_{K-1}, \bar{y}_{K-1})$, $b_1 = b\big(\sum_{k=1}^{K-1} n_k, \sum_{k=1}^{K-1} \bar{y}_k, n_K, \bar{y}_K\big)$, and $b_2 = b\big(n_K, \bar{y}_K, \sum_{k=1}^{K-1} n_k, \sum_{k=1}^{K-1} \bar{y}_k\big)$ for some functions a, b, and c. By setting $I_k = 0$ for $k = 1, \ldots, K-1$, we get

$$a(n_K, \bar{y}_K, \ldots) = b\left(n_K, \bar{y}_K, \sum_{k=1}^{K-1} n_k, \sum_{k=1}^{K-1} \bar{y}_k\right)$$

for all $(n_k, \bar{y}_k)_{k=1}^{K} \in \mathbb{R}_+^{2K}$. By scale and population independence, we get

$$a(n_K, \bar{y}_K, \ldots) = b(\pi_K, \theta_K, 1 - \pi_K, 1 - \theta_K) = w(\pi_K, \theta_K) \qquad (10)$$

for some function w. Due to symmetry, this result holds true for any k. By plugging (10) back into equation (9) and setting $I_j = 0$ for all $j \neq k$ for some $k = 1, \ldots, K-1$, we get

$$w\left(\frac{n_k}{\sum_{k=1}^{K} n_k}, \frac{\bar{y}_k}{\sum_{k=1}^{K} \bar{y}_k}\right) = w\left(\frac{\sum_{k=1}^{K-1} n_k}{\sum_{k=1}^{K} n_k}, \frac{\sum_{k=1}^{K-1} \bar{y}_k}{\sum_{k=1}^{K} \bar{y}_k}\right) w\left(\frac{n_k}{\sum_{k=1}^{K-1} n_k}, \frac{\bar{y}_k}{\sum_{k=1}^{K-1} \bar{y}_k}\right),$$

which generalizes to

$$w(ab, cd) = w(a, c)w(b, d)$$

for all $a, b, d, c \in (0, 1)$. Let $a = e^{x_1}$, $b = e^{x_2}$, $c = e^{x_3}$, $d = e^{x_4}$. Then, we have

$$\varphi(x_1 + x_2, x_3 + x_4) = \varphi(x_1, x_3) + \varphi(x_2, x_4)$$

for all $x_1, x_2, x_3, x_4 \in (-\infty, 0)$ where $\varphi(x, y) := \ln w(e^x, e^y)$. By Aczél and Dhombres

33

(1989) Corollary 2 on page 35, we get $\varphi(x, y) = c_1 x + c_2 y$ which implies $w(\pi, \theta) = \pi^{c_1} \theta^{c_2}$ for all $\pi, \theta \in (0, 1)$.

Now, using the weak reflexivity, we get

$$w(t\pi, t\theta) + w((1-t)\pi, (1-t)\theta) = w(\pi, \theta)$$

or

$$(t\pi)^{c_1}(t\theta)^{c_2} + ((1-t)\pi)^{c_1}((1-t)\theta)^{c_2} = \pi^{c_1}\theta^{c_2}$$

which is equivalent to

$$t^{c_1+c_2} + (1-t)^{c_1+c_2} = 1.$$

Since the left-hand side is strictly increasing in $c_1 + c_2$, then $c_1 + c_2 = 1$ is the unique solution. Thus, we have

$$W_K\big((I_1, n_1, \bar{y}_1), \ldots, (I_K, n_K, \bar{y}_K)\big) = f^{-1}\bigg(\sum_{k=1}^{K} \pi_k^{1-\alpha} \theta_k^{\alpha} f(I_k) \bigg)$$

for all $K \geq 2$ and $(I_k, n_k, \bar{y}_k)_{k=1}^{K} \in \mathbb{R}_{++}^{3K}$. $\qquad\square$

Due to continuity of W_K, these results extend to cases where $n_k = 0$ or $\bar{y}_k = 0$ for some k.

A.2   Proof of Proposition 1

The within-group inequality term in proposition 1 is an arithmetic average of subgroup inequalities with a geometric average weight structure in the subgroups' income and population shares. It therefore satisfies axioms 1 through 6 as a result of Theorem 1. Because generalized entropy indices are means-aggregative, axioms 7 requires that between-group inequality is zero when all subgroups have equal means. The between-group inequality term in proposition 1 is equal to aggregate inequality if all incomes are replaced by the relevant subgroup means and is therefore indeed zero if all subgroups have equal means. Similarly, the between-group inequality term is always constant for given aggregate characteristics and therefore satisfies axiom 8.

Next, we show that (2) is the unique decomposition for generalized entropy indices that satisfies axioms 1-8. Let $\alpha \in \mathbb{R}$ and suppose that $GE_\alpha(Y) = W(Y) + B(Y)$ is a decomposition of a generalized entropy index of parameter α that satisfies axioms 1-8. Then, by axiom 8, B is independent of subgroup income distributions. Hence, we can

set $GE_\alpha(Y_k) = 0$ for all $k = 1, \ldots, K$ without changing the value of B. By the zero within-group inequality axiom, $W(\bar{Y}) = 0$ and thus $B(Y) = GE_\alpha(\bar{Y})$. Finally, we get $W = \sum_{k=1}^{K} \pi_k^{1-\alpha} \theta_k^\alpha GE_\alpha(Y_k)$ by subtracting B from $GE_\alpha(Y)$.

A.3   Proof of Proposition 2

Let $q \in \mathbb{R}$. Since equation (3) always holds, the Foster-Shneyerov index with parameter q is Ω-aggregative with Ω containing the qth order moment.

The within-group inequality term in equation (3) satisfies Theorem 1 with $f(x) = x$ and $\alpha = 0$. Thus, (3) satisfies axioms 1-6. The inequality of power means[18] states that power means with different exponents are equal if and only if all arguments are equal. Thus, between-group inequality in (3) is zero if and only if all subgroups have equal qth order moments, that is, if and only if subgroup distributions are Ω-equivalent. Thus, axiom 7 holds. Finally, the condition in axiom 8 never holds for any Foster-Shneyerov index with $q \neq 1$, and in the case of $q = 1$, the index and the decomposition coincides with the generalized entropy index of parameter $\alpha = 0$, which is shown to satisfy all axioms in proposition 1. Thus, axiom 8 holds.

For uniqueness, let $FS_q(Y) = W(Y) + B(Y)$ be a decomposition that satisfies axioms 1-8. Now, consider a case where all subgroups are Ω-equivalent, that is, all subgroups have equal q-means. By axiom 7, between-group inequality is equal to zero and we get $W = FS_q(Y)$. Using equation (3), we get

$$FS_q(Y) = \sum_k \pi_k FS_q(Y_k) = W$$

for all $\pi_k \in [0, 1]$ and $FS_q(Y_k) \in \mathbb{R}_+$.

Using theorem 1, we get

$$f^{-1}\left(\sum_{k=1}^{K} \pi_k^{1-\alpha} \theta_k^\alpha f\left(FS_q(Y_k)\right)\right) = \sum_k \pi_k FS_q(Y_k)$$

or

$$\sum_k \pi_k^{1-\alpha} \theta_k^\alpha f\left(FS_q(Y_k)\right) = f\left(\sum_k \pi_k FS_q(Y_k)\right) \tag{11}$$

for all $\pi_k \in [0, 1]$, $\theta_k \in [0, 1]$, and $FS_q(Y_k) \in \mathbb{R}_+$, where f is a strictly monotonic function with $f(0) = 0$ and $\alpha \in \mathbb{R}$. Since cf gives the same W as f for any $c \neq 0$, we

---

[18]see e.g. Theorem 1 on p. 203 in Bullen (2013)

can assume $f(1) = 1$ without loss of generality. By setting $FS_q(Y_k) = 1$ for some k and $FS_q(Y_j) = 0$ for all $j \neq k$, we get $\pi_k^{1-\alpha}\theta_k^{\alpha} = f(\pi_k)$ for all $\pi_k, \theta_k \in [0, 1]$. Thus, we must have $\alpha = 0$ and hence $f(x) = x$ for all $x \in [0, 1]$. Now, equation (11) becomes

$$\sum_k \pi_k f(FS_q(Y_k)) = f\Big(\sum_k \pi_k FS_q(Y_k)\Big). \tag{12}$$

for all $\pi_k \in [0, 1]$ and $FS_q(Y_k) \in \mathbb{R}_+$. The only strictly monotonic solution to equation (12) is $f(x) = ax$ for some $a \neq 0$[19]. With normalization $f(1) = 1$ we get $a = 1$, and thus $W = \sum_k \pi_k FS_q(Y_k)$ and $B = FS_q(\mu_{q,i}u_{n_k})$.

A.4 Proof of Proposition 3

Suppose that $A_\varepsilon(Y) = W(Y) + B(Y)$ is a decomposition of Atkinson index of parameter ε that satisfies axioms 1-8. Since Atkinson index for any $\varepsilon < 1$ is a monotonic transformation of a corresponding generalized entropy indices, it is means-aggregative. Thus, by axiom 8, we get that B is always independent of subgroup income distributions. We may thus redistribute incomes such that $A_\varepsilon(Y_k) = 0$ for all $k = 1, \dots, K$ without changing the value of B. Hence, we get $B(Y) = A_\varepsilon(\bar{Y})$, where $\bar{Y}$ is derived from Y by replacing each individual's income by the respective subgroup mean. By subtracting B from A_ε, we get

$$W(Y) = \left(\sum_{k=1}^n \pi_k^\varepsilon \theta_k^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}} - \left(\sum_{k=1}^n \pi_k^\varepsilon \theta_k^{1-\varepsilon}(1 - A_{\varepsilon,i})^{1-\varepsilon}\right)^{\frac{1}{1-\varepsilon}}.$$

Now, W does not satisfy axiom 6 (replacement), which is a contradiction. Thus, there is no decomposition for Atkinson indices that satisfies all the proposed axioms.

A.5 Proof of Proposition 4

In the proof, we use Proposition 10 and the following lemma.

Lemma 6. *There exist two distinct distributions that share the same moments and Gini coefficients, which are all finite.*

Proof. Let X_c be a distribution with density function

$$f_c(x) = \big(1 + c\sin(2\pi \ln(x))\big) \frac{1}{\sqrt{2\pi}} \chi_{[0,\infty)}(x) \frac{1}{x} e^{-\frac{(\ln(x))^2}{2}}$$

---

[19]see Aczél (1966).

that depends on parameter c. Note that X_0 is the lognormal distribution corresponding to parameters $\mu = 0$, $\sigma = 1$. It can be shown that for $c \in [-1, 1]$, all moments of X_c are finite and do not depend on c.[20] We need to show that the Gini coefficients of X_c are equal for two different values of c. Since the Gini coefficient can be written as

$$G = 1 - 2 \int_0^1 L(F) dF,$$

it suffices to show that $\int_0^1 L_c(F) dF$ is equal for two different values of c, where L_c is the Lorenz curve of X_c.

Since

$$L_c(F(x)) = \frac{1}{\mu} \int_0^x t f_c(t) dt$$

where μ does not depend on c, we need to show that there exist two distinct values of c for which

$$A_c := \int_0^1 \int_0^{F^{-1}(x)} t f_c(t) dt dx$$

are equal. Now,

$$
\begin{aligned}
A_c &= \int_0^1 \int_0^{F^{-1}(x)} t f_c(t) dt dx = \int_0^1 \int_0^{F^{-1}(x)} (1 + c \sin(2\pi \ln(t))) \frac{1}{\sqrt{2\pi}} e^{-\frac{(\ln(t))^2}{2}} dt dx \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty \int_{F(t)}^1 (1 + c \sin(2\pi \ln(t))) e^{-\frac{(\ln(t))^2}{2}} dx dt \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty (1 + c \sin(2\pi \ln(t))) e^{-\frac{(\ln(t))^2}{2}} (1 - F(t)) dt \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty (1 + c \sin(2\pi \ln(t))) e^{-\frac{(\ln(t))^2}{2}} \left(1 - \int_0^t (1 + c \sin(2\pi \ln(x))) \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{(\ln(x))^2}{2}} dx \right) dt \\
&= \frac{1}{\sqrt{2\pi}} \int_0^\infty (1 + c \sin(2\pi \ln(t))) e^{-\frac{(\ln(t))^2}{2}} \left(1 - \int_{-\infty}^{\ln(t)} (1 + c \sin(2\pi s)) \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds \right) dt \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty (1 + c \sin(2\pi z)) e^{-\frac{z^2}{2}} e^z \left(1 - \int_{-\infty}^z (1 + c \sin(2\pi s)) \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds \right) dz \\
&= \frac{1}{\sqrt{2\pi}} (a + bc + dc^2),
\end{aligned}
$$

where

$$a = \int_{-\infty}^\infty e^{-\frac{z^2}{2}} e^z \left(1 - \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}} ds \right) dz$$

---

[20]See, for example, Schmüdgen (2017).

37

$$b = \int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{z^2}{2}}e^z dz - \int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{z^2}{2}}e^z \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}dsdz$$

$$- \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}e^z \int_{-\infty}^{z} \sin(2\pi s)\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}dsdz$$

$$c = \int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{z^2}{2}}e^z \int_{-\infty}^{z} \sin(2\pi s)\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}dsdz.$$

Now, if $b = 0$, then we have the result since either $d = 0$ and A_c does not depend on c or $d \neq 0$ and $A_c = A_{-c}$ for any $c \in [-1, 1]$.

First, note that

$$\int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{z^2}{2}}e^z dz = \int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{1}{2}(z-1)^2}e^{\frac{1}{2}}dz = e^{\frac{1}{2}}\int_{-\infty}^{\infty} \sin(2\pi x)e^{-\frac{1}{2}(x)^2}dx = 0$$

as an integral of an odd function. Thus,

$$b = \int_{-\infty}^{\infty} \sin(2\pi z)e^{-\frac{z^2}{2}}e^z \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}dsdz - \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}e^z \int_{-\infty}^{z} \sin(2\pi s)\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}}dsdz$$

$$= -e^{\frac{1}{2}}\int_{-\infty}^{\infty} \int_{-\infty}^{z} e^{-\frac{1}{2}(z-1)^2}e^{-\frac{s^2}{2}}\Big(\sin(2\pi z) + \sin(2\pi s)\Big)dsdz$$

$$= -e^{\frac{1}{2}}\int_{-\infty}^{\infty} \int_{-\infty}^{x+1} e^{-\frac{x^2}{2}}e^{-\frac{s^2}{2}}\Big(\sin(2\pi x) + \sin(2\pi s)\Big)dsdx$$

$$= -e^{\frac{1}{2}}\int_{-\infty}^{1} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}e^{-\frac{(x+t)^2}{2}}\Big(\sin(2\pi x) + \sin(2\pi(x+t))\Big)dxdt$$

$$= -e^{\frac{1}{2}}\int_{-\infty}^{1} e^{-\frac{1}{4}t^2}\int_{-\infty}^{\infty} e^{-(x+\frac{1}{2}t)^2}\Big(\sin(2\pi x) + \sin(2\pi(x+t))\Big)dxdt$$

$$= -e^{\frac{1}{2}}\int_{-\infty}^{1} e^{-\frac{1}{4}t^2}\int_{-\infty}^{\infty} e^{-(x+\frac{1}{2}t)^2}\Big(2\sin(\pi x + \pi(x+t))\cos(\pi x - (\pi(x+t)))\Big)dxdt$$

$$= -2e^{\frac{1}{2}}\int_{-\infty}^{1} e^{-\frac{1}{4}t^2}\int_{-\infty}^{\infty} e^{-(x+\frac{1}{2}t)^2}\Big(\sin(2\pi(x + \frac{1}{2}t))\cos(-\pi t)\Big)dxdt$$

$$= -2e^{\frac{1}{2}}\int_{-\infty}^{1} e^{-\frac{1}{4}t^2}\cos(-\pi t)\int_{-\infty}^{\infty} e^{-s^2}\sin(2\pi s)dsdt = 0,$$

since $\int_{-\infty}^{\infty} e^{-s^2}\sin(2\pi s)ds = 0$ as an integral of an odd function. $\square$

Suppose there exists a set of moments Ω such that the Gini coefficient is Ω-aggregative. Suppose Y_1 and Y_2 are distinct distributions that have the same values for the moments in Ω and the same Gini coefficients. Then Y_1 and Y_2 have distinct Lorenz curves. Consider a population partitioned into K subgroups such that each subgroup has distribution Y_1 and denote the vector of moments of Y_1 by Ω_1. Then, the aggregate Gini co-

efficient of the population is equal to $G = F\big(G(Y_1), \ldots, G(Y_1), \Omega_1, \ldots, \Omega_1; \pi_1, \ldots, \pi_K\big)$. By Proposition 10, we have

$$F\big(G(Y_1), \ldots, G(Y_1), \Omega_1, \ldots, \Omega_1; \pi_1, \ldots, \pi_K\big) = \Big(\sum_{k=1}^{K} \sqrt{\pi_k \theta_k G(Y_1)} \Big)^2 = G(Y_1),$$

since subgroup Lorenz curves are homothetic and $\theta_k = \pi_k$ for all k.

Next, consider another population partitioned into subgroups, where one subgroup has distribution Y_1 and the other subgroups have distribution Y_2, and let $\tilde{G}$ denote the aggregate Gini coefficient of the population. Since the subgroup Gini coefficients and moments are the same as before, we have

$$\tilde{G} = F\big(G(Y_1), \ldots, G(Y_1), \Omega_1, \ldots, \Omega_1; \pi_1, \ldots, \pi_K\big) = G(Y_1).$$

But since the subgroup Lorenz curves are not all homothetic, then Proposition 10 implies

$$\tilde{G} > G(Y_1),$$

which is a contradiction. Therefore, the Gini coefficient is Ω-nonaggregative.

A.6 Proof of Theorem 2

Let us first show that the decomposition with within-group inequality term (4) satisfies all the axioms. Since the within-group inequality term satisfies theorem 1 with $\alpha = \frac{1}{2}$ and $f(x) = x^{\frac{1}{2}}$, then it satisfies axioms 1-6.

As the Gini coefficient is Ω-nonaggregative, subgroup distributions are Ω-equivalent only if they are the same distribution. In this case, all subgroups have equal Gini coefficients and income share of each subgroup equals its population share. Then, both aggregate Gini coefficient and the within-group inequality term in (4) are equal to the common subgroup Gini coefficient. Thus, between-group inequality is equal to zero and axiom 7 holds.

Bhattacharya and Mahalanobis (1967) show that the aggregate Gini coefficient can be written as

$$G(Y) = \sum_{k} \pi_k \theta_k G(Y_k) + G(\bar{Y}) + R, \tag{13}$$

where $\bar{Y}$ is derived from Y by replacing each individuals income by the relevant subgroup mean and R depends on the amount of overlap between distributions and is zero

if and only if there is no overlap in incomes between subgroups. Hence, the condition in axiom 8 holds only if there can be no overlap in subgroup incomes. It is possible to create overlap with within-group transfers that keep subgroup Gini coefficients constant whenever two distributions have both positive population share and income share. Thus, Gini coefficient is means-aggregative only in the case of only one group, j, having both non-zero population share and non-zero income share. In this special case, equation (13) reduces to $G(Y) = \pi_j \theta_j G(Y_j) + G(\bar{Y})$ and the within-part in equation (4) reduces to $G^W(Y) = \pi_j \theta_j G(Y_j)$. Therefore, we get that $G^B(Y) = G(\bar{Y})$, which is constant under fixed aggregate characteristics. Hence, axiom 8 holds.

We next show that (4) is the unique within-group inequality term for the Gini coefficient that satisfies axioms 1-8. Suppose W is a within-part for the Gini coefficient that satisfies all eight axioms. First, by theorem 1, $W(G_1, \ldots, G_K, n_1, \ldots, n_K, \bar{y}_1, \ldots, \bar{y}_K) = f^{-1}\left(\sum_{k=1}^K \pi_k^{1-\alpha} \theta_k^\alpha f(G_k)\right)$, for all $(G_k, n_k, \bar{y}_k)_{k=1}^K \in \mathbb{R}_+^{3K}$, where f is a strictly monotonic function with $f(0) = 0$ and $\alpha \in \mathbb{R}$. Since cf gives the same W as f for any $c \neq 0$, we can assume $f(1) = 1$ without loss of generality.

Now, if there is only one group, j, with both non-zero population share and non-zero income share, then by axiom 8 and equation (13), we get

$$W = f^{-1}\left(\pi_j^{1-\alpha} \theta_j^\alpha f(G(Y_j))\right) = \pi_j \theta_j G(Y_j)$$

for all $\pi_j, \theta_j, G(Y_j) \in [0, 1]$. By taking f of both sides, we get

$$\pi_j^{1-\alpha} \theta_j^\alpha f\left(G(Y_j)\right) = f\left(\pi_j \theta_j G(Y_j)\right) \tag{14}$$

for all $\pi_j, \theta_j, G(Y_j) \in [0, 1]$. Further, by setting $G(Y_j) = 1$ we get

$$\pi_j^{1-\alpha} \theta_j^\alpha = f(\pi_j \theta_j) \tag{15}$$

for all $\pi_j, \theta_j \in [0, 1]$. By setting $\pi_j = 1$ we get $f(x) = x^\alpha$ for all $x \in [0, 1]$ and by setting $\theta_j = 1$ we get $f(x) = x^{1-\alpha}$ for all $x \in [0, 1]$. Thus, we have $\alpha = 1 - \alpha$ which implies $\alpha = \frac{1}{2}$ and $f(x) = x^{\frac{1}{2}}$ for all $x \in [0, 1]$. Therefore, the decomposition with the within-part $G^W(Y) = \left(\sum_{k=1}^n \sqrt{\pi_k \theta_k G(Y_k)}\right)^2$ is the unique decomposition for the Gini coefficient that satisfies axioms 1-8.

A.7   Proof of Proposition 5

Randomly permuting the group affiliations of a random sample of an infinite population is equivalent to replacing the incomes of that sample with a random draw from the aggregate income distribution. Equivalency between these two operations is a direct consequence of the anonymity property of inequality indices. That is, $I(P(y_1, \ldots, y_n)) = I(y_1, \ldots, y_n)$ for any permutation P. To prove proposition 5, it is useful to first prove the following two lemmas.

Lemma 7. *In an infinite population partitioned into K subgroups, drawing a random sample (with mass greater than zero) from the population and assigning the sampled observations to a new subgroup increases the within-group inequality term of the Gini coefficient (while keeping the aggregate Gini coefficient constant). That is,*

$$G^W\big((G_1, (1-\alpha)\pi_1, (1-\alpha)\theta_1), \ldots, (G_K, (1-\alpha)\pi_K, (1-\alpha)\theta_K), (G, \alpha, \alpha)\big)$$
$$\geq G^W\big((G_1, \pi_1, \theta_1), \ldots, (G_K, \pi_K, \theta_K)\big)$$

for all G_k, π_k, θ_k and $\alpha \in (0, 1]$.

Proof. Using the definition of the within-part of the Gini coefficient in equation (4) and the Brunn-Minkowski theorem, we get

$$G^W\big((G_1, (1-\alpha)\pi_1, (1-\alpha)\theta_1), \ldots, (G_K, (1-\alpha)\pi_K, (1-\alpha)\theta_K), (G, \alpha, \alpha)\big)$$
$$= \Big((1-\alpha)\sum_{k=1}^{K}\sqrt{\pi_k\theta_k G(Y_k)} + \alpha\sqrt{G(Y)}\Big)^2$$
$$\geq \Big((1-\alpha)\sum_{k=1}^{K}\sqrt{\pi_k\theta_k G(Y_k)} + \alpha\sum_{k=1}^{K}\sqrt{\pi_k\theta_k G(Y_k)}\Big)^2$$
$$= \Big(\sum_{k=1}^{K}\sqrt{\pi_k\theta_k G(Y_k)}\Big)^2$$
$$= G^W\big((G_1, \pi_1, \theta_1), \ldots, (G_K, \pi_K, \theta_K)\big).$$

$\square$

Lemma 8. *Splitting a subgroup into $m \geq 2$ subgroups weakly reduces the within-group inequality of the Gini coefficient. That is,*

$$G^W\big((G_1, \pi_1, \theta_1), \ldots, (G_K, \pi_K, \theta_K)\big)$$

$$\geq G^W\big((\tilde{G}_1, a_1\pi_1, a_1\theta_1), \ldots, (\tilde{G}_m, a_m\pi_1, a_m\theta_1), (G_2, \pi_2, \theta_2), \ldots, (G_K, \pi_K, \theta_K)\big),$$

for all G_k, π_k, θ_k, where $G_1 = G(Y_1)$, $\tilde{G}_k = G(\tilde{Y}_k)$ for $k = 1, \ldots, m$ with $Y_1 = \tilde{Y}_1 \cup \tilde{Y}_2 \cup \ldots \cup \tilde{Y}_m$, and $\sum_{j=1}^m a_j = 1$. Equality holds if and only if $\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_m$ are identical distributions.

Proof. To ease the notation, we prove the lemma for the case $m = 2$. Other cases are shown analogously.

Let Y_1 be the income distribution of subgroup 1 and let $\tilde{Y}_1 \cup \tilde{Y}_2$. Let $\tilde{\Lambda}_1$ and $\tilde{\Lambda}_2$ be the Lorenz areas of $\tilde{Y}_1$ and $\tilde{Y}_2$. Now,

$$|\tilde{\Lambda}_1 \oplus \tilde{\Lambda}_2| = |\Lambda_1| = (\tilde{\pi}_1 + \tilde{\pi}_2)(\tilde{\theta}_1 + \tilde{\theta}_2)G(\tilde{Y}_1 \cup \tilde{Y}_2),$$

where $\tilde{\pi}_k = \frac{a_k\pi_1}{\pi_1+\pi_2}$ and $\tilde{\theta}_k = \frac{a_k\theta_1}{\theta_1+\theta_2}$ for $k = 1, 2$. By the Brunn-Minkowski inequality

$$|\tilde{\Lambda}_1 \oplus \tilde{\Lambda}_2| \geq \left(\sqrt{\tilde{\pi}_1\tilde{\theta}_1\tilde{G}_1} + \sqrt{\tilde{\pi}_2\tilde{\theta}_2\tilde{G}_2}\right)^2,$$

where the inequality holds as equality if and only if the Lorenz areas of groups 1 and 2 are homothetic, or if the distributions are identical. Thus,

$$\begin{aligned}
&G^W\big((G_1, \pi_1, \theta_1), (G_2, \pi_2, \theta_2)\big) \\
&= \left(\sqrt{\pi_1\theta_1 G_1} + \sqrt{\pi_2\theta_2 G_2}\right)^2 \\
&= \left(\sqrt{(\tilde{\pi}_1 + \tilde{\pi}_2)(\tilde{\theta}_1 + \tilde{\theta}_2)G(\tilde{Y}_1 \cup \tilde{Y}_2)} + \sqrt{\pi_2\theta_2 G(Y_2)}\right)^2 \\
&= \left(\sqrt{|\tilde{\Lambda}_1 \oplus \tilde{\Lambda}_2|} + \sqrt{|\Lambda_2|}\right)^2 \\
&\geq \left(\sqrt{\tilde{\pi}_1\tilde{\theta}_1}\tilde{G}_1^{\frac{1}{2}} + \sqrt{\tilde{\pi}_2\tilde{\theta}_2}\tilde{G}_2^{\frac{1}{2}} + \sqrt{\pi_2\theta_2}G_2^{\frac{1}{2}}\right)^2 \\
&= G^W\big((\tilde{G}_1, a_1\pi_1, a_1\theta_1), (\tilde{G}_2, a_2\pi_2, a_2\theta_2), (G_2, \pi_2, \theta_2)\big),
\end{aligned}$$

with equality holding if and only if subgroups 1 and 2 have identical distributions.

$\square$

Let G denote the aggregate Gini coefficient. Now, by the subgroup replication principle, we have

$$\begin{aligned}
&G^W\big((G_1, (1-\alpha)\pi_1, (1-\alpha)\theta_1), \ldots, (G_K, (1-\alpha)\pi_K, (1-\alpha)\theta_K), (G, \alpha, \alpha)\big) \\
&= G^W\big((G_1, (1-\alpha)\pi_1, (1-\alpha)\theta_1), \ldots, (G_K, (1-\alpha)\pi_K, (1-\alpha)\theta_K), (G, \alpha\pi_1, \alpha\pi_1), \ldots, (G, \alpha\pi_K, \alpha\pi_K)\big).
\end{aligned}$$

Let $\tilde{G}_k$ denote the resulting Gini coefficient after $1 - \alpha$ sample of subgroup k's distribution is merged with $\alpha\pi_k$ sample of aggregate distribution. By lemma 8, merging any two groups must strictly increase within-group inequality unless the two groups have the same distribution of income, Thus, we get

$$G^W\big((G_1, (1-\alpha)\pi_1, (1-\alpha)\theta_1), \ldots, (G_K, (1-\alpha)\pi_K, (1-\alpha)\theta_K), (G, \alpha\pi_1, \alpha\pi_1), \ldots, (G, \alpha\pi_K, \alpha\pi_K)\big)$$
$$\leq G^W\big((\tilde{G}_1, \pi_1, \theta_1), \ldots, (\tilde{G}_K, \pi_K, \theta_K)\big)$$

Equality holds only if all subgroup have the same distribution of income. As the aggregate Gini coefficient is unchanged while the within-group inequality term increases, between-group inequality must decrease.

A.8 Proof of Proposition 6

We consider a redistribution scheme in which any income y_i is replaced by an income $\tilde{y}_i = y_i - \alpha(y_i - \mu)$ for $\alpha \in [0, 1]$. Let $\tilde{G}$ denote the Gini coefficient after this operation. First, note that the redistribution scheme reduces the aggregate Gini coefficient by α percent.

$$
\begin{aligned}
\tilde{G} &= \frac{\mathbb{E}\big[|\tilde{y}_i - \tilde{y}_j|\big]}{2\mu} \\
&= \frac{\mathbb{E}\Big[\big|y_i - \alpha(y_i - \mu) - \big(y_j - \alpha(y_j - \mu)\big)\big|\Big]}{2\mu} \\
&= (1 - \alpha)\frac{\mathbb{E}\big[|y_i - y_j|\big]}{2\mu}
\end{aligned}
$$

Now, it is easy to see that within group inequality also decreases by α percent under this redistribution scheme.

$$
\begin{aligned}
\tilde{G}^W &= \bigg(\sum_k \sqrt{\pi_k \tilde{\theta}_k \tilde{G}_k}\bigg)^2 \\
&= \bigg(\sum_k \sqrt{\pi_k \tilde{\theta}_k \frac{\mathbb{E}_k\big[|\tilde{y}_i - \tilde{y}_j|\big]}{2\tilde{\mu}_k}}\bigg)^2
\end{aligned}
$$

$$= \left(\sum_k \pi_k \sqrt{\frac{\mathbb{E}_k\left[|\tilde{y}_i - \tilde{y}_j|\right]}{2\mu}} \right)^2$$

$$= \left(\sum_k \pi_k \sqrt{(1 - \alpha)\frac{\mathbb{E}_k\left[|y_i - y_j|\right]}{2\mu}} \right)^2$$

$$= (1 - \alpha)\left(\sum_k \sqrt{\pi_k \theta_k G_k} \right)^2$$

$$= (1 - \alpha)G^W$$

where $\tilde{\theta}_k$ is subgroup k's income share after the replacement of incomes. Finally, since both the aggregate Gini coefficient as well as the within group term decrease by α percent, it must be the case the the between group terms also decreases by α percent.

A.9   Proof of Proposition 7

Note that the within group inequality term can be written as follows.

$$G^W = \left(\sum_k \sqrt{\pi_k \theta_k G_k} \right)^2$$

$$= \left(\sum_k \sqrt{\pi_k \theta_k \frac{\mathbb{E}_k\left[|y_i - y_j|\right]}{2\mu_k}} \right)^2$$

$$= \left(\sum_k \pi_k \sqrt{\frac{\mathbb{E}_k\left[|y_i - y_j|\right]}{2\mu}} \right)^2$$

But since lump-sum transfers between groups affect neither the mean absolute difference within groups nor the mean income in the population, it follows that such transfers do not affect within group inequality.

A.10   Proof of Proposition 8

Scale invariance follows directly from the scale invariance of the within-group inequality term and the inequality measure itself. To show translation invariance, note first that due to the arithmetic definition of the Gini coefficient in (5), adding some fixed amount z to each income in a population with average income equal to μ decreases the Gini coefficient by a factor of $\mu/(\mu + z)$. Let $\tilde{G}_k$ and $\tilde{\theta}_k$ denote the Gini coefficient and the

44

income share of subgroup k after translation. Within-group inequality after translation can then be written as follows

$$
\begin{aligned}
\tilde{G}^W &= \left(\sum_k \sqrt{\pi_k \tilde{\theta}_k \tilde{G}_k} \right)^2 \\
&= \left(\sum_k \sqrt{\pi_k \left(\pi_k \frac{\mu_k + z}{\mu + z} \right) \left(\frac{\mu_k}{\mu_k + z} G_k \right)} \right)^2 \\
&= \left(\sum_k \pi_k \sqrt{\frac{\mu_k}{\mu + z} G_k} \right)^2 \\
&= \frac{1}{\mu + z} \left(\sum_k \pi_k \sqrt{\mu_k G_k} \right)^2 \\
&= \frac{1}{\mu + z} \left(\sum_k \pi_k \sqrt{\frac{\theta_k}{\pi_k} \mu G_k} \right)^2 \\
&= \frac{\mu}{\mu + z} \left(\sum_k \sqrt{\pi_k \theta_k G_k} \right)^2 \\
&= \frac{\mu}{\mu + z} G^W.
\end{aligned}
$$

Within-group inequality also decreases by a same factor of $\mu/(\mu + z)$. Hence, the ratio of within-group inequality to between-group inequality remains unchanged.

A.11 Proof of Proposition 10

The inequality follows directly from the Brunn-Minkowski theorem.

A.12 Proof of Proposition 11

Let $G(S_m)$ denote the Gini coefficient computed on the collection of real numbers S_m. Similarly, let $G\left(\bigcup_{m=1}^{M} S_m \right)$ denote the Gini coefficient computed on the union of collections S_m. Inequality (6) can then be re-stated as

$$
2\mu N^2 G(S) \geq \left(\sum_{m=1}^{M} \sqrt{2\mu_m N_m^2 G(S_m)} \right)^2
$$

where μ_m and N_m are the mean and the cardinality of S_m, respectively, and μ and N are the mean and cardinality of the union of collections S_m.

Dividing both sides by $2\mu N^2$, we get

$$G(S) \geq \left(\sum_{m=1}^{M} \sqrt{\pi_m \theta_m G(S_m)} \right)^2$$

where $\pi_m = N_m/N$ and $\theta_m = (N_m \mu_n)/(N\mu)$. This inequality follows directly from proposition 10.

A.13   Proof of Proposition 12

Transfers that increase subgroup inequality by the Lorenz criterion shift the subgroup Lorenz curves outwards so that the pre-transfers Lorenz region are subsets of the post-transfers Lorenz regions. Since the Minkowski addition is monotonic, that is $|A \oplus B| \leq |\tilde{A} \oplus \tilde{B}|$ if $A \subseteq \tilde{A}$ and $B \subseteq \tilde{B}$, such transfers also shift the population Lorenz curve outwards. Hence, the aggregate Gini coefficient must increase.