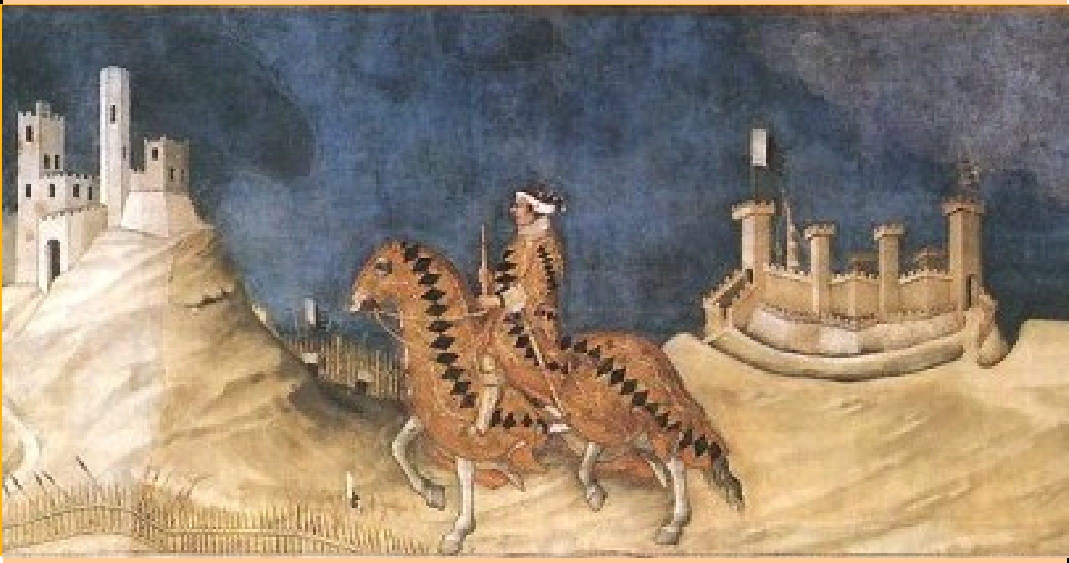
UNIVERSITÀ DEGLI STUDI DI SIENA

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Ernesto Savaglio

Multidimensional Inequality with Variable Household Weight

n. 667 – Dicembre 2012



Abstract. We establish the multidimensional couterpart of the fundamental theorem on the inequality measurement due to Hardy, Littlewood and Polya (HLP, 1934) for the case of household with different weights

JEL Classification: D31, D63, I31

Ernesto Savaglio, DEC, University of Pescara, DEPS, University of Siena and GRASS

1. Introduction

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 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 more general setting of distributions of households with different needs/compositions endowed with different goods. The importance of defining an order-preserving function in such a setting relies on the one-to-one correspondence between isotone functions, maps

[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.

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.

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. Theorem

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 interpret an element $x_{i,j}$ of a multivariate distribution $X \in \Gamma_n^k$ 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 $\mathbb{M}^{m,n}$ the set of all *row-stochastic* $m \times n$ matrices R, namely nonnegative rectangular matrices with all its row sums equal to one. 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.

We now introduce a binary relational system to compare (in terms of relative inequality) *multi-dimensional distributions* (of individual/attributes) 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 1. *Let* $(X, \mathbf{p}) \in \Lambda_m^k$ *and* $(Y, \mathbf{q}) \in \Lambda_n^k$*. Then, we say that* $(X, \mathbf{p})$ *is less unequal than* $(Y, \mathbf{q})$*, denoted to as* $(X, \mathbf{p}) \prec_h (Y, \mathbf{q})$*, if there exists a matrix* $R \in \mathbb{M}^{m,n}$ *such that*

$$X = RY \quad \textit{and} \quad \mathbf{q} = \mathbf{p}R.$$

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

Then, we have:

Theorem 1. *Let $(X,\mathbf{p}) \in \Lambda_m^k$ and $(Y,\mathbf{q}) \in \Lambda_n^k$. The following conditions are equivalent:*

(i)**:** $(X,\mathbf{p}) \prec_h (Y,\mathbf{q})$*;*

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

$$\sum_{i=1}^{m} p_i \phi(\mathbf{x}_i) \le \sum_{j=1}^{n} q_j \phi(\mathbf{y}_j).$$

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'. Indeed:

First notice that $X = RY$ and $\mathbf{q} = \mathbf{p}R$ for some $R \in \mathbb{M}^{m,n}$ mean that $x_i = \sum_{j=1}^{n} r_{ij} y_j$ for any $i \in \{1, ..., m\}$ and $q_j = \sum_{i=1}^{m} p_i r_{ij}$ for any $j \in \{1, ..., n\}$. Then, for any $\phi : \mathbb{R}^k \to \mathbb{R}$ continuous and convex, $X = RY$ implies by Jensen inequality that:

$$\phi\left(\sum_{j=1}^{n} r_{ij} y_j\right) \le \sum_{j=1}^{n} r_{ij} \phi(\mathbf{y}_j) \quad \text{for any } 1 \le i \le m,$$

namely:

$$\phi(\mathbf{x}_i) \le \sum_{j=1}^{n} r_{ij} \phi(\mathbf{y}_j) \quad \text{for any } 1 \le i \le m.$$

Now, if we multiply both members by p_i and we sum, then we get:

$$\begin{aligned}
\sum_{i=1}^{m} p_i \phi(\mathbf{x}_i) &\le \sum_{i=1}^{m} p_i \sum_{j=1}^{n} r_{ij} \phi(\mathbf{y}_j) = \\
&= p_1 \sum_{j=1}^{n} r_{ij} \phi(\mathbf{y}_j) + ... + p_m \sum_{j=1}^{n} r_{ij} \phi(\mathbf{y}_j) = \\
&= \sum_{j=1}^{n} \phi(\mathbf{y}_j) \left[p_1 r_{1j} + ... + p_m r_{mj}\right] = \\
&= \sum_{j=1}^{n} \phi(\mathbf{y}_j) \sum_{i=1}^{m} p_i r_{ij} = \\
&= \sum_{j=1}^{n} \phi(\mathbf{y}_j) q_j,
\end{aligned}$$

i.e.:

$$\sum_{i=1}^{m} p_i \phi(\mathbf{x}_i) \le \sum_{j=1}^{n} q_j \phi(\mathbf{y}_j)$$

as required.

In order to prove that $(ii) \Rightarrow (i)$, assume (ii) holds for $X \in \Gamma_m^k$ and $Y \in \Gamma_n^k$ and that $\mathbf{p} = (p_1, ..., p_m) \in \Xi_m$ and $\mathbf{q} = (q_1, ..., q_n) \in \Xi_n$. Then, define the set:

$$\mathfrak{C}(X,Y) = \left\{ R = (r_{ij}) \in \mathbb{M}^{m,n} \mid \mathbf{x}_i = \sum_{j=1}^{n} r_{ij}\mathbf{y}_j, \text{ for any } 1 \leq i \leq m \right\}$$

as the set of all $(m \times n)$ row-stochastic matrices such that $\mathbf{x}_i$ is in the convex hull of the rows of Y, i.e $\mathbf{x}_i \in \mathit{conv}(Y)$, where $\mathit{conv}(Y)$ is defined as the set of all convex combinations of the points $\mathbf{y}_j$ of Y with $j \in \{1, ..., n\}$.

The set $\mathfrak{C}(X,Y)$ is convex, because the set of all row-stochastic matrices is closed by convex combination. Moreover, it is closed. Indeed, suppose not, namely that $\mathfrak{C}(X,Y)$ is open. Then, its complement $\mathfrak{C}'(X,Y) = \left\{ R = (r_{ij}) \in \mathbb{M}^{m,n} \mid \mathbf{x}_i > \sum_{j=1}^{n} r_{ij}\mathbf{y}_j, \text{ for any } 1 \leq i \leq m \right\}$ is closed. But, $\mathbf{x}_i - \sum_{j=1}^{n} r_{ij}\mathbf{y}_j > 0$ is an open halfspace in $\mathbb{R}^{m,n}$, that is an open set by Theorem 1.7.1 pg. 33 in Webster (1994). Hence, $\mathfrak{C}(X,Y)$ must be closed. In other terms, for any i, $\mathbf{x}_i = \sum_{j=1}^{n} r_{ij}\mathbf{y}_j$ is a hyperplane in $\mathbb{R}^{m,k}$ dimensions that by definition has no interior points, namely no points that are the centre of some open ball which lies in A. Since, the interior of any set is open, and no points in $\mathbb{M}^{m,n}$ are interior points of the set, then that set cannot be open. Therefore, $\mathfrak{C}(X,Y)$ is closed.

Finally, $\mathfrak{C}(X,Y)$ is non-empty. To prove that $\mathfrak{C}(X,Y) \neq \emptyset$ means to show that if (ii) holds than $X = RY$ for some $R \in \mathbb{M}^{m,n}$, i.e. there exists at least one $R \in \mathfrak{C}(X,Y)$ such that $\mathbf{x}_i = \sum_{j=1}^{n} r_{ij}\mathbf{y}_j$ for any $1 \leq i \leq m$ that means that any entry of X is a convex combination of the (column) entries of Y. Now, in order to prove that $\mathfrak{C}(X,Y) \neq \emptyset$, it is sufficient to show that if condition (ii) holds, then $\mathbf{x}_i = \mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n)$ for any $i = \{1, ..., m\}$, where $\mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n)$ represents the convex hull of all vectors $\mathbf{y}_1, ..., \mathbf{y}_n$. So, suppose by contradiction that there is a $\mathbf{x}_j \notin \mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n)$ for some $j \in \{1, ..., m\}$ and then consider the function $\phi : \mathbb{R}^k \to \mathbb{R}$ defined as $\phi(\mathbf{t}) = d(\mathbf{t}, \mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n))$, where d is the Haussdorf distance in $\mathbb{R}^k$ of a point from a set. We observe that $\phi(\mathbf{t})$ is a convex function since $\mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n)$ is a non-empty convex subset of $\mathbb{R}^k$. Then, we notice that $\phi(\mathbf{x}_j) > 0$, $\phi(\mathbf{y}_l) = 0$ for any $1 \leq l \leq n$. Thus, it follows that $\underbrace{\sum_{i=1}^{m} p_i \phi(\mathbf{x}_i)}_{>0} > \underbrace{\sum_{i=1}^{n} q_j \phi(\mathbf{y}_i)}_{=0 \text{ for any } j}$ which contradicts condition (ii) above. Hence, $\mathbf{x}_i = \mathit{conv}(\mathbf{y}_1, ..., \mathbf{y}_n)$ for any $i = \{1, ..., m\}$, and $\mathfrak{C}(X,Y)$ must be non-empty as required. To sum up, we have proved that if condition (ii) above holds then $X = RY$.

In order to prove that $\mathbf{q} = \mathbf{p}R$, take a $R = (r_{i,j}) \in \mathfrak{C}(X,Y)$ and define a vector $\mathbf{s}(R) = (s_1(R), ..., s_n(R)) \in \mathbb{R}^n$ as follows $s_j(R) = \sum_{i=1}^{m} p_i r_{i,j}$, for any $1 \leq j \leq n$. Then, define a set:

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

and note that $\mathcal{F}(X,Y)$ is convex. Indeed, for any R, $R' \in \mathfrak{C}(X,Y)$ and $\alpha \in [0,1]$ $\alpha \sum_{i=1}^{m} p_i r_{ij} + (1-\alpha) \sum_{i=1}^{m} p_i r'_{ij} = \sum_{i=1}^{m} p_i \left[\alpha r_{ij} + (1-\alpha) r'_{ij}\right]$ or equivalently $\alpha \mathbf{s}(R) + (1-\alpha)\mathbf{s}\left(R'\right) = \mathbf{s}\left(\alpha R + (1-\alpha) R'\right)$.

Since the set of row-stochastic matrices is closed by convex combination, $\mathbf{s}\left(\alpha R+(1-\alpha)R^{'}\right)\in\mathcal{F}(X,Y)$, i.e. $\mathcal{F}(X,Y)$ is convex.

Moreover, $\mathcal{F}(X,Y)$ is closed. Indeed, it is the set of linear combinations of the elements of a set that is closed and convex.

Finally, we need to show that there is a matrix $R^{*}=\left(r^{*}_{i,j}\right)\in\mathfrak{C}(X,Y)$ such that $\mathbf{s}(R^{*})=\mathbf{q}$, namely $\mathbf{q}\in\mathcal{F}(X,Y)$. In order to prove that, suppose by contradiction that $\mathbf{q}\notin\mathcal{F}(X,Y)$. Then, since $\mathcal{F}(X,Y)$ is a closed and convex set, by the Hahn-Banach theorem (see Rudin (1977))[4], it follows that there exists a $\gamma\in\mathbb{R}$ and a vector $(t_1,...,t_n)\in\mathbb{R}^n$ such that the following inequality holds:

$$\sum_{j=1}^{n}t_jq_j<\gamma<\sum_{j=1}^{n}t_js_j \text{ for any } (s_1,...,s_n)\in\mathcal{F}(X,Y). \tag{2.1}$$

Consider first the left inequality in 2.1, i.e. $\sum_{j=1}^{n}t_jq_j<\gamma$. Take a vector $\mathbf{y}\in conv(\mathbf{y}_1,...,\mathbf{y}_n)$ and define the simplex:

$$\Delta(\mathbf{y})=\left\{(\zeta_1,...,\zeta_n)\in[0,1]^n:\mathbf{y}=\sum_{j=1}^{n}\zeta_j\mathbf{y}_j,\ \sum_{j=1}^{n}\zeta_j=1\right\},$$

namely the set of all points whose scalar product with $\mathbf{y}_j$ is $\mathbf{y}$. The set $\Delta(\mathbf{y})$ is a compact, i.e. a closed and bounded, set because $[0,1]$ is a closed and bounded interval. Then, define a convex function $C:conv(\mathbf{y}_1,...,\mathbf{y}_n)\to\mathbb{R}$ as follows:

$$C(\mathbf{y})=\max\left\{\sum_{j=1}^{n}\zeta_jt_j:(\zeta_1,...,\zeta_n)\in\Delta(\mathbf{y})\right\} \text{ for } \mathbf{y}\in conv(\mathbf{y}_1,...,\mathbf{y}_n).$$

Since $C(\mathbf{y}_j)=t_j\zeta_j<t_j$ for any $1\leq j\leq n$ because $\zeta_j\in[0,1]$, we have:

$$\sum_{j}^{n}q_jC(\mathbf{y}_j)<\sum_{j}^{n}q_jt_j<\gamma. \tag{2.2}$$

Take now the right inequality in 2.1, i.e. $\gamma<\sum_{j=1}^{n}t_js_j$. Then, for any $R=(z_{ij})\in\mathbb{M}^{m,n}$, we have:

$$\gamma<\sum_{j}^{n}t_js_j(Z)=\sum_{j=1}^{n}t_j\sum_{i=1}^{m}p_ir_{i,j}=\sum_{i}^{m}\sum_{j}^{n}p_it_jr_{i,j}.$$

According to the previous result $\mathbf{x}_i=conv(\mathbf{y}_1,...,\mathbf{y}_n)$ for any $1\leq i\leq m$, so the number $\zeta_j(\mathbf{x}_i)$ are well-defined for $1\leq i\leq m$ and $1\leq j\leq m$. Further, note that $\left(\zeta_j(\mathbf{x}_i)\right)\in\mathbb{M}^{m,n}$ and therefore:

$$\sum_{i}^{m}\sum_{j}^{n}p_it_jz_{i,j}=\sum_{i}^{m}p_iC(\mathbf{x}_i). \tag{2.3}$$

[4]The Hahn-Banach Theorem in the separation form states that if A and B are two non-empty disjoint convex subsets of a real normed linear space X, with A that is compact and B that is closed, then A and B can be strictly separated by a closed hyperplane. This means that there exists a nonzero continuous linear functional ϕ and a real number α such that $\phi(x)<\alpha<\phi(y)$ for all $x\in A$ and $y\in B$.

Expressions 2.2, 2.3 together entail:

$$\sum_{j=1}^{n} q_j C\left(\mathbf{y}_j\right) < \sum_{i=1}^{m} p_i C\left(\mathbf{x}_i\right).$$

that contradicts condition (ii) above and thus the thesis that $\mathbf{q} \in \mathcal{F}\left(X, Y\right)$ i.e. $\mathbf{q} = \mathbf{p}R$ that completes the proof. □

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.