ON MULTIDIMENSIONAL INEQUALITY WITH VARIABLE DISTRIBUTION MEAN

ERNESTO SAVAGLIO

Abstract. We compare inequality of alternative populations of individuals, who differed in many characteristics besides income. To do so, we extended the notion of Generalized Lorenz preorder to a context of multivariate distributions with different marginals and showed, by convex analysis, that some conditions, relevant in the analysis of multidimensional inequality, were equivalent to the ordering we introduced.

JEL Classification Numbers: D31, D63, I31.

1. Introduction and overview

The present paper is devoted to the problem of ranking multidimensional distributions in terms of inequality starting from a certain partial preordering that ranks matrices representing the distribution of commodities among people. In particular, we show that some basic results of the theory of majorization for income distributions, namely the well-known Hardy-Littlewood-Pólya (henceforth (HLP)) Theorem on inequality measurement (see HLP [8] and Marshall and Olkin [18]) can be extended to multivariate distributions with different marginals.

1.1. **Relationship to the literature.** Economic literature on inequality measurement is mainly concerned with the comparison of *univariate* indices of well-being, which record differences in the distribution of income (and/or wealth) within and between populations. However, such an approach is considered an inadequate basis for comparing individual disparities because people differ in many aspects besides income. The analysis of different individual attributes is indeed crucial for understanding and evaluating inequality among persons. Thus, a recent research trend focuses on the development of criteria for ranking *multivariate* distributions of individual attributes.[1]

In fact, the current economic literature on multidimensional inequality measurement has followed three different approaches. The first uses some generalized stochastic dominance criteria in order to rank multivariate distributions in a way that is consistent with some classes of social welfare functions (typically [1], [2], [7], and [9]). The second axiomatically generalizes some well-known classes of unidimensional inequality indices to a multidimensional context ([6], [9],

Date: September 18, 2008.

Key words and phrases. multidimensional inequality, generalized Lorenz preordering, price-majorization.

A preliminary version of the present work was presented at the Second Meeting of the Society for the Study of Economic Inequality (ECINEQ) held in Berlin in July 2007 and at the 9th Meeting of the Society of Social Choice and Welfare held in Montreal in June 2008. I thank the participants of both workshops and specially John Weymark for many useful comments. The usual disclaimers apply.

[1]See e.g. Maasoumi [16] and/or Savaglio [22] for a survey.

[15], [26], and [27]). The third applies some tools of convex analysis to compare different distributions of goods within and between populations (see [9], [10], [11], [12], [21], [22]). The first two approaches present several problems as pointed out in Dardanoni [4] and Savaglio [22]. Indeed, multidimensional inequality measurement via dominance conditions[2] introduces sometimes unjustified value-judgements and new complications: discriminatory criteria require preferences' unanimity over classes of evaluation functions which often exhibit unclear requirements on their third and fourth derivatives. At the same time, multidimensional inequality measurement via generalization of well-known classes of inequality indices obscures the level of detail at which comparisons should be undertaken and obtains ambiguous distributional rankings as (among others) Dardanoni [4] pointed out. On the contrary, the third approach has several attractive analytical properties. It is a visual way of studying multivariate distributions, not only making it possible to fully describe mutidimensional disparity, but also to compare distributions with respect to location and dependency. In order to study disparity in a context of more than one (income) variable, we therefore follow this last approach without any further ado.

1.2. **Motivation.** In social choice theory, the question of ordering social states of a population of n anonymous individuals, identical in all respects except for their incomes, generally consists in ranking personal income distributions in terms of inequality. The principal tool for analyzing disparity of income distributions is the Lorenz curve and the related Lorenz dominance criterion,[3] which apply whenever distributions are defined over a fixed population and have identical means. The foregoing assumptions severely restrict the usefulness of this approach in many important practical situations such as international comparisons between countries. Some scholars (see e.g. [18], [24]) have therefore introduced the notion of generalized Lorenz (GL) preordering, which ranks distributions with different means over a fixed population size and solves this puzzling situation. In fact, if we draw two n-person samples from the heterogeneous populations of two countries A and B, then according to GL preordering, the distribution of incomes of A, whose average is μ_A, can be compared with the distribution of incomes of B, whose average is μ_B, with $\mu_A \neq \mu_B$, by simply scaling up the Lorenz curve of both distributions by their respective mean income. GL criterion thus applies to the situation when the marginal distributions under examination differ as would generally be the case with international comparisons or for a given country over a long period. However, GL preordering is unfortunately not sufficient for everyday needs because it was developed for the case in which all non-income attributes are regarded as irrelevant for the purposes of inequality. On the contrary, we know that economists often confront (and are urged to do so) different and welfare-relevant non-income personal characteristics between and within countries in different periods of time. The aim and motivation of the present contribution is then twofold: (a) to provide a rigorous framework for inequality evaluation of heterogeneous distributions, representing populations of individuals characterized by more than one welfare-relevant attribute, and (b) to extend analysis of the standard of living intertemporally or cross-nationally to a multidimensional context. Many empirical illustrations (see e.g. [5], [13], [14], [17], [20]) have indirectly shown that having a GL criterion for multidimensional distributions could be very useful in practice. Indeed, the empirical applications of

[2] See Trannoy [25] for a wide survey of the topic.

[3] See e.g. Marshall and Olkin [18] chapter 1 for a formal definition of the Lorenz curve and the Lorenz criterion.

theoretical multidimensional inequality analysis, including (multidimensional) inequality indices and stochastic dominance techniques, are either information-losing exercises or choices of (sometimes) *ad hoc* criteria. On the contrary, the GL-extension for multivariate distributions and its equivalent conditions, discussed below, is less demanding, because when satisfied, checking it provides detailed information about where the crucial changes in the distribution of income and non-income attributes are.

1.3. **Content.** Here, we analyze inequality in a context of more than one (income) variable, addressing the problem of how to compare populations of individuals endowed with a finite set of attributes distributed in an uneven affluence. To do this, we extend the notion of GL preordering (or dually weak majorization) (see [18], [24]) to a context of multiple individual attributes. In particular, we compare multivariate distributions in terms of inequality when the means of their marginals differ. We represent a multidimensional distribution as a matrix, whose generic entry consists in the quantity of the kth good, $k = 1, ..., m$, allocated to the ith group of individuals, $i = 1, ..., n$. A preorder of different distribution matrices is defined according to their level of disparity and the properties of such a preorder are provided. We compare our preorder of matrix distributions with the main inequality criteria for ranking multidimensional distributions discussed in economic literature (see e.g. [9], [10], [11], [18], [21]). Finally, using certain tools of convex analysis, we show that such a preorder can be replaced by the order defined as the inclusion of the (convex polytope generated by the) columns (and of course rows) of a distribution matrix in the convex hull defined by the set of all convex combinations of the columns (and of course rows) of another distribution matrix and analogously by a (social) evaluation function that records the level of inequality of alternative individual distributions of goods.

The organization of the paper is as follows. In section 2, we discuss our basic notation and definitions, in section 3 presents our results and in the final section some concluding remarks.

2. Notation and definitions

We consider a fixed population of individuals $N = \{1, ..., n\}$, where $n \geq 2$, distinguished according to their level of income as follows: $y_1 \geq y_2 \geq \cdots \geq y_n$, with y_i the income of the ith individual in the distribution $\mathbf{y} \in \mathbb{R}^n$. The concept that the components of a distribution $\mathbf{y}$ are 'more spread out' than the components of a distribution $\mathbf{x}$ has been studied among others by Hardy, Littlewood and Pölya [8], who showed that for any $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y}$ is 'less unequal than' $\mathbf{x}$, denoted

$$\mathbf{y} \preceq \mathbf{x} \tag{2.1}$$

if and only if $\mathbf{y} = P\mathbf{x}$ for some doubly stochastic matrix P,[4] or equivalently if and only if

$$\sum_{i=1}^{n} g(x_i) \geq \sum_{i=1}^{n} g(y_i) \tag{2.2}$$

[4] A doubly stochastic matrix is a square semipositive matrix, with the sum of all components of each row or column equals to one.

holds for all continuous convex functions $g : \mathbb{R} \to \mathbb{R}$. Moreover, equation 2.1 can be combined with a result due to Birkhoff and Von Neumann (see [18], chap. 2, p.19) in order to show that the set

$$\{\mathbf{y} : \mathbf{y} \preceq \mathbf{x}\}$$

is the convex hull of points obtained by permuting the components of $\mathbf{x}$.

As a matter of fact, the crucial assumption for comparison of the two distributions is that $\mathbf{x}$ and $\mathbf{y}$ have identical means. However, considerations on efficiency (higher incomes are more desirable than lower incomes) and international comparisons (GDP in France differs from GDP in Italy) forced researchers to check for conditions that enable us to compare distributions where the total income distributed over different populations can differ. Moreover, scholars have isolated the problem of comparing distributions with different means by the issue of analyzing income inequality among population of different sizes. Indeed, let us suppose $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$ are two vector distributions with means $\overline{x}$ and $\overline{y}$ respectively, where $\overline{x} \neq \overline{y}$, and whose n components are ordered in terms of decreasing incomes, then distribution $\mathbf{x}$ can be considered more unequal than $\mathbf{y}$, denoted as $\mathbf{y} \preceq_{weak} \mathbf{x}$, if and only if

$$\sum_{i=1}^{k} x_i \geq \sum_{i=1}^{k} y_i, \qquad \text{for } k = 1, ..., n. \tag{2.3}$$

Condition 2.3 is referred to as distribution $\mathbf{y}$ is *weakly majorized* by $\mathbf{x}$ (see [18] chapter 1) or dually that $\mathbf{y}$ is dominated by $\mathbf{x}$ according to *generalized Lorenz preordering* (see e.g. [24]).

In what follows, we extend the notion of ordering between vectors in 2.3 to the case of rectangular matrices, namely multidimensional distributions representing a population of n agents among which real-valued attributes are distributed. Indeed, we consider a fixed population of individuals $N = \{1, ..., n\}$ with $n \geq 2$, distinguished by a set $M = \{1, ..., m\}$ of attributes, with $m \geq 2$ in order to avoid trivial qualifications. A distribution matrix, denoted $\mathbf{X} = (x_1, ..., x_m)$, which is a collection of m column vectors, is a matrix, where x_j, for $j = 1, ..., m$, are all column vectors of length n, of the following form:

$$\begin{array}{ccl}
 & \overbrace{\begin{array}{cccc} a & b\ \dots & l\ \dots & m \end{array}} & \longleftarrow \text{attributes} \\
\text{people} \downarrow & & \\
\mathbf{X} = & \left[\begin{array}{cccc} x_{1,a} & x_{1,b} & \cdots & x_{1,m} \\ . & . & & . \\ . & . & x_{i,l} & . \\ . & . & & . \\ . & . & . & . \\ x_{n,a} & \cdots & . & x_{n,m} \end{array}\right] &
\end{array} \tag{2.4}$$

As mentioned above, let element $x_{i,j} \in \mathbf{X}$ be the quantity of the jth real-valued attribute (such as net annual flow of the jth commodity) belonging to the ith individual. The ith row of $\mathbf{X}$ is denoted row_i or $x_{i,\cdot}$, the jth column col_j or $x_{\cdot,j}$, and $\mathcal{A} \subset \mathbb{R}^{n,m}$ is the real vector space of (n, m) matrices with non-negative real entries. A nonnegative square matrix $\mathbf{R}$ (i.e. a matrix such that $x_{i,j} \geq 0$ for every $x_{i,j} \in \mathbf{X}$ and with the same number of rows and columns) with all its row sums equal to 1 is said to be *row-stochastic* or *Markovian* due to its role in the theory of discrete Markov chains. When the sum of all components of each column of a row-stochastic matrix is equal to one, matrix $\mathbf{P}$ is said to be *double stochastic* (see footnote 2). If each row and column of a doubly stochastic matrix has a single unit and all other entries are zero, matrix Π is said to be a *permutation* matrix.[5]

In order to establish when a given distribution $\mathbf{X}$ is more unequal than $\mathbf{Y}$, several notions of matrix-ranking have been introduced in the economic literature (see e.g. [9], [10], [11], [18], [21]). The following criterion compares *multivariate distributions with the same marginals*:

Definition 1. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, *then* $\mathbf{X}$ *is said to majorize* $\mathbf{Y}$, *written* $\mathbf{Y} \prec \mathbf{X}$, *if there exists a* $n \times n$ *doubly stochastic matrix* $\mathbf{P}$, *such that* $\mathbf{PX} = \mathbf{Y}$.

The foregoing definition extends the notion of majorization on integers and the idea of transfer first introduced by Muirhead [19] for the unidimensional case to multivariate distributions. It essentially means that the average is a *smoothing-operation*, which makes the components of $\mathbf{Y}$ more spread out than the components of $\mathbf{X}$.

It is now well established (see e.g. [10], [11], [18]), that $\mathbf{Y} \prec \mathbf{X}$ is tantamount to requiring that the following inequality:

$$\sum_{i=1}^{n} f\left(x_i\right) \geq \sum_{i=1}^{n} f\left(y_i\right) \tag{2.5}$$

hold for any convex function $f : \mathbb{R}^m \to \mathbb{R}$ defined on the real-vector space of the rows of all distribution matrices in $\mathcal{A}$. Finally, it is straightforward to check that Definitions 1 and 2.5 are generalization of 2.1 and 2.2 respectively.

3. Results

3.1. **A new multidimensional inequality criterion and its properties.**

In contrast to the ordering $\prec$ discussed in the previous section, we now introduce a binary relational system that compares *multidimensional distributions of individual attributes with different means* in terms of their relative inequality:

Definition 2. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$ *be two matrices. Then* $\mathbf{Y}$ *is said to be w-majorized by* $\mathbf{X}$, *written* $\mathbf{Y} \ll_w \mathbf{X}$, *if there exists an* $n \times n$ *row-stochastic matrix* $\mathbf{R}$, *such that* $\mathbf{RX} = \mathbf{Y}$.

In words, if we interpret the rows of a matrix as individuals, it is as if we compare a population with a set of attributes (the columns) in terms of inequality at different periods of time, when individual relative affluence is changed. An alternative interpretation considers an international institution that wants to rank different countries in terms of their inequality. The multivariate distributions then represent population samples, with the same size for each nation, to be compared according to a finite set of *essential* attributes such as income, education, health, needs, ability etc. We now illustrate Definition 2 by considering the following:

[5]Notice that the permutation matrices represent the extreme points of the set of doubly stochastic matrices (see [18]).

Example 1. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathbb{R}^{2,3}$ *be two matrices representing a population of two individuals endowed with the same three goods with different affluence and distribution:*

$$\mathbf{X} = \begin{pmatrix} 2 & 5 & 1 \\ 1 & 1 & 7 \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} 1.8 & 4.2 & 2.2 \\ 1.5 & 3 & 4 \end{pmatrix}$$

Since there exists a row-stochastic matrix:

$$\mathbf{R} = \begin{pmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{pmatrix}$$

such that $\mathbf{RX} = \mathbf{Y}$, $\mathbf{Y} \ll_w \mathbf{X}$, *then we conclude that* $\mathbf{X}$ *has a greater level of disparity than* $\mathbf{Y}$ *as each column* j, *for* $j = 1, ...m$, *of the latter is a convex combination of the corresponding column of the former.*

In fact, *w-majorization* amounts to the replacement of each entry of the distribution matrix $\mathbf{Y}$ by averages of the column distribution components of $\mathbf{X}$, namely $\mathbf{Y} \ll_w \mathbf{X}$ if and only if there exists a row-stochastic matrix $\mathbf{R}$ such that $\mathbf{R}x_{\cdot,j} = y_{\cdot,j}$ for any $j = 1, ..., m$.[6] The ordering $\ll_w$ is a *preorder* on set $\mathcal{A}$, i.e. a reflexive and transitive binary relation. Moreover, if we denote with $\mathbf{Z}_S$ the submatrix of $\mathbf{Z}$ induced by the columns indexed by the elements in $S \subseteq \{1, ..., m\}$, then $\mathbf{Y} \ll_w \mathbf{X}$ entails $\mathbf{Y}_S \ll_w \mathbf{X}_S$ for each S. In words, if we fix the marginals of all but s columns, an ordering on $\mathcal{A}$ defines on s-attribute-ordering on the possible column vectors for the s essential characteristics. If for each subset of attributes, these conditional orderings are independent of the values of the columns of the attributes whose marginals have been fixed, then we say that the inequality ordering we consider has a *separability* property of sort. Moreover, if $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$ and $\mathbf{P}$, $\mathbf{Q}$ are two $(n \times n)$ *permutation* matrices, then $\mathbf{Y} \ll_w \mathbf{X}$ entails $\mathbf{PY} \ll_w \mathbf{QX}$. In words, we identify a kind of *anonymity* requirement which prevents the ordering $\ll_w$ from paying attention to the identities of individuals. The foregoing two properties (i.e. separability and anonymity), are provided directly by the definition of $\ll_w$, since the set of row-stochastic matrices is closed under matrix products. Finally, note that both properties allow *w-majorization* to be consistent with an *additive* evaluation of multidimensional inequality (for a discussion on such a topic, see among others [3] and [23]).

3.2. **On the relation between *w-majorization* and some multidimensional inequality criteria.** The ordering $\ll_w$ differs from the ordering $\ll$ studied in [21], which compares matrices with a different number of rows (therefore with different population units), but the same number of columns. In fact, it is easy to prove that the binary relation $\ll_w$ is nested between the matrix majorization $\prec$ and *vp-majorization* $\ll$ analyzed in [21] (i.e. $\prec$ implies $\ll_w$, which implies $\ll$). But $\ll_w$ does not in general imply $\prec$ as it is straightforward to check in Example 1 above. On the contrary, if $\mathbf{X}$ and $\mathbf{Y}$ have the same column marginal distributions then $\ll_w$ implies $\prec$ as shown by the following:

Proposition 1. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$ *and* $\mathbf{P}$ *be a square matrix with non-negative entries and such that* $\mathbf{PX} = \mathbf{Y}$. *If* $\mathbf{Y} \ll_w \mathbf{X}$ *and* $e^t\mathbf{X} = e^t\mathbf{Y}$, *then* $\mathbf{Y} \prec \mathbf{X}$.

[6] Note that *w-majorization* is equivalent to 2.3 for the univariate case.

Proof. It is known ([18] chap. 15 and [10]) that in a binary relational system comparing rectangular matrices $\mathbf{X}$, $\mathbf{Y} \in \mathbb{R}^{n,m}$, if a suitable column vector $\mathbf{e} = (1, 1, ..., 1)$ is added to both matrices under consideration, then their relative ranking is preserved. In particular, if $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, then $\mathbf{Y} \ll_w \mathbf{X}$ if and only if $\mathbf{A} = [\mathbf{Y}, \mathbf{e}] \ll_w [\mathbf{X}, \mathbf{e}] = \mathbf{B}$. Now, suppose $\mathbf{A} \ll_w \mathbf{B}$, then there exists a row-stochastic matrix $\mathbf{P}$ such that $\mathbf{PB} = \mathbf{A}$. Post-multiplying both sides by the so-called Moore-Penrose pseudoinverse $\mathbf{B}^{-1}$ of $\mathbf{B}$, we get $\mathbf{PBB}^{-1} = \mathbf{AB}^{-1}$. But $\mathbf{BB}^{-1} = \mathbf{O}$ is the orthogonal projection matrix onto $\mathbf{B}$, then $\mathbf{AB}^{-1} = \mathbf{PO}$, that has non-negative entries by assumption. In order to obtain the required result, we need to show that $\mathbf{PO}$ (and of course $\mathbf{AB}^{-1}$) is doubly stochastic, namely that $\mathbf{POe} = \mathbf{e}$ and $\mathbf{e}^t\mathbf{PO} = \mathbf{e}^t$. Since $\mathbf{Oe} = \mathbf{e}$, then $\mathbf{Pe}$, but $\mathbf{Pe} = \mathbf{e}$, because $\mathbf{P}$ is row-stochastic, hence $\mathbf{POe} = \mathbf{e}$. On the other hand, $\mathbf{e}^t\mathbf{PO} = \mathbf{e}^t\mathbf{PBB}^{-1} = \mathbf{e}^t\mathbf{AB}^{-1}$, but $e^t\mathbf{A} = e^t\mathbf{B}$ by assumption, thus $\mathbf{e}^t\mathbf{BB}^{-1} = \mathbf{e}^t\mathbf{O} = \mathbf{e}^t$. Therefore $\mathbf{PO} = \mathbf{AB}^{-1}$ is doubly stochastic. Recalling that $\mathbf{A} \ll_w \mathbf{B}$ is equivalent to $\mathbf{Y} \ll_w \mathbf{X}$, we obtain the expected result. □

Koshevoy [10] recently introduced standard tools of convex analysis into the theory of inequality measurement. He defines the multivariate extension of the Lorenz preordering as being induced by inclusion among the zonotopes of multivariate distributions.[7] More in general, if

$$\mathbb{H} = co\left\{(\mathbf{z}_{i,1}, ..., \mathbf{z}_{i,m}),\ i = 1, ..., n\right\} \tag{3.1}$$

denotes the convex hull of a matrix $\mathbf{Z} \in \mathbb{R}^{n,m}$, namely the set of convex combinations of the row-vectors of $\mathbf{Z}$, then, given two matrices $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{n,m}$, we say that $\mathbf{Y}$ has lower inequality than $\mathbf{X}$ if it lies in the convex hull of all column (and of course all row) permutations of $\mathbf{X}$ and we write:

$$\mathbf{Y} \subseteq co\,(\mathbf{X})\,. \tag{3.2}$$

In words, the idea is to consider each multidimensional distribution as the boundary of a proper chosen set and order these sets by inclusion. Since there is no natural complete order of the m-dimensional space when $m > 1$, in this way it is possible to avoid the questionable ordering of vectors typical of the multidimensional inequality indices and adopt a more elegant and intuitive geometric approach to the analysis of disparity when individuals differ in several attributes.[8] In particular, it holds true that:

Proposition 2. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, *then* $\mathbf{Y} \ll_w \mathbf{X}$ *if and only if* $\mathbf{Y} \subseteq co\,(\mathbf{X})$.

Proof. ($\Rightarrow$) Take $\mathbf{Y} \ll_w \mathbf{X}$, then $\mathbf{PX} = \mathbf{Y}$ which is tantamount to:

$$\mathrm{row}_i(\mathbf{Y}) = [y_{i,1}, ..., y_{i,m}] = \sum_{k=1}^{m} p_{i,k}\,\mathrm{col}_k\,(\mathbf{X}) \quad \text{for } i = 1, ..., n,$$

i.e. $\mathrm{row}\,(\mathbf{Y}) \in co\,(\mathrm{col}\,(\mathbf{X}))$ as required.

($\Leftarrow$) Assume $\mathbf{Y} \subseteq co\,(\mathbf{X})$, that is equivalent to $\mathrm{row}_i\,(\mathbf{Y}) = \sum_{i,k} p_{i,k}\,\mathrm{col}_{i,k}$, for $i = 1, ..., n$, i.e. a system of n linear equation in m variables (with $n \geq m$), a solution of which is a matrix $\mathbf{P}$ (with the constraint that $\sum_k \mathbf{p}_{i,k} = 1$), such that $\mathbf{PX} = \mathbf{Y}$, namely $\mathbf{Y} \ll_w \mathbf{X}$. □

[7] Note that a zonotope generated by the column vectors of a matrix distribution is the finite Minkowski sum of line segments in $\mathbb{R}^m$.

[8] See Weymark [29] for a survey of the approach which evaluates multivariate disparity according to alternative classes of multidimensional inequality indices and Dardanoni [4] and Savaglio [22] for a criticism of that approach.

The foregoing result introduces a new characterization of multidimensional disparity. The set inclusion of convex hulls defines an order which reflects the dispersion of column vectors. The convex hull of a matrix representing the distribution of m real-valued attributes of a finite population also fully describes the underlying distribution and therefore no information is lost, unlike in the case of multidimensional inequality indices[9]. Finally, as the set of convex combinations of column-vectors of the distribution matrix has a graphic representation, the inclusion of the convex polytope associated with multidimensional distributions indicates the degree of inequality and is visual, as shown by the following prominent:

Example 2. *Let $X, Y \in R_+^{3,2}$ be two matrices of a three-individual population endowed with two goods, but with different affluence:*

$$\mathbf{X} = \begin{pmatrix} 1 & 1 \\ 3 & 5 \\ 18 & 20 \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} 6.5 & 7.5 \\ 8.1 & 9.4 \\ 8.8 & 10.6 \end{pmatrix}.$$

The total amount of both attributes in $\mathbf{X}$ is almost all owned by the third individual, while it is spread throughout the population of $\mathbf{Y}$. Since there exists a row-stochastic matrix:

$$\mathbf{P} = \begin{pmatrix} 0.5 & 0.2 & 0.3 \\ 0.4 & 0.2 & 0.4 \\ 0.1 & 0.5 & 0.4 \end{pmatrix},$$

such that $\mathbf{PX} = \mathbf{Y}$, then we conclude that the columns of $\mathbf{Y}$ relay to the set of all convex combinations of the row vectors of $\mathbf{X}$ (see Figure 1.(a)):

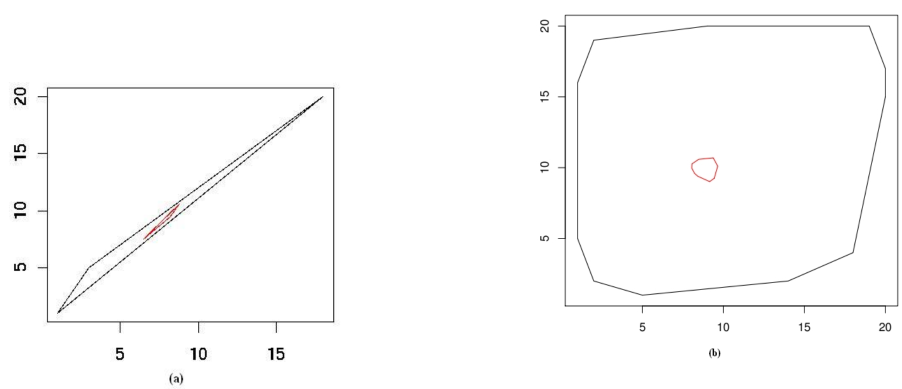

Figure 1

In order to make inclusion of two convex hulls generated by the preordering induced by w-majorization more graphically evident, figure 1.(b) depicts two matrices representing two populations of one-hundred individuals respectively who differ in only two welfare-relevant attributes.

Another important notion of matrix majorization (originally proposed by Kolm [9]), is that of *price majorization*, namely a matrix $\mathbf{Y}$ is said to be *price-majorized* by $\mathbf{X}$, denoted as $\mathbf{Y} \prec_p \mathbf{X}$,

[9] See Savaglio [22] for a critical discussion.

if $\mathbf{Yp} \prec \mathbf{Xp}$ for all $\mathbf{p} \in \mathbb{R}^m$, where $\mathbf{Zp} = ((\mathbf{z}_{1,\cdot}, \mathbf{p}), ..., (\mathbf{z}_{n,\cdot}, \mathbf{p}))$ for $\mathbf{Z} \in \mathbb{R}^{n,m}$ and $\prec$ is interpreted in the sense of (standard) majorization (see [18], chap.1) between vector distributions. In words, this means that if we post-multiply the endowment of each individual by the same coefficients (weights), the disparity between two populations does not chance. The notion of price majorization is useful for comparing nonmonetary quantities such as individual talents, health, education, etc. It tests whether a given distribution of individual resources is less spread out than another, irrespective of the values assigned to those resources. Unfortunately, individual attributes are thus reduced to monetary quantities, losing information. As a consequence of the previous proposition, we now establish the relation between $\prec_p$ and $\ll_w$:

Corollary 1. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, $\mathbf{Y}$ *price-majorized by* $\mathbf{X}$ *entail* $\mathbf{Y}$ *w-majorized by* $\mathbf{X}$.

Proof. Suppose $\mathbf{X}$ is price-majorized by $\mathbf{Y}$ but that $\mathbf{Y}$ is not w-majorized by $\mathbf{X}$. This implies that there is at least a $\mathrm{row}_i(\mathbf{Y}) \notin co(\mathrm{col}(\mathbf{X}))$. Therefore, there exists a hyperplane such that: $\langle \mathrm{row}_i(\mathbf{Y}), \mathbf{p}\rangle \geq t$ and $\langle \mathrm{row}_j(\mathbf{X}), \mathbf{p}\rangle < t$, for all $j = 1, ..., n$, $\mathbf{p} \in \mathbb{R}^m$ and $t > 0$, and such that $\langle\cdot\rangle$ is the inner product of two vectors. However, by definition of price-majorization:

$$y_1 = \mathrm{row}_1(\mathbf{Y})\,\mathbf{p} \geq y_i = \mathrm{row}_i(\mathbf{Y})\,\mathbf{p} \geq t > \mathrm{row}_1(\mathbf{X})\,\mathbf{p} = x_1$$

which is a contradiction and therefore the required result. □

Since price-majorization reduces the difficult problem of comparing multidimensional distributions to the easier task of testing (standard) majorization between distribution vectors, the former is an efficient method of checking when the preordering induced by the relation $\ll_w$ holds. However, in general, $\ll_w$ does not imply $\prec_p$, as it is possible to verify by considering $\mathbf{p} = (2, 1, 0)$ in Example 1. Finally, since for any $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, $\mathbf{Y} \prec_p \mathbf{X}$ is equivalent to requiring the weak set-inclusion of the Lorenz zonotopes [10] (see [10] Theorem 1) of the foregoing distribution matrices, then the multivariate generalization of the Lorenz criterion, introduced by Koshevoy [10], induces a ranking on $\mathcal{A}$ which is a subordering of the w-majorization preordering.

3.3. **A normative approach to multidimensional inequality measurement.** We know that w- majorization only provides a preordering of the possible distributions of attributes. Following the traditional approach to unidimensional inequality measurement in the economic literature, we introduce the class of functions preserving the preordering[11] $\ll_w$ into the present multidimensional framework. Since *w-majorization* represents a subordering of the ranking induced by vp-majorization $\ll$ studied in [21] (i.e. $\ll_w$ implies $\ll$), then the class of functions preserving $\ll_w$ must be larger than that related to $\ll$ (see [21], Theorem 3). Indeed, if we define a set $\Upsilon(\mathbb{R}^m)$ as the set of all real-valued convex functions defined over a nonempty compact convex set $C \subset \mathbb{R}^{k,m}$, we establish a counterpart of result 2.2:

Proposition 3. *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathcal{A}$, *then the following statements are equivalent:*

(*i*) $\mathbf{Y} \ll_w \mathbf{X}$;

[10] A (Lorenz) zonotope of a (n, d)-matrix is a center symmetric convex polyhedron in $\mathbb{R}^{d+1}_+$, namely a polytope which can be obtained as the Minkowski sum of finitely many closed line segments in $\mathbb{R}^d$ connecting the origin to the endpoint of each vector (see [30] for a wide survey)

[11] A function preserving the ranking induced by a given binary relation is referred to as order-preserving, isotonic or convex in the sense of Schur (see [18])

(*ii*) *For each integer* $1 \leq k \leq n$ *and* $\gamma \in \Upsilon_k(\mathbb{R}^m)$,

$$\max_k \gamma(x_k) \geq \max_k \gamma(y_k) \tag{3.3}$$

Proof. ($\Rightarrow$) According to Proposition 2, *w-majorization* is equivalent to $\operatorname{row}(\mathbf{Y}) \subseteq co(\operatorname{col}(\mathbf{X}))$. Now, $\operatorname{row}(\mathbf{Y})$ is convex, because any $\operatorname{row}_i(\mathbf{Y}) = \sum_{k=1}^{m} r_{i,k} \operatorname{col}_k(\mathbf{X})$, where $r_{i,j}$ are the entries of some row-stochastic matrix $\mathbf{R}$ such that $\mathbf{Y} \ll_w \mathbf{X}$, and $co(\operatorname{col}(\mathbf{X}))$ is also convex by definition. From [28], we know that given two convex sets A and B, $A \subseteq B$ if and only if $\max_{z \in A} \phi(z) \leq \max_{z \in B} \phi(z)$ for every convex function defined over $A \cup B \subseteq \mathbb{R}^{k,m}$ and $z \in \mathbb{R}^m$, which is the required result.

($\Leftarrow$) Let us assume that $\max_k \gamma(x_k) \geq \max_k \gamma(y_k)$ for every real-valued convex function $\gamma \in \Upsilon_k(\mathbb{R}^m)$. This is equivalent to saying that the set of all rows of $\mathbf{Y}$ is convex and a subset of the convex set of all columns of $\mathbf{X}$, i.e. that $\operatorname{row}(\mathbf{Y}) \subseteq co(\operatorname{col}(\mathbf{Y}))$, or $\mathbf{Y} \ll_w \mathbf{X}$. $\square$

$\max \gamma(x)$ can be interpreted as a function that evaluates the relative disparity of each person in a society, with convexity of γ that dually captures individual inequality aversion. Moreover, according to the normative approach to inequality measurement, it could be considered the dual of a corresponding (multidimensional) inequality index and/or used to construct one. Since social evaluation ranks alternative distributions of attributes according to their social desirability, it could be worth investigating the value judgements underlying the expression in 3.3, in order to axiomatically build a new class of multidimensional indices of disparity. However, this task is best left to further research.

4. Final remarks

In exploring the possibility of theoretically extending inequality analysis from the univariate to the multivariate setting, our main aim was to provide a multidimensional counterpart of the generalized Lorenz preordering for the case in which people differ in many attributes besides income.

We know that the approach that analyzes multidimensional inequality by measure indices (see [29]) is problematical because it requires judgements on the relative importance of the various individual attributes, the degree of substitution between them and the degree of inequality aversion in society. At the same time, the approach to multidimensional inequality via stochastic dominance (see [25]) introduces further layers of complexity to the measurement of disparity in several dimensions as pointed out in [22], drawbacks that need to be overcome in order to use multidimensional inequality analysis in comparative studies.

On the contrary, using simple tools of convex analysis, the present study shows how our approach can be of some interest for the analysis of inequality when several (individual) characteristics are considered simultaneously and offers significant scope for further developments in the study of multidimensional inequality.

All in all, relatively little progress has been made in extending the theory of inequality measurement from univariate to the multivariate case. Studies on multidimensional disparity comparisons are very few, the problem is really complex and much work remains to be done.

References

[1] Atkinson A. B. and F. Bourguignon (1982) - The comparison of multidimensional distributions of economic status - *Review of Economic Studies* 49, 181-201.

[2] Atkinson A. B. and F. Bourguignon (1987) - Income distribution and differences in needs - in G.R. Feiwel (ed.), *Arrow and the Foundation of the Theory of Economic Policy*, Macmillian, London.

[3] Bossert, W. and J.A. Weymark (1996) - Utility in social choice. In: Barberà, S., P.J. Hammond, C. Seidl (eds) - *Handbook of utility Theory*, vol 2. Extensions. Boston: Kluwer.

[4] Dardanoni, V. (1995) - On multidimensional inequality measurement - In: Dagum, C. and A. Lemmi (Eds.), *Income Distribution, Social Welfare, Inequality,and Poverty.* Vol. 6 of *Research on Economic Inequality.* JAI Press, Stamford, CT, 201–207.

[5] Decancq, K., and M.A. Lugo (2008) - Measuring inequality in well-being: a proposal based on the multidimensional Gini index - mimeo

[6] Gajdos, T. and J. A. Weymark (2005) - Multidimensional generalized Gini indices - *Economic Theory* 26, 471-496.

[7] Gravel, N. and P. Moyes (2006) - Ethically robust comparisons of distributions of two individual attributes - IDEP working paper no 06.04.

[8] Hardy, G. H., Littlewood, J.E., Polya, G. (1934/1952) - *Inequalities* - London: Cambridge University Press.

[9] Kolm, S.C. (1977) - Multidimensional egalitarianisms - *Quarterly Journal of Economics* 91, 1-13.

[10] Koshevoy, G. (1995) - Multivariate Lorenz majorization - *Social Choice and Welfare* 12, 93-102.

[11] Koshevoy, G. (1998) - The Lorenz zonotope and multivariate majorizations - *Social Choice and Welfare* 15, 1-14.

[12] Koshevoy and Mosler (1996) - The Lorenz zonoid of a multivariate distribution - *Journal of American Statistical Association* 91, 873-882.

[13] Lugo, M.A. (2007) - Comparing multidimensional indices of inequality: methods and application - *Research on Economic Inequality* 14, 213-236.

[14] Lugo, M.A. and E. Maasoumi (2008) - Multidimensional poverty measures from an information theory perspective - ECINEQ WP 2008 - 85.

[15] Maasoumi, E. (1986) - The measurement and decomposition of multidimensional inequality - *Econometrica* 54, 991-997.

[16] Maasoumi, E. (1999) - Multidimensioned approaches to welfare analysis - In J. Silber (ed.) *Handbook of Income Inequality Measurement.* Boston, MA: Kluwer Academic Publishers, pp. 437-477.

[17] Maasoumi, E., and G. Nickelsburg (1988) - Multidimensional measures of well-being and an analysis of inequality in the Michigan data - *Journal of Business and Economic Statistics* 6, 327–334.

[18] Marshall A. W. and I. Olkin (1979) - *Inequalities: Theory of Majorization and its Applications* - Academic Press.

[19] Muirhead, R.F. (1903) - Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters - Proceeding Edimburgh Mathematical Society 21, 144-157.

[20] Quadrado L., W. Heijman and H. Folmer (2001) - Multidimensional Analysis of Regional Inequality: The case of Hungary - *Social Indicators Research* 56, 21-42.

[21] Savaglio, E. (2006) - On multidimensional inequality with variable population size - *Economic Theory* 28, 85-94.

[22] Savaglio, E. (2006) - Three approaches to the analysis of multidimensional inequality - in Farina, F. and E. Savaglio (eds), *Inequality and Economic Integration*, Routledge, London.

[23] Sen, A.K. (1997) - On economic inequality, 2nd edn. Extended edition with J. Foster. Oxford: Clarendon Press.

[24] Shorrocks, A. F. (1983) - Ranking income distributions - *Economica*, 50, 3-17.

[25] Trannoy, A. (2006) - Multidimensional egalitarianism and the dominance approach: a lost paradise? - in Farina, F. and E. Savaglio (eds), *Inequality and Economic Integration*, Routledge, London.

[26] Tsui, K. Y. (1995) Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Sen-Kolm approach - *Journal of Economic Theory* 67, 251-265.

[27] Tsui, K. Y. (1999) - Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation - *Social Choice and Welfare* 16, 145-157.

[28] Webster, R. (1994) - *Convexity* - Oxford University Press.

[29] Weymark, J. (2006) - The normative approach to the measurement of multidimensional inequality - in Farina, F. and E. Savaglio (eds), *Inequality and Economic Integration*, Routledge, London.

[30] Ziegler, G. (1997) - *Lectures on polytopes* - Springer-Verlag

DMQTE, University of Chieti-Pescara -viale Pindaro 42, 65127 - Pescara (I) and DEP, University of Siena - piazza S. Francesco 7, 53100 - Siena (I)

E-mail address: ernesto@unich.it