::ON MULTIDIMENSIONAL INEQUALITY WITH VARIABLE HOUSEHOLD WEIGHT

ERNESTO SAVAGLIO

52 RIUNIONE SIE

TEMA: PIL, benessere e felicità, Povertà e distribuzione del reddito
PRELIMINARY DRAFT

Abstract. The present work establishes a suitable multidimensional extension of a fundamental theorem of inequality measurement, namely the characterization of the Lorenz (or dually majorization) preorder of real-valued univariate (income) distributions, via a class of reasonable welfare functions (or dually inequality indices) due to Hardy, Littlewood and Polÿa (henceforth HLP, (1934)).

JEL Classification: D31, D63, I31

1. Introduction

TO BE WRITTEN

Theoretical analysis of inequality compares discrete distributions of individual (or household) incomes. Such a 'univariate approach' is now widely considered to be inadequate because it does not take into account that people differ in many aspects (as. e.g. gender, life expectancy, needs, etc) besides income and therefore individual disparities can arise in more than one dimension. Moreover, the standard classes of unidimensional inequality indices do not provide sufficient information on individual deprivation because, taking income as the unique explanatory variable, they neglect fundamental problems as e.g. individual lack of access to health care or to education.

However, the problem of extending the theory of inequality measurement from the unidimensional to the multidimensional setting is a quite complex and unexplored research field.[1] Several orderings have been used to compare multidimensional distributions in terms of inequality, but there is not yet a criterion that "universally" recognizes a redistribution of resources as more equitable than another one.

In the present work, we address the issue of assessing multidimensional inequality by introducing a new ordering that compares (discrete) multivariate distributions representing households that differ in several characteristics besides income and that could have different size and weights.

The relevance of the problem arises from the following considerations.

The Lorenz curve and the related Lorenz dominance criterion [2] are the principal tools for ranking income distributions in terms of inequality. They apply whenever distributions are defined over a

Key words and phrases. Multidimensional Inequality, Socila Welafre Functions, Equivalence Scales, Lorenz Criterion, Hardy-Littlewood-Polya Theorem.

[1] For an almost complete analysis of the economic literature on multidimensional inequality see the three surveys of Savaglio, Trannoy and Weymark in [4].

[2] See e.g. Marshall and Olkin [9] chapter 1 for a formal definition of the Lorenz curve and the Lorenz criterion.

fixed population and have identical means. These assumptions severely restrict the usefulness of this approach in many important practical situations.[....]

The present work establishes a suitable multidimensional extension of a fundamental theorem of inequality measurement, namely the characterization of the Lorenz (or dually majorization) preorder of real-valued univariate (income) distributions, via a class of reasonable welfare functions (or dually inequality indices) due to Hardy, Littlewood and Polÿa (henceforth HLP, (1934)). In particular, we compare multidimensional distributions in terms of inequality starting from a certain partial preordering that ranks matrices representing the distribution of commodities among households with different weights.

Our analogue to the classic HLP theorem consists in proving that a version of the result due to Schur and Ostrowski on the class of majorization order-preserving functions (see Marshall and Olkin (1979) ch. 1) also holds in our setting of distributions of individual goods. In particular, we focus on the class of real-valued functions which preserve the majorization preorder as defined on the space of height vectors. The importance of defining an order-preserving function in such a setting relies on the one-to-one correspondence between isotone functions, maps that preserve the majorization relation, and the so-called social evaluation functions (SEF)). SEFs are widely used to define inequality indices, which in turn provide the basis for welfare comparisons between and within populations by equity-concerned policy-makers.

Motivation. When non-income attributes are regarded as relevant for the purposes of inequality and the observations of common data sets usually are weighted. Indeed, we know that economists often confront (and are urged to do so) different and welfare-relevant non-income personal characteristics between and within countries. Therefore, it seems natural to investigate criteria for ranking multivariate distributions with different population, where multiple attributes of well-being have to be compared simultaneously. Moreover, scholars typically analyze grouped data where the frequency denotes the weights of the groups of households in a population. More recently, theoretical and empirical works on the equivalence scales have emphasized the importance of weights for distributive analysis in the case of heterogeneous households.[3]

2. Preliminaries

Let $N = \{1, ..., n\}$ with $n \geq 2$ be a fixed set of individuals and $\aleph = \left\{ z \in \mathbb{R}^n : \sum_{i=1}^{n} z_i = T \right\}$ be the set of all real-valued distribution vectors of a total amount of income T. An element $z_i \in z$ with $i = 1, ..., n$ represents the income received by the i-th individual. We denote with $z_\downarrow = \left(z_{[1]}, ..., z_{[n]}\right)$ the vector whose components are rearranged from the richest to the poorest. Then, the most applied criterion to compare an income distribution x with another one y in terms of inequality is the Lorenz (or dually majorization) preorder (LP) that establishes:

Definition 1. *For any x,$y \in \aleph$, x is a more even distribution of income than y, denoted to as $x \prec_{LP} y$, if:*

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]} \quad k = 1, ...n - 1. \tag{2.1}$$

[3] See e.g. Ebert (????)

In a celebrated result of their joint effort, Hardy, Littlewood and Polya (see Marshall and Olkin [9] ch.1) have proved the so-called fundamental result of the (unidimensional) inequality measurement theory, which states:

Theorem 1 (HLP, (1929))**.** *For any x, $y \in \aleph$, the following two conditions are equivalent:*

(1) *$x \prec_{LP} y$ if $x = By$ for some $n \times n$ bistochastic matrix B;*[4]

(2) *the inequality*

$$\sum_{i=1}^{n} f(x_i) \leq \sum_{i=1}^{n} f(y_i)$$

holds for any $f : \mathbb{R} \rightarrow \mathbb{R}$ that is convex and symmetric, that is invariant under any permutation of the coordinates.[5]

In words, Theorem 1 says that comparing two income distributions x and y according to the Lorenz preorder is equivalent to the fact that distribution x is obtained from y through a finite number of income transfers that take place from the richer to the poorer[6] or it is tantamount to saying that there exists an additive class of (convex) evaluation functions consistent with (i.e. preserving) the ordering induced by the Lorenz criterion.

In what follows, our aim is to extend the foregoing result to the more general case in which people could differ with respect to *several characteristics besides income* and the households of a *populations of variable size* could have *different weights*. In order to pursue our aim, for $k \in \mathbb{N}$ and for each $n \in \mathbb{N}$, we define with:

$$\Gamma_n^k = \left\{ (\mathbf{x}_1, ..., \mathbf{x}_n) : \mathbf{x}_i \in \mathbb{R}^k, 1 \leq i \leq n \right\}$$

the real vector space of (n, k) matrices with real entries, where the generic $\mathbf{x}_i$ is a row vector of length k. We could interprete an element $x_{i,j}$ of a multivariate ditribution $X \in \Gamma_n^k$ as the quantity of the jth (real-valued) attribute (such as the net annual flow of the jth commodity) belonging to the ith individual. The ith row of X is denoted row_i or $\mathbf{x}_{i,\cdot}$, the jth column col_j or $\mathbf{x}_{\cdot,j}$. Given $X = (\mathbf{x}_1, ..., \mathbf{x}_n) \in \Gamma_n^k$, we denote with X^T its transpose and with $\mathbb{M}^{m,n}$ the set of all *row-stochastic* $m \times n$ matrices R, namely nonnegative rectangular matrices with all its row sums equal to one.[7] Finally, we define:

$$\Xi_n = \left\{ (a_1, ..., a_n) : a_i \in (0, 1) \,,\, 1 \leq i \leq n,\ \sum_{i=1}^{n} a_i = 1 \right\}.$$

the set of all possible weight systems and with $\Lambda_n^k = \Gamma_n^k \otimes \Xi_n$ the space of all possible pairs of distribution matrices and (households') weights.

3. Results

3.1. **A new multidimensional inequality criterion.**

We now introduce a binary relational system to compare (in terms of relative inequality) *multidimensional distributions* (of individual/attributes)

[4]A bistochastic matrix B is a square matrix where all entries are non-negative and the sum of all rows and columns are equal to one.

[5]See Marshall and Olkin [9] for a wide review of the present proposition.

[6]Notice that a bistochastic matrix B just collects such a (finite) sequence of income transfers (see Marshall and Olkin [9] chapters 1 and 2).

[7]Any row-stochastic matrix whose sum of all components of each column is equal to one is of course a *bistochastic* matrix (see footnote 4).

in the most general case in which a *population* could *differ in size* and its members (individuals, households, groups) (could) have *different weights*. Therefore, we state:

Definition 2. *Let $(X, \mathbf{a}) \in \Lambda_m^k$ and $(Y, \mathbf{b}) \in \Lambda_n^k$. Then, we say that $(X, \mathbf{a})$ is less unequal than $(Y, \mathbf{b})$, denoted to as $(X, \mathbf{a}) \prec_h (Y, \mathbf{b})$, if there exists a matrix $R \in \mathbb{M}^{m,n}$ such that*

$$X^T = RY^T \quad \textit{and} \quad \mathbf{b} = \mathbf{a}R.$$

In words, if we interpret the rows of a matrix as households with different endowments, then it is as if we would rank different countries (of course with different population size) in terms of inequality by also considering the fact the households are not homogeneus within and between countries. Indeed, households differ not only with respect to their income, but also in size, number of adults, number and age of children, the health status of family members, the place where they live and other important socio-economical characteristics. Since these different household characteristics are observable, they are typically employed to assign a value reflecting the needs and the standard of living of the respective households. In most empirical situations, scholars in fact take into account of such diversity among households by using equivalence scales, namely factors expressing how the necessities of a household grow with each additional member. The preordering $\prec_h$ proposed here could then be considered as a theoretical justification to the use of equivalence scales in the general case in which households differ for several characteristics besides income and needs.

According to Definition 2, the cost of comparability of two multivariate distributions in the present general setting is that two contraints must be satisfied at a time: the elements of both the more even distribution matrix and of its weight vector has to be a convex combination of both the elements of the more unequal distribution matrix and of its weight vector. The benefits of our criterion is that it allows for possibly different weighting schemes and impose a minimalistic structure on the equality-enhancing transformations to be used in order to make a multivariate distribution "smoother". Moreover, preordering $\prec_h$ represents a theoreticl guide to a policy maker who would like to ajust bundles of households' characteristics (or attributes) by means of a scaling factor and then comparing adjusted multidimensional distributions by means of a suitable inequality criterion.[8]

We now illustrate Definition 2 by considering the following:

Example 1. *Let $(X, \mathbf{a}) \in \Lambda_3^2$ and $(Y, \mathbf{b}) \in \Lambda_2^2$ be two matrices representing a population of three and two individuals respectively endowed with the same two goods with different affluence and two weights' systems representing for example two equivalence scales of two different States:*

$$Y^T = \begin{bmatrix} 3 & 1 \\ 1 & 4 \end{bmatrix} \quad \textit{and} \quad X^T = \begin{bmatrix} 2 & 2.5 \\ 1.7 & 3 \\ 1.8 & 2.8 \end{bmatrix}$$

$$\mathbf{a} = (0.2,\ 0.3,\ 0.5) \ \textit{and}\ \mathbf{b} = (0.4,\ 0.6)\,.$$

Since there exists a row-stochastic matrix:

$$R = \begin{bmatrix} 1/2 & 1/2 \\ 1/3 & 2/3 \\ 2/5 & 3/5 \end{bmatrix}$$

[8]In a unidimensional context, see Blackorby and Donaldson (1993) for for a wide analysis of adult-equivalence scales and their potential implications for applied welfare economics.

such that $X^T = RY^T$ and $\mathbf{b} = \mathbf{a}R$, $(X, \mathbf{a}) \prec_h (Y, \mathbf{b})$ then we conclude that Y has a greater level of disparity than X as each column j, for $j = 1, 2$, of the latter is a convex combination of the corresponding column of the former.[9]

Example 2. *Consider as a suitable example of weight vectors $a = (0.281, 0.273, 0.208, 0.178, 0.059)$ representing the percentage of Italian households with one, two,..., five or more components respectively as reported by ISTAT for the year 2009 and $b = (28.8, 58.9, 12.3)$ representing the percentage of single adult, monogamous and polygamous households respectively as reported by de Laiglesia and Morrisson (2008) for Cote d'Ivoire. Then, if a Statistical Boureau want to compare Italy with Cote d'Ivoire in terms of multidimensional (i.e. with respect to several monetary and non-monetary indicators) inequality taking into account that the family structure of such African country differs from Italian one could eventualy use the criterion we propose.*

3.2. **A representation theorem on weights and measures of the multivariate distributions.** We know that h-criterion only provides a preordering of the possible distributions of attributes. Then, following the traditional approach to unidimensional inequality measurement in the economic literature, we introduce the class of functions preserving[10] the preordering $\prec_h$ into the present multidimensional framework by stating the following generalization of the Theorem 1 above:

Theorem 2. *Let $(X, \mathbf{a}) \in \Lambda_m^k$ and $(Y, \mathbf{b}) \in \Lambda_n^k$. The following conditions are equivalent:*

(i)**:** $(X, \mathbf{a}) \prec_h (Y, \mathbf{b})$;

(ii)**:** *For any continuous convex function $\phi : \mathbb{R}^k \to \mathbb{R}$*

$$\sum_{i=1}^{m} a_i \phi(\mathbf{x}_i) \leq \sum_{j=1}^{n} b_i \phi(\mathbf{y}_j).$$

Proof. The implication $(i) \Rightarrow (ii)$ is an immediate consequence of Jensen's inequality, i.e. that a 'convex transformation of a mean is less or equal than the mean after a convex transformation'.

Then, assume (ii) holds and define the following set:

$$\mathcal{M}(X, Y) = \left\{ A = (a_{ij}) \in \mathbb{M}^{m,n} \mid \mathbf{x}_i = \sum_{j=1}^{n} a_{ij} \mathbf{y}_j,\ 1 \leq i \leq m \right\}.$$

We first show that $\mathcal{M}(X, Y)$ is a non-empty closed convex subset of $\mathbb{M}^{m,n}$. Now, that $\mathcal{M}(X, Y)$ is closed and convex is obvious. To prove that $\mathcal{M}(X, Y) \neq \varnothing$, it is sufficient to show that if Theorem 2.(ii) holds, then $\mathbf{x}_i = conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ for any $i = \{1, ..., m\}$, where $conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ represents the convex hull of vectors $\mathbf{y}_1, ..., \mathbf{y}_n$. So doing, suppose that $\mathbf{x}_j \notin conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ for some $j \in \{1, ..., m\}$ and then consider the function $\phi : \mathbb{R}^k \to \mathbb{R}$ defined as $\phi(\mathbf{t}) = \delta(\mathbf{t}, conv(\mathbf{y}_1, ..., \mathbf{y}_n))$, where δ is the Euclidean distance in $\mathbb{R}^k$ and note that $\phi(\mathbf{x}_j) > 0$, $\phi(\mathbf{y}_l) = 0$ for $1 \leq l \leq n$, and $\phi(\mathbf{t})$ is a convex function since $conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ is a non-empty convex subset of $\mathbb{R}^k$. Thus, it follows that $\sum_{i=1}^{m} a_i \phi(\mathbf{x}_i) >$

[9]It is worth reminding here that in a multidimensional context the effect of an equality-enhancing transformation could be ambiguous. Indeed, the rearrangement of the components of a multivariate distribution by means of a row-stochastic matrix R might sometimes produce an uneven allocation of attributes (see e.g. Dardanoni [2] Savaglio in [4] p. 279), because the obtained multidimensional distribution could be smoother in the tails and more polarized in the middle with an uncertain total effect of the inequality-decreasing operator R on the whole distribution.

[10]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 Marshall and Olkin [9] ch. 2)

$\sum_{j=1}^{n} b_i \phi(\mathbf{y}_j)$ which contradicts condition (ii) in Theorem 2. Hence, $\mathbf{x}_i = conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ for any $i = \{1, ..., m\}$, as required.

Given $M = (m_{i,j}) \in \mathcal{M}(X,Y)$, we now define a vector $\mathbf{s}(M) = (s_1(M), ..., s_n(M)) \in \mathbb{R}^n$ by $s_j(M) = \sum_{i=1}^{m} m_{i,j} a_i$, $1 \le j \le n$. Set

$$\mathcal{F}(X,Y) = \{\mathbf{s}(M) : M \in \mathcal{M}(X,Y)\}.$$

Note that $\alpha \mathbf{s}(M) + (1-\alpha)\mathbf{s}(N) = \mathbf{s}(\alpha M + (1-\alpha)N)$ for any $M, N \in \mathcal{M}(X,Y)$ and $\alpha \in [0,1]$, so that $\mathcal{F}(X,Y)$ is a convex subset of $\mathbb{R}^n$. Moreover, since $\mathcal{M}(X,Y)$ is closed the same must be true for $\mathcal{F}(X,Y)$. We finally need to show that if $\mathbf{b} \in \mathcal{F}(X,Y)$, then there exists a matrix $M_0 \in \mathcal{M}(X,Y)$ such that $\mathbf{s}(M_0) = \mathbf{b}$. In order to show that, assume that $\mathbf{b} \notin \mathcal{F}(X,Y)$. Since $\mathcal{F}(X,Y)$ is a closed convex set, it follows from the Hahn-Banach theorem (see [10]) that there exists a $c \in \mathbb{R}$ and $(t_1, ..., t_n) \in \mathbb{R}^n$ such that:

$$\sum_{j=1}^{n} t_j s_j < c < \sum_{j=1}^{n} t_j b_j \text{ for } (s_1, ..., s_n) \in \mathcal{F}(X,Y). \tag{3.1}$$

Given an arbitrary vector $\mathbf{y} \in conv(\mathbf{y}_1, ..., \mathbf{y}_n)$, we define the set:

$$\Delta(\mathbf{y}) = \left\{ (\lambda_1, ..., \lambda_n) \in [0,1]^n : \mathbf{y} = \sum_{j=1}^{n} \lambda_j \mathbf{y}_j, \ \sum_{j=1}^{n} \lambda_j = 1 \right\}.$$

Note that $\Delta(\mathbf{y})$ is compact for any $\mathbf{y} \in conv(\mathbf{y}_1, ..., \mathbf{y}_n)$ since it is complete (closed) and totally bounded. We may therefore define a function $C : conv(\mathbf{y}_1, ..., \mathbf{y}_n) \to \mathbb{R}$ by setting

$$C(\mathbf{y}) = \max \left\{ \sum_{j=1}^{n} \lambda_j t_j : (\lambda_1, ..., \lambda_n) \in \Delta(\mathbf{y}) \right\} \text{ for } \mathbf{y} \in conv(\mathbf{y}_1, ..., \mathbf{y}_n).$$

Let $\lambda_j(\mathbf{y}) \in [0,1]$, $1 \le j \le n$, be such that $C(\mathbf{y}) = \sum_{j=1}^{n} t_j \lambda_j(\mathbf{y})$. Since $\lambda_j(\mathbf{y}_i) = \delta_{i,j}$, one has $C(\mathbf{y}_j) \ge t_j$ for any $1 \le j \le n$ and thus:

$$\sum_{j}^{n} b_j C(\mathbf{y}_j) \ge \sum_{j}^{n} b_j t_j > c. \tag{3.2}$$

On the other hand, the left inequality in 3.1 implies that for any matrix $M{=}(m_{i,j}) \in \mathcal{M}(X,Y)$ one has:

$$c > \sum_{j}^{n} t_j s_j(M) = \sum_{i}^{m} \sum_{j}^{n} a_i t_j m_{i,j}. \tag{3.3}$$

Let us now consider the extension $\widehat{C}$ of the function C to $\mathbb{R}^k$ given by:

$$\begin{cases} \widehat{C}(\mathbf{y}) = C(\mathbf{y}) \quad \text{if } \mathbf{y} \in conv(\mathbf{y}_1, ..., \mathbf{y}_n) \text{ and} \\ \widehat{C}(\mathbf{y}) = 0 \quad \text{for } \mathbf{y} = \mathbb{R}^k \backslash conv(\mathbf{y}_1, ..., \mathbf{y}_n). \end{cases}$$

Now, $\widehat{C}(\mathbf{y})$ is a concave function on $\mathbb{R}^k$and by a previous argument $\mathbf{x}_i \in conv(\mathbf{y}_1, ..., \mathbf{y}_n)$, for any $1 \le i \le m$, so that the numbers $\lambda_j(\mathbf{x}_i)$ are well-defind for any $1 \le i \le m$, $1 \le j \le n$. Note that $\lambda_j(\mathbf{x}_i) \in \mathcal{M}(X,Y)$ and that by 3.2 and 3.3 the following holds:

$$\sum_{i=1}^{m} a_i \widehat{C}(\mathbf{x}_i) = \sum_{i=1}^{m} \sum_{j=1}^{n} a_i t_j \lambda_j(\mathbf{x}_i) < \sum_{j=1}^{n} b_j t_j \le \sum_{i=1}^{n} b_i \widehat{C}(\mathbf{y}_i).$$

However, this contradics condition (ii) of Theorem 2 and thus the thesis. It follows that $\mathbf{b} \in \mathcal{F}(X, Y)$, which complete the proof. □

COMMENTS: TO BE WRITTEN

If $m = n$ and $a_i = b_j = \frac{1}{n}$, $1 \leq i, j \leq n$, then the following criterion compares *multivariate distributions with the same marginals*:

Definition 3. *Let X, $Y \in \Gamma_n^k$, then Y is said to majorize X, written $X \prec_u Y$, if there exists a $n \times n$ bistochastic matrix B, such that $X = BY$.*

The foregoing definition extends the notion of majorization under Theorem 1.(i) to the multivariate case. It essentially means that the average is a *smoothing-operation*, making the components of X more spread out than the components of Y. Then, Theorem 2 and Birkhoff-Von Neumann's Theorem (see Marshall and Olkin (1979) ch.1) imply that $\prec_h$ induces $\prec_u$ and therefore:

Corollary 1 (Marshall and Olkin [9] and Koshevoy [7, 8])**.** *If $X, Y \in \Gamma_n^k$ then the following conditions are equivalent:*

(i) *There exists some bistochastic matrix B such that $X = BY$;*

(ii) *The inequality $\sum_{i=1}^n \phi(\mathbf{x}_i) \leq \sum_{i=1}^n \phi(\mathbf{y}_i)$ holds for any convex function $\phi : \mathbb{R}^k \to \mathbb{R}$.*

It is worth noticing that for $k = 1$, Corollary 1 amounts to Theorem 1 stated above.

4. Conclusions

TO BE WRITTEN

References

[1] Blackorby, C. and D. Donaldson (1993) - Adult-equivalence scales and the economic implementation of interpersonal comparisons of well-being - *Social Choice and Welfare* 10, 335-361.

[2] Dardanoni, V.

[3] de Laiglesia J. R. and C. Morrisson (2008) - Household Structures and Savings: Evidence from Household Surveys - OECD w.p. 267.

[4] Farina, F. and E. Savaglio (eds) (2007) - *Inequality and Economic Integration* - Routledge, London.

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

[6] Ebert, U. and P. Moyes (2000) - Consistent income tax structure when households are heterogeneous - *Journal of Economic Theory* 90, 116-150.

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

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

[9] Marshal A. W. and I. Olkin (1979) - *Inequalities: Theory of Majorization and its applications* - Academic Press, New York.

[10] Rudin, W. (1991) - *Functional Analysis* - McGraw-Hill.

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

DMQTE- Università "G. D'Annunzio" of Pescara, viale Pindaro, 42 - Pescara (Italy)

E-mail address: ernesto@unich.it