GREDI

Groupe de Recherche en Économie et Développement International

Cahier de Recherche / Working Paper
10-27

Pair-Based Decomposable Inequality Measures

Stéphane Mussard

UNIVERSITÉ DE SHERBROOKE

Pair-Based Decomposable Inequality Measures*

Mussard Stéphane †
Lameta
Université Montpellier I, France

November, 2010

Abstract

Four axioms are introduced in order to characterize the family of pair-based decomposable inequality measures, which is embraced in the class of weakly decomposable inequality measures. Three axioms, namely, normalization by pairs, aggregation by pairs, and decomposition by pairs enable the pair-based family of inequality measures to be deduced and to be decomposed into within- and between-group components. The weights of population shares that bring out those within- and between-group estimators have the particularity to be unique and to sum to unity. By invoking the fourth axiom of symmetry by pairs, it is proved that pair-based inequality measures and their two decomposed components are $U-$statistics, so that, statistical information may be inferred.

Key-words and phrases: Characterization, Pair-based Aggregation, Pair-based Decomposition, $U-$statistics.

JEL Codes: D63, D31, C43.

*I am greatly indebted to Shlomo Yitzhaki for very helpful comments and suggestions. Sure, the usual disclaimer applies.

†Lameta, Université Montpellier I, Avenue Raymond Dugrand - Site de Richter - C.S. 79606, 34960 Montpellier Cedex 2 France. Tel: 33 (0)4 67 15 83 82 / Fax : 33 (0)4 67 15 84 67 - e-mail: smussard@adm.usherbrooke.ca, Associate researcher at Grédi, Université de Sherbrooke.

I. Introduction

For more than thirty years, decomposable inequality measures have been modelled on the basis of stringent axioms. Particularly, the *additive decomposition* axiom, i.e., the canonical decomposition introduced by Shorrocks (1980), has played an important role (see also Bourguignon (1979) for a less demanding requirement, i.e., the aggregation principle). It is common practice, for empirical investigations, to partition the population into many closed groups.[1] In this respect, the use of the additive decomposition axiom brings out two estimators, namely, a within-group income inequality component and a between-group income inequality one. Although this axiom provides the ability to characterize many indices available to address various empirical inquiries, the very drawback of this principle of decomposition lies in the nature of the between-group component, which exhibits an estimator based on the mean incomes of the groups only. As mentioned, e.g. in Ebert (2010), the crude structure of such a between-group component may lead to progressive (rich-to-poor) transfers, when they occur between members of distinct groups, that increase income inequalities between those groups.

Ebert (2010) introduced a weaker decomposition axiom that underlies the structure of income inequalities between each and every pairs of incomes, which are drawn from group pairwise. This attractive feature allows a new class (family) of weakly decomposable inequality measures to be proposed. It embraces the usual variance, the variance of logarithms and Gini's mean difference, among others. Precisely, Ebert (2010) introduced two new axioms: a decomposition principle and an aggregation one.

(i) The decomposition axiom postulates that the overall inequality is decomposed into within- and between-group components. Those components are weighted within-group inequality indices and weighted between-group inequality indices, respectively, where indices are specified on the basis of the inequality between income pairs. The between-group inequality component captures the inequality between each and every pairs of groups. Thus, full information about the shape of all income distributions (taken two by two) is available.

(ii) The aggregation axiom enables one to conceive the overall inequality as the inequality between one person and the rest of the population. The idea of group-pairwise inequality is still captured. Actually, as shown by Ebert (2010), this aggregation principle is a particular case of the decomposition principle (and sometimes equivalent under many conditions).

In consequence, by invoking either the first axiom or the second one, it is possible to characterize weakly decomposable inequality measures, for which the problems mentioned above (the crude between-group estimator and the difficulty linked with transfers between groups) may be solved.[2]

[1] Contrary to the deprivation literature, in which individuals compare their income to any given person belonging to particular groups they choose, that is, opened groups, the different groups are fixed *a priori* in the paper, being e.g., male/female, age groups, etc. These are, on the contrary, closed groups.

[2] It is worth mentioning that the literature is not silent about points one and two. Many materials were offered in the literature in a different manner to bring in either: the existence of another type of decomposition rule or the concept of pair-based inequality measures. On the one hand, Chameni (2006) introduced the Gini of order α (G^α) – G^α is actually equivalent to the one-parameter family of inequality measure $\hat{K}^\varepsilon$ characterized by Ebert (2010) – and demonstrated that G^α and the coefficient of variation squared are based on the same structure of decomposition, relying on three components. On the other hand, Kolm (1999) introduced the pair-based inequality measures, conceived as the aggregation of inequalities between each and every pairs of incomes. Contrary to Kolm

In this paper, I propose to conceive another class of weakly decomposable inequality measures, so-called *pair-based decomposable inequality measures*. This class will be designed with a general functional form, not with precise and operational indices, for which examples will be given only. For that purpose, I use four axioms in which the notion of pairwise combinations reaches its climax, that is, all axioms are defined on the concept of pairs. The axioms of symmetry by pairs and normalization by pairs will be of secondary importance. The two central axioms used in the characterization, related to those mentioned above, are those of decomposition and aggregation by pairs.

(i) The axiom of decomposition by pairs (DECP) postulates the idea that the overall inequality is composed of the inequality between each and every pairs of groups k and h weighted by w_{kh}. This weight will be deduced and formalized in the characterization. Picking one group k and measuring the inequality between this group k and itself produces a within-group inequality index. The within-group inequality component is then the weighted average of all within-group inequality indices weighted by w_{kk}. Again, the analytical form of the weight w_{kk} has to be deduced in the characterization such as the two inequality components (within and between) sum up to the overall income inequality exactly. DECP is quite close to Ebert's axiom. The difference lies in the structure of the weights: if $k = h$, then w_{kh} and w_{kk} coincide.

(ii) The axiom of aggregation by pairs (AGGP) introduced in the paper underlies the structure of the pair-based inequality indices. For any given group, say k, the inequality index related to group k takes into account the inequality between each and every pairs of incomes within this group. On the other hand, the between-group inequality index, e.g. the inequality between groups k and h, brings out the inequality between each and every pairs of incomes issued from groups k and h only.

In consequence, AGGP characterizes within- and between-group inequality indices, whereas DECP yields the structure of within- and between-group inequality components (with their weights w_{kk} and w_{kh}, respectively). AGGP and DECP coincide if there is one person per group. The fact that AGGP is different from the aggregation axiom used in Ebert's (2010) approach implies four fundamental results.

- Combining DECP and AGGP leads to a subclass of weakly decomposable inequality measures, the so-called pair-based decomposable inequality measures in which more concern about pairs is introduced: overall inequality is deduced from inequality between pairs of groups, and inequality indices are deduced from inequality between pairs of incomes.

- The pair-based inequality measures have the particularity to yield a unique weighting scheme. The weights w_{kk} and w_{kh}, associated with within- and between-group inequality indices respectively, sum to unity whereas this particularity does not hold systematically for all weakly decomposable inequality measures.

- The characterization is managed with a high degree of freedom since only three pair-based axioms are introduced. In particular, the principle of population traditionally used in this literature to compare distributions of heterogeneous sizes is relaxed. As it will be pointed out at the end of the paper, those pair-based axioms may be combined with traditional requirements in order to decompose well-known indices, e.g. Gini's mean difference.

(1999) and Chameni (2006), the merit of Ebert's (2010) contribution is to carry out a proper axiomatic foundation to capture both concepts.

• The within- and between-group pair-based inequality measures (and components) are proved to be $U-$statistics under few conditions. In consequence, they are asymptotically normal.

The aim of this note is to prove the above points. I present the notations and the axioms in Sections II and III, respectively. In Section IV, I expose the main results. Finally, I close in Section V.

II. Notations

Let $\mathbb{N} := \{0, 1, 2, \ldots, n\}$ be the set of nonnegative integers and let $\mathbb{N}^* := \mathbb{N} \setminus \{0\}$. Let $\mathbb{R}_+^k$ $[\mathbb{R}_+^h]$ be the $k-$dimensional $[h-$dimensional$]$ nonnegative Euclidean space, $k, h \in \mathbb{N}^*$. A pair-based inequality measure is defined to be a function $I : \mathbb{R}_+^k \times \mathbb{R}_+^h \to \mathbb{R}_+$ computed both on $k-$dimensional and $h-$dimensional income vectors $\mathbf{x}_k := (x_1, x_2, \ldots, x_{n(\mathbf{x}_k)})$, $\mathbf{x}_h := (x_1, x_2, \ldots, x_{n(\mathbf{x}_h)})$, where $n(\mathbf{x}_k)$ [resp. $n(\mathbf{x}_h)$] denotes the size of the vector $\mathbf{x}_k$ [resp. $\mathbf{x}_h$].
The population is partitioned into K groups. I denote $\mathbf{x}_k$, from now on, as the $n(\mathbf{x}_k)-$dimensional income vector of group k, $n(\mathbf{x}_k)$ being the size of group k, for all $k \in \{1, \ldots, K\}$. I also note x_{ik} [resp. x_{jh}], the income of the $i-$th person being in group k [resp. the income of the $j-$th person in group h]. For simplicity, I shall use in the sequel $I(\mathbf{x}_k, \mathbf{x}_k)$, $I(\mathbf{x}_k, \mathbf{x}_h)$ and $I(\mathbf{x}, \mathbf{x})$ for the pair-based inequality indices related to: group k for all $k \in \{1, \ldots, K\}$, to groups k and h for all $k \neq h \in \{1, \ldots, K\}$, and to the overall population, respectively. Also, $n(\mathbf{x}) \equiv n$ is used to represent the size of the overall population.

III. Axioms Based on Pairwise Combinations

III.1. Four axioms

All pair-based inequality indices are normalized. When the inequality is computed over two identical and egalitarian income vectors (not necessarily of the same size), the pair-based inequality index takes the zero value:

Axiom III.1 NorMalization by Pairs:

$$I(\mathbf{x}_k, \mathbf{x}_h) = 0 \text{ whenever } \mathbf{x}_k = c \cdot \mathbb{1}_{n(\mathbf{x}_k)} \text{ and } \mathbf{x}_h = c \cdot \mathbb{1}_{n(\mathbf{x}_h)} \text{ , for all } k, h \in \{1, 2, \ldots, K\} \quad \text{(NMP)}$$

where $\mathbb{1}_{n(\mathbf{x}_k)}$ [resp. $\mathbb{1}_{n(\mathbf{x}_h)}$] symbolizes a $n(\mathbf{x}_k)-$dimensional [resp. $n(\mathbf{x}_h)-$dimensional] vector of ones and where $c \in \mathbb{R}_+$.

In order to respect a minimal principle of justice, the pair-based inequality measures are imposed to be symmetric in groups. The symmetry by pairs of groups is an anonymity principle postulating that the pair-based inequality index does not depend on how the different groups are labeled:

Axiom III.2 SYMmetry by Pairs:

$$I(\mathbf{x}_k, \mathbf{x}_h) = I(\mathbf{x}_h, \mathbf{x}_k) \text{ for all } k, h \in \{1, \ldots, K\} \text{ .} \quad \text{(SYMP)}$$

As it will be shown in the proof below, this principle is not necessary to derive the main result. The SYMP axiom implies however more concern for statistical inference (see Corollary (IV.1)).

The pair-based inequality indices are aggregative by income pairs:

Axiom III.3 AGGregation by Pairs:

$$I(\mathbf{x}_k, \mathbf{x}_h) = f\left(\sum_{i=1}^{n(\mathbf{x}_k)} \sum_{j=1}^{n(\mathbf{x}_h)} I(x_{ik}, x_{jh})\right) \text{ for all } k, h \in \{1, \dots, K\} \;, \tag{AGGP}$$

for any given function $f\left(\cdot\;;\; n(\mathbf{x}_k), n(\mathbf{x}_h)\right)$ *such as* $f : \mathbb{R}_+ \to \mathbb{R}_+$.

The overall pair-based inequality index (computed over the entire population) is decomposable by group pairs. Formally:

Axiom III.4 DEComposition by Pairs:

$$I(\mathbf{x}, \mathbf{x}) = \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} I(\mathbf{x}_k, \mathbf{x}_h) \;, \tag{DECP}$$

where w_{kh} *is the weight associated with the pair-based inequality index* $I(\mathbf{x}_k, \mathbf{x}_h)$.

III.2. Some Remarks

The decomposition by pairs enables particular (and sometimes between-group) income transfers to be performed since it captures all inequalities between members of different groups (as well as those of the same group, see Ebert's (2009) concentration principle). It is worth mentioning that these pair-based income inequalities are rather difficult to exhibit with traditional measures based on additive decomposability, see Shorrocks (1980).[3] Ebert (2010) uses a little more demanding axiom of AGGregation:

$$I(\mathbf{x}, \mathbf{x}_{n+1}) = \gamma(n+1) I(\mathbf{x}, \mathbf{x}) + \delta(n+1) \sum_{i=1}^{n} I(x_i, \mathbf{x}_{n+1}) \;.^{4} \tag{AGG}$$

Compared with AGG, AGGP does not impose any form of separability on f. This will be done by combining AGGP with DECP. On the other hand, DECP is very closed to Ebert's axiom of DEComposition:

$$\begin{aligned} I\left(\mathbf{x}_1, \mathbf{x}_2, n(\mathbf{x}_1), n(\mathbf{x}_2)\right) &= \alpha_1\left(n(\mathbf{x})\right) I(\mathbf{x}_1, n(\mathbf{x}_1)) + \alpha_2\left(n(\mathbf{x})\right) I(\mathbf{x}_2, n(\mathbf{x}_2)) \\ &\quad + \beta\left(n(\mathbf{x})\right) \sum_{i=1}^{n(\mathbf{x}_1)} \sum_{j=1}^{n(\mathbf{x}_2)} I(x_{i1}, x_{i2}, 2) \;, \end{aligned} \tag{DEC}$$

where α_1, α_2 are the weights associated with the within-group inequality indices and β the weight associated with the between-group inequality index.

[3] Chameni (2006) proved however that particular additive decomposable indices bring out all pairwise income differences, such as the coefficient of variation squared.

[4] $\gamma(n+1)$ and $\delta(n+1) \in \mathbb{R}_{++}$ are weight functions, where $\mathbb{R}_{++}$ is $n-$dimensional positive Euclidean space.

IV. Results

Let $\mathscr{D}$ be the set of all pair-based decomposable inequality measures such as:

$$\mathscr{D} := \left\{ I : \mathbb{R}_+^n \times \mathbb{R}_+^n \to \mathbb{R}_+ \mid I \text{ respects NMP, AGGP and DECP} \right\} .$$

Theorem IV.1 *All pair-based decomposable inequality measures $I \in \mathscr{D}$ imply, for $K \in \mathbb{N}^*$:*

$$\sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} = \sum_{k=1}^{K} \sum_{h=1}^{K} \frac{n(\mathbf{x}_k) n(\mathbf{x}_h)}{[n(\mathbf{x})]^2} = 1 .$$

Proof.

I proceed in three steps. The first step is rather informative. It outlines the linearity of the f function. The second step lies in the determination of the constant of f. Finally, step 3 provides the w_{kh} weights that sum to unity.

Step 1: Adapting AGGP to the entire population yields:

$$I(\mathbf{x}, \mathbf{x}) = f\left(\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j) \right) .$$

Equalizing the last expression with DECP in which $I(\mathbf{x}_k, \mathbf{x}_h)$ is replaced by AGGP such as there is one person per group provides:

$$f\left(\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j) \right) = \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} f\left(I(x_{ik}, x_{jh}) \right) .$$

Let $u_{ij} := I(x_i, x_j)$ and $u_{ij}^{kh} := I(x_{ik}, x_{jh})$, x_{jh} being the $j-$th individual's income of group h and x_{ik} the $i-th$ individual's income of group k. If there is one person per group and if $\mathbf{x}$ is an egalitarian distribution: $x_1 = x_2 = \cdots = x_{n(\mathbf{x})}$, thus $u_{ij} = u \in \mathbb{R}_+$ for all $i, j \in \{1, 2, \ldots, n(\mathbf{x})\}$ and:

$$f([n(\mathbf{x})]^2 u) = \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} f(u) .$$

Since $u_{ij}^{kh} = u_{ij} = u$ for all $i \in \{1, \ldots, n(\mathbf{x}_k)\}$, for all $j \in \{1, \ldots, n(\mathbf{x}_h)\}$ and for all $k, h \in \{1, \ldots, K\}$, it follows that $w_{kh} =: w$ for all $k, h \in \{1, \ldots, K\}$. Therefore:

$$f([n(\mathbf{x})]^2 u) = K^2 w f(u) .$$

Remark that, because there is one individual per group: $n(\mathbf{x}) = K$. Thus a Cauchy-based equation is deduced:

$$f\left([n(\mathbf{x})]^2 u \right) = [n(\mathbf{x})]^2 w f(u)$$

Let $[n(\mathbf{x})]^2 u =: y$. Hence,

$$f(y) = w \frac{y}{u} f(u) = w \frac{y}{u} f\left(\frac{y}{[n(\mathbf{x})]^2} \right) .$$

Then,

$$f(u) = f\left(\frac{y}{[n(\mathbf{x})]^2}\right) = \frac{u}{wy} f(y) \ .$$

Let $y = 1$ and $\frac{f(1)}{w} =: c$, so finally:

$$f(u) = cu \ . \tag{S1}$$

This equation holds for groups with one individual only. In step 2 below, I generalize the result to any number of individuals per group.

Step 2: The second step also aims at determining the constant c. For that purpose, I split Step 2 into two parts: **(i)** an induction reasoning with just one group in the population, and **(ii)**, an induction reasoning with two groups in the population.

(i) There is one group in the population $\mathbf{x}_k = \mathbf{x}$ and the AGGP axiom can be rewritten as, equivalently:

$$I(\mathbf{x}_k, \mathbf{x}_k) = f\left(\sum_{i=1}^{n(\mathbf{x}_k)} \sum_{j=1}^{n(\mathbf{x}_k)} I(x_{ik}, x_{jk})\right) \iff I(\mathbf{x}, \mathbf{x}) = f\left(\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j)\right) \ .$$

- $n(\mathbf{x}) = 1$: Consider one person in the population (group k) with income x_i, hence:

$$I(x_i, x_i) = f\left(I(x_i, x_i)\right) \iff f = \mathrm{Id}_{\mathbb{R}_+} \ , \ \text{if} \ \ n(\mathbf{x}) = 1 \ . \tag{S2.a}$$

- $n(\mathbf{x}) = 2$: Suppose the population is composed of two individuals i and j only, with income $x_i = x_j$ such as i and j belong to the sole group k. Applying AGGP to $I(\mathbf{x}_k, \mathbf{x}_k)$ or equivalently to $I(\mathbf{x}, \mathbf{x})$ produces:

$$I(\mathbf{x}_k, \mathbf{x}_k) = I(\mathbf{x}, \mathbf{x}) = f\left(I(x_i, x_i) + I(x_i, x_j) + I(x_j, x_i) + I(x_j, x_j)\right) = f\left(4I(x_i, x_j)\right) \ .$$

  As $x_i = x_j$, invoking NMP gives:

$$I(\mathbf{x}, \mathbf{x}) = 0 = I(x_j, x_i) \ .$$

  It follows that:

$$I(x_j, x_i) = f\left(4I(x_i, x_j)\right) \iff u = f\left(4u\right) \ .$$

  The f function is defined as $f(\cdot \ ; \ n(\mathbf{x}), n(\mathbf{x}))$, such as $f : \mathbb{R}_+ \to \mathbb{R}_+$ and $n(\mathbf{x}) = 2$. Letting $t = 4u$, I deduce two possible functional forms:

$$f(t) = \begin{cases} \frac{1}{n(\mathbf{x})n(\mathbf{x})} \cdot t \\ \frac{1}{2n(\mathbf{x})} \cdot t \ . \end{cases} \tag{S2.b}$$

  The unique solution being consistent with Eq.(S2.a) is the first solution of Eq.(S2.b). Indeed, substituting $n(\mathbf{x}) = 1$ into $f(t) = \frac{1}{n(\mathbf{x})n(\mathbf{x})} \cdot t$ brings out $f(t) = t$, for all $t \in \mathbb{R}_+$. On the contrary, plugging $n(\mathbf{x}) = 1$ into the second solution of Eq.(S2.b), i.e., $f(t) = \frac{1}{2n(\mathbf{x})} \cdot t$ gives $f(t) = \frac{1}{2} \cdot t$, which contradicts Eq.(S2.a). Then AGGP outlines the following solution:

$$f(t) = \frac{1}{n(\mathbf{x})n(\mathbf{x})} \cdot t, \ \text{for all} \ t \in \mathbb{R}_+ \ \text{and} \ n(\mathbf{x}) \in \{1, 2\} \ . \tag{S2.c}$$

- $\underline{n(\mathbf{x}) = n-1}$: Suppose the relationship above holds for all integers $n(\mathbf{x}) \in \{1, 2, \dots, n-1\}$:

$$f(t) = \frac{1}{n(\mathbf{x})n(\mathbf{x})} \cdot t, \text{ for all } t \in \mathbb{R}_+ \text{ and } n(\mathbf{x}) \in \{1, 2, \dots, n-1\} \ . \tag{S2.d}$$

- $\underline{n(\mathbf{x}) = n}$: Suppose now that n individuals have the same income: $x_1 = \cdots = x_i = x_j = \cdots = x_n$. I deduce from AGGP and NMP, as before, that:

  $$I(x_i, x_j) = 0 = I(\mathbf{x}, \mathbf{x}) = f\left(I(x_1, x_1) + \cdots + I(x_i, x_j) + \cdots + I(x_n, x_n)\right) = f\left(n^2 I(x_i, x_j)\right) \ .$$

  Letting $t = n^2 u$ yields:

  $$f(t) = \frac{1}{n(\mathbf{x})n(\mathbf{x})} \cdot t, \text{ for all } t \in \mathbb{R}_+ \text{ and } n(\mathbf{x}) \in \mathbb{N}^* \ . \tag{S2.e}$$

  Equation (S2.a) respects the relation depicted in equation (S2.c). As equation (S2.c) is true and equation (S2.e) is also true, thus equation (S2.e) is true for all integers $n \in \mathbb{N}^*$. Remembering that one group was postulated (that is $\mathbf{x} = \mathbf{x}_k$), thereby the solution may be rewritten for any given $k \in \{1, 2, \dots, K\}$ such as:

  $$f(t) = \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_k)} \cdot t, \text{ for all } t \in \mathbb{R}_+ \ , \text{ for all } k \in \{1, 2, \dots, K\} \text{ and } n(\mathbf{x}_k) \in \mathbb{N}^* \ . \tag{S2.f}$$

(ii) I now consider two groups in the population with income vectors $\mathbf{x}_1$ and $\mathbf{x}_2$.

- $\underline{n(\mathbf{x}_1) = n(\mathbf{x}_2) = 1}$: Consider one person per group (i in group 1 and j in group 2), such as $x_{i1} = x_{j2}$. Using AGGP and the same reasoning as before provides:

  $$I(x_{i1}, x_{j2}) = f\left(I(x_{i1}, x_{j2})\right) \iff f = \mathrm{Id}_{\mathbb{R}_+} \text{ if } n(\mathbf{x}_1) = n(\mathbf{x}_2) = 1 \ . \tag{S2.g}$$

- $\underline{n(\mathbf{x}_1) = n(\mathbf{x}_2) = 2}$: Using the same approach as in **(i)**, that is, with equal incomes and the application of NMP, i.e., $I(x_{i1}, x_{j2}) = u = 0$ for all $i, j \in \{1, 2\}$ yields:

  $$I(x_{i1}, x_{j2}) = f\left(4I(x_{i1}, x_{j2})\right) \iff u = f(4u) \ .$$

  I retrieve two similar solutions:

  $$f(t) = \begin{cases} \frac{1}{n(\mathbf{x}_1)n(\mathbf{x}_2)} \cdot t \\ \frac{1}{n(\mathbf{x}_1)+n(\mathbf{x}_2)} \cdot t \ . \end{cases} \tag{S2.h}$$

  Remark again that the second solution in (S2.h) can be dropped since it contradicts (S2.g) [just replace $n(\mathbf{x}_1)$ and $n(\mathbf{x}_2)$ by 1].

- $\underline{n(\mathbf{x}_1) = n_1 - 1, \ (\mathbf{x}_2) = n_2 - 1}$: Consider the solution is true for any given $n(\mathbf{x}_1) \in \{1, 2, \dots, n_1 - 1\}$ and $n(\mathbf{x}_2) \in \{1, 2, \dots, n_2 - 1\}$:

  $$f(t) = \frac{1}{n(\mathbf{x}_1)n(\mathbf{x}_2)} \cdot t \text{ for all } t \in \mathbb{R}_+, \ n(\mathbf{x}_1) \in \{1, 2, \dots, n_1 - 1\}, \ n(\mathbf{x}_2) \in \{1, 2, \dots, n_2 - 1\} \ . \tag{S2.i}$$

- $\underline{n(\mathbf{x}_1) = n_1,\ n(\mathbf{x}_2) = n_2}$: Suppose that all individuals $(n_1 + n_2)$ have all the same income: $x_{11} = \cdots = x_{i1} = \cdots = x_{n_1 1} = x_{12} = \ldots = x_{j2} = \cdots = x_{n_2 2}$. I deduce from AGGP and NMP:

$$I(x_{i1}, x_{j2}) = 0 = I(\mathbf{x}_1, \mathbf{x}_2) = f\left(\sum_{i=1}^{n(\mathbf{x}_1)} \sum_{j=1}^{n(\mathbf{x}_2)} I(x_{i1}, x_{j2})\right) = f\left(n_1 n_2 \cdot I(x_{i1}, x_{j2})\right) \ .$$

Hence, letting $t = n_1 n_2 \cdot u$ gives:

$$f(t) = \frac{1}{n(\mathbf{x}_1)n(\mathbf{x}_2)} \cdot t \ , \ \text{for all } t \in \mathbb{R}_+ \text{ and } n(\mathbf{x}_1),\ n(\mathbf{x}_2) \in \mathbb{N}^* \ . \tag{S2.j}$$

Equation (S2.g) respects the first relation depicted in (S2.h). As (S2.h) is true and (S2.j) is also true, thus equation (S2.j) is true for all integers in $\mathbb{N}^*$. Remark that the result is available for any given $\mathbf{x}_1 = \mathbf{x}_k$ and $\mathbf{x}_2 = \mathbf{x}_h$, $k \neq h \in \{1, \ldots, K\}$:

$$f(t) = \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} \cdot t \text{ for all } t \in \mathbb{R}_+ \ , \ \text{for all } k \neq h \in \{1, \ldots, K\} \text{ and } n(\mathbf{x}_1),\ n(\mathbf{x}_2) \in \mathbb{N}^* \ . \tag{S2.k}$$

- Recap: Merging now the results of **(i)** and **(ii)**, that is, equations (S2.f) and (S2.k), the most general solution of f is given by:

$$f(t) = \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} \cdot t, \text{ for all } t \in \mathbb{R}_+, \text{ for all } h, k \in \{1, 2, \ldots, K\} \text{ and } n(\mathbf{x}_k), n(\mathbf{x}_h) \in \mathbb{N}^* \ . \tag{\textbf{S2}}$$

Note, from (**S1**) [and also from (**S2**)] that f is monotone, consequently f^{-1} exists. Therefore, I can proceed to Step 3.

Step 3: Combining AGGP and DECP yields:

$$I(\mathbf{x}, \mathbf{x}) = f\left(\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j)\right) = \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} I(\mathbf{x}_k, \mathbf{x}_h) \ .$$

As $f^{-1} : \mathbb{R}_+ \to \mathbb{R}_+$ is valued to be $f^{-1}(t) = [n(\mathbf{x})]^2 \cdot t$ [see (S2.e)], this entails:

$$\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j) = f^{-1}\left(\sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} I(\mathbf{x}_k, \mathbf{x}_h)\right) = [n(\mathbf{x})]^2 \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} I(\mathbf{x}_k, \mathbf{x}_h) \ .$$

Using AGGP again yields:

$$\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j) = [n(\mathbf{x})]^2 \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} \cdot f\left(\sum_{i=1}^{n(\mathbf{x}_k)} \sum_{j=1}^{n(\mathbf{x}_h)} I(x_{ik}, x_{jh})\right) \ .$$

From (**S2**), it follows that:

$$\sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} I(x_i, x_j) = [n(\mathbf{x})]^2 \sum_{k=1}^{K} \sum_{h=1}^{K} w_{kh} \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} \sum_{i=1}^{n(\mathbf{x}_k)} \sum_{j=1}^{n(\mathbf{x}_h)} I(x_{ik}, x_{jh}) \ .$$

Rearranging the terms gives:

$$\sum_{i=1}^{n(\mathbf{x})}\sum_{j=1}^{n(\mathbf{x})} I(x_i,x_j) = \sum_{k=1}^{K}\sum_{h=1}^{K}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_h)} I(x_{ik},x_{jh})\frac{w_{kh}\,[n(\mathbf{x})]^2}{n(\mathbf{x}_k)n(\mathbf{x}_h)} \ .$$

Note that:

$$\sum_{i=1}^{n(\mathbf{x})}\sum_{j=1}^{n(\mathbf{x})} I(x_i,x_j) = \sum_{k=1}^{K}\sum_{h=1}^{K}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_h)} I(x_{ik},x_{jh}) \ .$$

So, finally:

$$w_{kh} = \frac{n(\mathbf{x}_k)n(\mathbf{x}_h)}{[n(\mathbf{x})]^2} \text{ for all } k,h \in \{1,2,\ldots,K\} \text{ and } n(\mathbf{x}), n(\mathbf{x}_k), n(\mathbf{x}_h) \in \mathbb{N}^*. \qquad (\mathbf{S3})$$

Since $\sum_{k=1}^{K}\sum_{h=1}^{K} n(\mathbf{x}_k)n(\mathbf{x}_h) = [n(\mathbf{x})]^2$, hence $\sum_{k=1}^{K}\sum_{h=1}^{K} w_{kh} = 1$. This completes the proof. ■

The use of the three axioms NMP, AGGP, and DECP yields the following family of pair-based inequality measures:

$$I(\mathbf{x},\mathbf{x}) = \sum_{i=1}^{n(\mathbf{x})}\sum_{j=1}^{n(\mathbf{x})} \frac{1}{[n(\mathbf{x})]^2} I(x_i,x_j) \ . \qquad (1)$$

Four attractive features related to the pair-based inequality measures may be itemized:

(i) This family is decomposable by pairs such as full information may be captured. On the one hand, the information is related to income inequalities between group pairs (*via* DECP):

$$I(\mathbf{x},\mathbf{x}) = \sum_{k=1}^{K}\sum_{h=1}^{K} w_{kh} I(\mathbf{x}_k,\mathbf{x}_h) = \underbrace{\sum_{k=1}^{K} w_{kk} I\left(\mathbf{x}_k,\mathbf{x}_k\right)}_{I_w} + \underbrace{\sum_{k=1}^{K}\sum_{\forall h\neq k} w_{kh} I(\mathbf{x}_k,\mathbf{x}_h)}_{I_{bp}} \ ,$$

where I_w and I_{bp} represent the within- and the between-group inequality components, respectively. On the other hand, the information is related to income inequalities between income pairs (*via* the intersection between DECP and AGGP):

$$I(\mathbf{x},\mathbf{x}) = \underbrace{\sum_{k=1}^{K}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_k)} \frac{1}{[n(\mathbf{x}_k)]^2} w_{kk} I(x_{ik},x_{jk})}_{I_w} + \underbrace{\sum_{k=1}^{K}\sum_{\forall h\neq k}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_h)} \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} w_{kh} I(x_{ik},x_{jh})}_{I_{bp}} \ . \qquad (2)$$

The first term I_w denotes the inequality between each and every pairs of income within each group of the population and I_{bp} represents the inequality between each and every pairs of incomes drawn from each and every pairs of groups.

(ii) The induction reasoning exposed in step 2 entails a high degree of freedom. Indeed, few axioms are used to perform the characterization. Usually, the Population Principle (PP) is invoked in order to extend the set of individuals from 1 to $n \in \mathbb{N}^*$.[5] In this approach the PP principle is removed. Besides, it is respected by many inequality indices as those presented in the third point below.

(iii) This class $\mathscr{D}$ comprises, among others, Gini's Mean Difference (GMD), the Variance of Logarithms (VL), Ebert's (2010) one-parameter family of inequality measures (L^{ε}):

$$GMD(\mathbf{x},\mathbf{x}) = \frac{1}{[n(\mathbf{x})]^2} \sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} |x_i - x_j| \ , \ n(\mathbf{x}) \geq 1$$

$$VL(\mathbf{x},\mathbf{x}) = \frac{1}{[n(\mathbf{x})]^2} \sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} \frac{|\ln x_i - \ln x_j|^2}{2} \ , \ n(\mathbf{x}) \geq 1$$

$$L^{\varepsilon}(\mathbf{x},\mathbf{x}) = \frac{1}{[n(\mathbf{x})]^2} \sum_{i=1}^{n(\mathbf{x})} \sum_{j=1}^{n(\mathbf{x})} |x_i - x_j|^{\varepsilon} \ , \text{ for all } \ \varepsilon \geq 1 \text{ and } n(\mathbf{x}) \geq 1 \ .$$

(iv) Another feature related to $\mathscr{D}$ is the possibility to combine traditional axioms of the inequality measurement literature so that inequality indices are completely characterized and decomposed into subgroups. Let me take the simple example of the characterization of the GMD decomposition. For that purpose, three well-known axioms of the literature on deprivation measures are introduced, see e.g. Ebert and Moyes (2000):

Axiom IV.1 NorMalization to 1:

$$I(\mathbf{x}_k,\mathbf{x}_h) = 1 \text{ whenever } \mathbf{x}_k = \mathbb{1}_{n(\mathbf{x}_k)} \text{ and } \mathbf{x}_h = \mathbb{0}_{n(\mathbf{x}_h)} \ , \text{ for all } k,h \in \{1,2,\dots,K\} \ , \qquad (\mathrm{NM}^1)$$

where $\mathbb{0}_{n(\mathbf{x}_h)}$ *represents a* $n(\mathbf{x}_h)-$*dimensional vector of zeros.*

Axiom IV.2 HOMogeneity of degree 1:

$$I(\lambda\mathbf{x}_k,\lambda\mathbf{x}_h) = \lambda I(\mathbf{x}_k,\mathbf{x}_h) \ , \text{ for all } \lambda \in \mathbb{R}\setminus\{0\} \text{ and } k,h \in \{1,2,\dots,K\} \ . \qquad (\mathrm{HOM}^1)$$

Axiom IV.3 INVariance:

$$I_b(\mathbf{x}_k + \delta\mathbb{1}_{n(\mathbf{x}_k)},\mathbf{x}_h + \delta\mathbb{1}_{n(\mathbf{x}_h)}) = I_b(\mathbf{x}_k,\mathbf{x}_h) \ , \text{ for all } \delta \in \mathbb{R} \text{ and } k,h \in \{1,2,\dots,K\} \ . \qquad (\mathrm{INV})$$

Example IV.1 *A pair-based inequality measure* $I \in \mathscr{D}$ *respects* NM^1, HOM^1, INV *and* SYMG *if, and only if, it is the Gini Mean Difference*

$$GMD(\mathbf{x}_k,\mathbf{x}_h) = \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} \sum_{i=1}^{n(\mathbf{x}_k)} \sum_{j=1}^{n(\mathbf{x}_h)} |x_{ik} - x_{jh}| \ , \ \forall k \neq h.$$

[5] Let $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$. Concatenating ℓ times the $\mathbf{x}$ vector yields: $\mathbf{x}^{(\ell)} = \{\underbrace{x_1,\dots,x_1}_{\ell \text{ times}},\dots,\underbrace{x_n,\dots,x_n}_{\ell \text{ times}}\}$. If the inequality measure satisfies the population principle, then:

$$I(\mathbf{x}^{(\ell)}) = I(\mathbf{x}), \text{ for all } \ell \in \mathbb{N}^*\setminus\{1\} \ . \qquad (\mathrm{PP})$$

It follows that $GMD(\mathbf{x},\mathbf{x})$ is decomposable into GMD_w and GMD_{bp} as follows:

$$GMD(\mathbf{x},\mathbf{x}) = \underbrace{\sum_{k=1}^{K} w_{kk} GMD\left(\mathbf{x}_k,\mathbf{x}_k\right)}_{GMD_w} + \underbrace{\sum_{k=1}^{K}\sum_{\forall h\neq k}^{K} w_{kh} I(\mathbf{x}_k,\mathbf{x}_h)}_{GMD_{bp}}$$

$$= \underbrace{\sum_{k=1}^{K}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_k)} \frac{1}{[n(\mathbf{x}_k)]^2} w_{kk}\left|x_{ik}-x_{jk}\right|}_{GMD_w} + \underbrace{\sum_{k=1}^{K}\sum_{\forall h\neq k}^{K}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_h)} \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)} w_{kh}\left|x_{ik}-x_{jh}\right|}_{GMD_{bp}} .$$

Proof.

• **Necessity:** Let $\xi : \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ be a pair-based inequality index being in the $\mathscr{D}$ set that satisfies NM[1], HOM[1], INV and SYMG. Suppose two groups in the population with one person per group: $\mathbf{x}_k = x_k$ and $\mathbf{x}_h = x_h$, $K = 2$. Invoking INV leads to: $\xi(x_k, x_h) = \xi(x_k + \delta, x_h + \delta)$. Let $\delta = -x_h$, thus: $\xi(x_k, x_h) = \xi(x_k - x_h, 0)$. By HOM[1], $\lambda\xi(x_k, x_h) = \xi(\lambda x_k, \lambda x_h)$. Let $\lambda = x_k - x_h$, hence: $\xi(x_k, x_h) = (x_k - x_h)\xi(1, 0)$. From NM[1], $\xi(1, 0) = 1$, therefore: $\xi(x_k, x_h) = (x_k - x_h)$. By SYMG, this yields: $\xi(x_k, x_h) = |x_k - x_h| = GMD(x_i, x_j)$. Applying the result of Theorem IV.1 gives:

$$\xi(\mathbf{x}_k,\mathbf{x}_h) = \frac{1}{n(\mathbf{x}_k)n(\mathbf{x}_h)}\sum_{i=1}^{n(\mathbf{x}_k)}\sum_{j=1}^{n(\mathbf{x}_h)} |x_{ik} - x_{jh}| = GMD(\mathbf{x}_k,\mathbf{x}_h) .$$

The decomposition of $GMD(\mathbf{x},\mathbf{x})$ into GMD_w and GMD_{bp} follows directly.

• **Sufficiency:** It is left for the reader. ■

(v) Although the symmetry by pairs (SYMP) has not been invoked in the characterization of $\mathscr{D}$, SYMP enables asymptotic properties to be inferred:

Corollary IV.1 *Consider all pair-based decomposable inequality measures $I \in \mathscr{D}$ satisfying SYMP. Choose a $n(\mathbf{x})-$dimensional sample partitioned into K disjointed i.i.d. subsamples of size $n(\mathbf{x}_k)$ [$n(\mathbf{x}_k)$ being sufficiently large] for all $k \in \{1, 2, \ldots, K\}$, then the unbiased estimators $\hat{I}(\mathbf{x},\mathbf{x})$, $\hat{I}(\mathbf{x}_k,\mathbf{x}_h)$ for all $k, h \in \{1, 2, \ldots, K\}$, $\hat{I}_w$ and $\hat{I}_{bp}$ are all $U-$statistics, consequently, asymptotically normal.*

Proof.

From Theorem IV.1, the entire class of pair-based decomposable inequality measures rewrites as follows:

$$\theta(F) = \int\int I\left(x, y\right) dF(x) dF(y) ,$$

where $F(x)$ represents the c.d.f. of the population of size $m(\mathbf{x})$. SYMP implies $I(x, y)$ to be a 'kernel' for $\theta(F)$, i.e., a symmetric function (in its two first arguments here). Consequently, following the standard results in this literature, see e.g. Hoeffding (1948), or Schechtman (2005) for related stratification measures, the following estimator of $\theta(F)$ is an $U-$statistics if it rewrites:

$$U = \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{B}} \frac{I\left(x_{\alpha 1}, \ldots, x_{\alpha m(\mathbf{x})}; x_{\beta 1}, \ldots, x_{\beta m(\mathbf{x})}\right)}{\begin{pmatrix} n(\mathbf{x}) \\ m(\mathbf{x}) \end{pmatrix} \begin{pmatrix} n(\mathbf{x}) \\ m(\mathbf{x}) \end{pmatrix}} =: \hat{I}(\mathbf{x}, \mathbf{x}) \approx I(\mathbf{x}, \mathbf{x}) \ ,$$

where $\mathcal{A}$ is the number of unordered subsets of $m(\mathbf{x})$ integers taken without replacement among the set of $n(\mathbf{x})$ elements. U is asymptotically normal when $n(\mathbf{x})$ is sufficiently large. Hence, its variance may be computed. Of course, additional information about the structure of the kernel is required, e.g., one has to check whether I is an unbiased estimator for θ in order to get precisely the expectation and the standard deviation of the estimator $\hat{I}(\mathbf{x}, \mathbf{x})$. The same demonstration applies for within-group inequality indices (if $k = h$) and between-group inequality indices (if $k \neq h$):

$$U_{kh} = \sum_{\alpha \in \mathcal{A}} \sum_{\beta \in \mathcal{B}} \frac{I\left(x_{\alpha 1}, \ldots, x_{\alpha m(\mathbf{x}_k)}; x_{\beta 1}, \ldots, x_{\beta m(\mathbf{x}_h)}\right)}{\begin{pmatrix} n(\mathbf{x}_k) \\ m(\mathbf{x}_k) \end{pmatrix} \begin{pmatrix} n(\mathbf{x}_h) \\ m(\mathbf{x}_h) \end{pmatrix}} =: \hat{I}(\mathbf{x}_k, \mathbf{x}_h) \approx I(\mathbf{x}_k, \mathbf{x}_h) \ , h, k \in \{1, 2, \ldots, K\} \ ,$$

where $\mathcal{A}$ [resp. $\mathcal{B}$] is the number of unordered subsets of $m(\mathbf{x}_k)$ [resp. $m(\mathbf{x}_h)$] integers taken without replacement among the set of $n(\mathbf{x}_k)$ [resp. $n(\mathbf{x}_h)$] elements. If U_{kh} represent unbiased estimators of $\theta(F_k, F_h)$ for all $h, k \in \{1, 2, \ldots, K\}$ where

$$\theta(F_k, F_h) = \int \int I\left(x_k, x_h\right) dF_k(x_k) dF_h(x_h) \ ,$$

and where F_k [resp. F_h] is the c.d.f. of the population of group k [resp. h], then U_{kk} for all $k \in \{1, 2, \ldots, K\}$ are one-sample $U-$statistics and U_{kh} for all $k \neq h \in \{1, 2, \ldots, K\}$ are two-sample $U-$statistics. These are asymptotically normal. Remark that the estimators of within- and between-group inequality components are linear combinations of $U-$statistics (see Eq. (2)):

$$\hat{I}_w = \sum_{k=1}^{K} w_{kk} U_{kk} \ , \ \hat{I}_{bp} = \sum_{k=1}^{K} \sum_{h \neq k}^{K} w_{kh} U_{kh} \ ,$$

consequently, these inequality components are asymptotically normal. ■

V. Concluding Remarks

In this note, the notion of pairwise combinations associated with inequality measures reaches its climax. The class of pair-based inequality measures $\mathscr{D}$ is characterized with a high degree of freedom, as in Ebert (2010) in the case of weakly decomposable inequality measures. Indeed, few conditions (actually three axioms) are laid down about the structure of the function f and the weights w_{kh}. The class $\mathscr{D}$ implies the existence of within- and between-group inequality components whose weights are unique and sum to one. Then, the overall inequality index is a convex combination of within- and between-group inequality indices.

The well-suited feature relying on the $\mathscr{D}$ class of pair-based inequality measures is the possibility to rewrite the pair-based inequality components (I_w and I_{bp}) with within- and between-group inequality indices associated with the weights w_{kk} and w_{kh}, where $w_{kh} = w_{kk}$ if $k = h$. It follows that within- and between-group indices as well as within- and between-group inequality components may be $U-$statistics if SYMG applies on both within- and between-group indices, for which the variance may be derived.

References

[1] Bourguignon, F. (1979), "Decomposable Inequality Measures", *Econometrica*, 47, 901-920.

[2] Chameni Nembua, C. (2006), "Linking Gini to Entropy: Measuring Inequality by an Interpersonal Class of Indices", *Economics Bulletin*, 4(5), 1-9.

[3] Ebert, U. (2009), "Taking Empirical Studies Seriously: The Principle of Concentration and the Measurement of Welfare and Inequality", *Social Choice and Welfare*, 32(4), 555-574.

[4] Ebert, U. (2010), "The Decomposition of Inequality Reconsidered: Weakly Decomposable Measures", *Mathematical Social Sciences*, 60(2), 94-103.

[5] Ebert, U. and P. Moyes (2000), "An axiomatic Characterization of Yitzhaki's Index of Individual Deprivation", *Economics Letters*, 68, 263-270.

[6] Hoeffding, W. (1948), "A class of Statistics With Asymptotically Normal Distributions", *Annals of Statistics*, 19, 293-325.

[7] Kolm, S.-C. (1999), *Rational Foundations of Income Inequality Measurement*, in Silber, J. (eds.), Handbook of Income Inequality Measurement, Kluwer Academic Publishers, 19-94.

[8] Schechtman, E. (2005), "Stratification: Measuring and Inference", *Communications in Statistics – Theory and Methods*, 34(11), 2133-2145.

[9] Shorrocks, A. F. (1980), "The Class of Additively Decomposable Inequality Measures", *Econometrica*, 48, 613-625.