Mathematical
Inequalities
& Applications
Volume 6, Number 1 (2003), 81–86

A NOTE ON INEQUALITY CRITERIA

ERNESTO SAVAGLIO

(*communicated by A. Laforgia*)

Abstract. The *absolute differentials ordering* (ADO) and the *relative differentials ordering* (RDO) have been introduced as suitable alternative inequality criteria to *Lorenz ordering* (LO). We provide two new alternative proofs that ADO and RDO are sub-orderings of LO. Furthermore, we point out some "paradoxical situations", where these two different partial orderings fail to rank alternative income distributions.

1. Introduction

In his seminal paper, Atkinson (1970) has provided an elegant justification for the use of Lorenz curves as a means of measuring income inequality within the utilitarian framework. LO is supported by the results of Hardy, Littlewood and Pòlya (HLP, 1934), Dasgupta, Sen and Starret (1973), Rothschild and Stiglitz (1973), and Marshall and Olkin (1979). They relate it to the *Pigou-Dalton Principle* and to the welfarist approach to inequality.

Nevertheless, several theorists, like Chateauneuf (1996) and Moyes (1994), have recently noted the existence of situations where LO fails to be a suitable inequality criterion and it gives rise to "pathological situations" in ranking income distributions. For that reason, these scholars have introduced two alternative inequality criteria to LO, denominated by Moyes (1994) as *absolute differentials ordering* (ADO) and *relative differentials ordering* (RDO). Such partial orderings have been introduced, in the literature of economic inequality, by Preston (1990) and Moyes (1994). In the field of the *Theory of Majorization*, ADO and RDO first appeared in the research of Marshall, Walkup and Wets (1967).

Our concern is to analyze the relation between these two partial orderings and the Lorenz criterion. We focus our attention on the fact that ADO and RDO are sub-orderings of LO. This means that, in some circumstances, they have a greater normative and descriptive content than LO has. But, sometimes, they contradict unquestioned and clear criteria about how to reduce inequality: e.g. the Pigou-Dalton principle.

Mathematics subject classification (2000): 26D99 *JEL classification*: D31, D63, I31.

Key words and phrases: Inequality, orderings, transfers.

I would like to thank professors C. d'Aspremont, A. Laforgia, M. Scarsini, A. Trannoy and especially G. Bloise. The usual disclaimer applies.

This text presents research results of the Belgian Program of Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the author.

Paper MIA-06-07

Furthermore, there exist some situations where they, in turn, involve “paradoxical situations” in ranking income distributions.

This paper consists of three sections. In section 2 we explain our basic definitions and crucial concepts. In section 3 we show that ADO and RDO imply LO providing an alternative proof of this well-known result. In order to prove this relation we apply a very simple technique of *Theory of Majorization* (see Marshall and Olkin, 1979 chapters 1 and 2). A further proof of such a relation is also provided in a different setting by relating these inequality criteria to the sign-change orderings (see Karlin, 1968). Moreover, in this section, we point out the pros and cons of adopting ADO and RDO. Finally, some remarks on issues related to the use of partial orderings conclude the paper.

2. Notation and definitions

We consider a given finite population $N = \{1, \dots, i, \dots, n\}$ of individuals. A ranked *income distribution*

$$x = (x_1, \dots, x_i, \dots, x_n)$$

is a finite collection of positive real numbers such that $0 < x_1 \leqslant \dots \leqslant x_i \leqslant \dots \leqslant x_n$ and $\sum_i x_i = 1$. Let x_i be interpreted as the income of individual i in the population N. The set of all ranked income distributions for the population N is denoted by $\aleph = \left\{x \in \mathbb{R}^n_+ : x_1 \leqslant \dots \leqslant x_i \leqslant \dots \leqslant x_n \text{ and } \sum_i x_i = 1\right\}$.

An *inequality criterion* $\preceq$ is an ordering on $\aleph$ and can be identified with a subset $\preceq$ of $\aleph \times \aleph$. When $x \preceq y$, we shall say that y is (weakly) more unequal than x.

The three inequality criteria considered in this work are defined below.

DEFINITION 1 (Inequality Criteria). Given two income distributions $x, y \in \aleph$, we say that:

1. x is less unequal than y in the sense of Lorenz, denoted $x \preceq_{LO} y$, whenever
$$\sum_{i=1}^{k} x_i \geqslant \sum_{i=1}^{k} y_i, \quad \text{for all } k \in \{1, \dots, n-1\};$$
2. x is less unequal than y for the absolute differentials ordering (ADO), denoted $x \preceq_{ADO} y$, if
$$x_{i+1} - x_i \leqslant y_{i+1} - y_i \ \ \forall i \in \{1, \dots, n-1\};$$
3. x is less unequal than y for the relative differentials ordering (RDO), denoted $x \preceq_{RDO} y$, if:
$$\frac{x_{i+1}}{x_i} \leqslant \frac{y_{i+1}}{y_i} \quad \forall i \in \{1, \dots, n-1\}.$$

On the set $\aleph$ of all income distribution profiles, a normative requirement to evaluate inequality is the well-known “Principle of Transfers” of Pigou-Dalton (PDT). The notion of a vector x majorized by another vector y in the sense of “Pigou-Dalton” implies that, denoting the income of individual by y_k, if $y_i < y_j$ and if an amount δ of income is transferred from individual j to i, then income inequality is diminished provided $\delta \leqslant y_j - y_i$. When x is obtained by y through a sequence of PDT, then this is equivalent to the replacement of y_i and y_j by $y_i + \delta$ and $y_j - \delta$. Note that, if $\delta \leqslant y_j - y_i$, this amounts to the replacement of y_i and y_j by averages. Then, given x, $y \in \aleph$, we can obtain x from y through repeated averages of two incomes at a time. This is theoretically tantamount to say:

DEFINITION 2 (S-majorization). Given $x, y \in \aleph$, x is majorized by y, denoted $x \preceq_S y$, if $x = yP$ for some doubly stochastic matrix P.[1]

This characterization of vector majorization in terms of doubly stochastic matrix confirms the fact that averaging is a smoothing operation, i.e. x is a smoothing of y in the sense that each income level of x is a convex combination of income levels in y. We know by HLP (1934)[2], that LO is equivalent to PDT and S-majorization. Such a result makes LO a normative and descriptive criterion for analyzing inequality (see Sen, 1997). Hence, it is worth to pursue if this equivalence also holds for ADO and RDO.

3. The relation between ADO, RDO and LO

3.1. Comparing inequality criteria

We start to analyze why ADO is a suitable inequality criterion.

Let us consider two vectors of income profiles $x, y \in \aleph$, that are ADO-ranked. Set $x = y + h \quad \forall i \in \{1, \ldots, n\}$, where h represents a vector of elementary transfers among individuals such that $\sum_{i=1}^{n} h_i = 0$. Hence x can be obtained from y by transferring amounts of income from richer to poorer people. It shows that $x \preceq_{ADO} y$ if and only if $h_{i+1} \leqslant h_i \quad \forall i \in \{1, \ldots, n-1\}$ with $\sum_{i=1}^{n} h_i = 0$. It means that it is possible to obtain, from a given income distribution vector, another one which shows less inequality, simply by transferring shares of income from a richer individual to a poorer one in an absolute progressive way: a suitable requirement for an inequality criterion.

Analogously for RDO, if, given two vector $x, y \in \mathbb{R}^n$, we interpret y_i as the individual income before tax and $x_i = f(y_i)$ as this income vector after taxation, we can see that $x \preceq_{RDO} y$ is in accordance with progressive income taxation, if RDO is equivalent to $\frac{f(y)}{y}$ non-increasing on $(0, +\infty)$. In other words, if we suppose that x results from y through redistributive effects due to a progressive income tax, we can order the two vectors as RDO does. In such a case, RDO will appear to be a suitable inequality criterion.

Now a question arises: "If ADO and RDO are suitable inequality criteria, may we consider them as an alternative to LO?" Or, in other words, what is the relation between LO and ADO or RDO? An alternative way to check such a relation is by using a simple technique of *Theory of Majorization* as we show in the following proposition:[3]

[1] A square matrix is said to be doubly stochastic if all its entries are nonnegative and each of its rows and columns sums to one.

[2] This result is also reproduced in Dasgupta *et alii* (1973) as Lemma 2.

[3] Alternative proofs of this well-known result are for examples in Marshall and Olkin (1979) proposition B.1 on page 129 or in Chateauneuf (1996) proposition 6 on page 13.

PROPOSITION 1. *ADO and RDO are suborderings of LO.*

Proof. Let us assume x, $y \in \aleph$ with $x \neq y$ and that are ADO-ranked. It means that $y_i + h_i = x_i \ \forall i \in \{1, \ldots, n\}$, where h_i represents a vector of elementary transfers among individuals such that $h_i \geqslant h_{i+1}$ and $\sum_{i=1}^{n} h_i = 0$. This implies that there exist a k and j, such that k is the smallest index such that $x_k < y_k$ and j is the largest index, smaller than k, such that $x_j > y_j$. Such a pair k, j must exist, since the smallest index i for which $x_i \neq y_i$ must satisfy $x_i > y_i$. By choice of k and j, we have the following chain of inequalities:

$$y_j < x_j \leqslant x_k < y_k. \tag{1}$$

Let $h = \min(y_k - x_k; x_j - y_j)$, $(1-\lambda) = \frac{h}{y_k - y_j}$ and let

$$y^\circ = (y_1, \ldots, y_{j-1}, y_j + h, y_{j+1}, \ldots, y_{k-1}, y_k - h, y_{k+1}, \ldots, y_n).$$

From (1) we know $\lambda \in (0,1)$. Using some algebra, we verify that

$$y^\circ = \lambda y + (1-\lambda)(y_1, \ldots, y_{j-1}, y_k, y_{j+1}, \ldots, y_{k-1}, y_j, y_{k+1}, \ldots, y_n), \tag{2}$$

In a compact form (2) can be written as $y^\circ = yT$ for $T = \lambda I + (1-\lambda)Q$, where Q is a permutation matrix that interchanges the *jth* and *kth* coordinates. Consequently, as this T-operator is an inequality reducing transformation (thereafter T-transform), $y^\circ \prec y$. It is also true that $x \prec y^\circ$. Then this means that, by successive applications of a finite number of T-transforms, we can derive x from y through a finite product of elementary transfers of linear kind. The fact that T-transforms are doubly stochastic and the product of T-transformations is still doubly stochastic guarantees the existence of a doubly stochastic matrix P, such that $x = yP$. So ADO implies S-majorization. Then the equivalence between S-majorization and LO, proved by HLP (1934) and quoted above, points out what is the relation between ADO and LO and completes the claim.

b) The proof that RDO implies LO is essentially the same.[4] □

Thus ADO or RDO do not capture the notion of Schur transformation, because as we showed A*DO* implies *S-majorization*, while the reverse (i.e. *S-majorization* ⇒ *ADO*) does not hold. Indeed, if we take e.g. the two vectors $y = (10, 15, 20, 25)$ and $x = (11, 14, 21, 24)$, we know that y is dominated in the sense of Lorenz by x, while y is not dominated in the sense of ADO by x, denoted $x \not\succeq_{ADO} y$. Nonetheless, Lorenz majorization guarantees that there exists a square matrix P such that $yP = x$, which means that ADO or RDO are *proper* subordering of LO.

An alternative proof of the *proposition 1* can be provided by relating all binary relations in *definition 1* to the sign-changed orderings.[5]

Indeed, let $S(g)$ be the number of sign changes of the function $g(\cdot)$, which can be properly defined. A natural condition on $y - x$, corresponding to y being in some

[4] It is enough to suppose that there exists a S-*convex* function f, i.e. a transformation reducing inequality that maps y into x, multiplying the vector y for a bistochastic matrix P.

[5] See Karlin (1968) vol. I, chapters 5 and 6.

sense more variable than x, is

$$S(y-x)=1 \quad \text{with sign sequence } -,\ +.$$

For $x,\ y \in \aleph$, it is well known that $x \preceq_{LO} y$ if and only if $S(y-x) \leqslant 1$. As $x \preceq_{ADO} y$ means $h_{i+1} \leqslant h_i \quad \forall i \in \{1,\ldots,n-1\}$ with $\sum_{i=1}^{n} h_i = 0$, then $x \preceq_{ADO} y$ only if $S(y-x) \leqslant 1$ with sign sequence, $-$, $+$. Hence it is evident that ADO implies LO.[6]

3.2. Polarization and Transfers

After having shown the suitable characteristics of ADO and RDO and their relation with LO, we have to notice that if we adopt ADO and RDO as alternative inequality criteria to LO, we can no longer make use of important tools to study inequality such as the PDT or S-majorization, which have a normative and descriptive appeal as Dasgupta and *alii* (1973), Rothschild and Stiglitz (1973) and Sen (1997) pointed out. Moreover, we can observe that ADO and RDO sometimes fail to rank income distributions, giving rise to extreme "paradoxical situations".

Indeed, let us assume the existence of an hypothetical decision maker concerned for equity, who must choose between two income distributions $y=(10,15,20,25)$ and $x=(0,0,0,70)$. Then, if he decides according to the Lorenz criterion, he will prefer y to x (i.e. $y \preceq_{LO} x$). If he decides according to ADO he will be not able to conclude that x represents an evident unfair distribution with respect to y. Note that such an ambiguity of ADO and RDO does not depend on the particular distribution x. We continue to obtain such paradoxical situation with vector like $x=(1,1,1,67)$, $x=(2,2,2,62)$, etc. ADO and RDO then can not rank some high-polarized distribution, which are evidently Lorenz dominated by y and represent unequal allocations of income.

If only one transfer of income takes place from richer people to poorer ones, the transfer must be from the richest to the poorest in order to satisfying ADO: an extreme requirement for a suitable inequality criterion. Finally, ADO and RDO, not considering the distribution as a "whole", cannot rank income vectors which show small differences among quantiles: if $y=(1,1,4,5)$ and $x=(2,3,3,3)$, we can see that $x \preceq_{LO} y$, but $x \npreceq_{ADO} y$ and $x \npreceq_{RDO} y$. Analogously with $y=(3,3,5,5)$, $x=(3,4,4,5)$: $x \preceq_{LO} y$, while, $x \npreceq_{ADO} y$ and $x \npreceq_{RDO} y$. One should notice that ADO and RDO lead to paradoxes different with respect to the LO ones (see Chateauneuf, 1996). While if $x \preceq_{LO} y$ with $y=(10,15,20,25)$ and $x=(11,14,21,24)$, we *could* think y less unequal than x at least when we consider subgroups $\{2,3\}$[7], there is no doubt that $x=(0,0,0,70)$ is more unequal than y even if $x \npreceq_{ADO} y$ and $x \npreceq_{RDO} y$. In other words, ADO and RDO are "more incomplete" (partial) orderings and they lead to more "pathological situations" than LO does. Hence, they cannot be considered suitable substitutes of Lorenz ordering. ADO and RDO arouse too extreme paradoxes in ranking income distributions.

---

[6] A similar argument shows that RDO implies LO.

[7] Inequality increases as we pass from y to x (i.e. $y_3-y_2=20-15=5<x_3-x_2=21-14=7$).

4. Conclusion

The proposal of using two different inequality criteria, namely ADO and RDO, like alternatives to LO led us to focus on the characteristics of these two partial orderings. They capture more information concerning income distribution than LO does, taking into account the differences (ratios) in the income of the groups of individuals belonging to a given distribution. However, they are normatively inconsistent: in some situations, they are not coherent with some unquestioned and clear criteria like the PDT about how to reduce inequality. Moreover, some examples have showed that several paradoxes, in ranking alternative income distributions, arise when we use these two partial orderings.

The conclusion is that every partial ordering faces some situation it cannot rank. We then have to use every inequality criterion, paying much attention to its limits and overall to its normative and descriptive contents.

REFERENCES

[1] ATKINSON, A. B., *On the measurement of Inequality*, Journal of Economic Theory, **2** (1970), 244–263.

[2] CHATEAUNEUF, A., *Decreasing inequality: an approach through non-additive models*, w.p.96.58 (1996) CERMSEM, Université de Paris I.

[3] DASGUPTA, P., A. K. SEN, AND D. STARRET, *Notes on the measurement of inequality*, Journal of Economic Theory **6** (1973), 180–187.

[4] KARLIN, S., *Total Positivity, volume I*, Standford University Press, Standford, California, 1968.

[5] HARDY, G. H., J. E. LITTLEWOOD, AND G. POLYA, *Inequalities*, Cambridge University Press, London, 1934.

[6] MARSHALL, A. W., AND I. OLKIN, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.

[7] MARSHALL, A. W., D. W. WALKUP AND R. J. B. WETS, *Order-preserving functions: applications to majorization and order statistics*, Pacific Journal of Mathematics, **23** (1967), 569–585.

[8] MOYES, P., *Inequality reducing and inequality preserving transformations of incomes: symmetric and individualistic transformations*, Journal of Economic Theory, **63** (1994), 271–298.

[9] PRESTON, I., *Ratios, differences, and inequality indices*, Institute for Fiscal Studies Working Paper W90-9, 1990.

[10] ROTHSCHILD, M. AND J. E. STIGLITZ, *Some further results on the measurement of inequality*, Journal of Economic Theory, **6** (1973), 188–204.

[11] SEN, A. K., *On Economic Inequality* – (Second eds), extended edition with J. Foster, Oxford, Clarendon Press, 1997.

(Received December 15, 2001)

Ernesto Savaglio
C.O.R.E.
Louvain-la-Neuve
Belgium
e-mail: ernesto@core.ucl.ac.be