From Unidimensional to Multidimensional Inequality, Welfare and Poverty Measurement.

Claudio Zoli
Department of Economics
University of Verona, Italy*

January 2009
Winter School on "Inequality and Welfare Theory", Canazei

Abstract

In this handout we review results concerning representation of partial orders via stochastic orders and investigate their applications to some classes of stochastic dominance conditions applied in inequality, poverty and welfare measurement. The results obtained in an unidimensional framework are extended to multidimensional analysis. In particular we provide an overview of the main issues concerning aggregation of multidimensional distributions into synthetic indicators as the Human Development Index or Social Welfare Functions. Moreover we explore the potential for multidimensional evaluations based on the partial orders induced by different criteria of majorization. The lecture is divided into 4 parts: (i) an introduction to basic results concerning unidimensional evaluations of inequality, welfare and poverty (ii) an illustration of the problems of aggregation of evaluations when applied in the multidimensional context where individuals exhibit various attributes, (iii) the discussion of the potentials and limits of the application of generalizations of the majorization approach to comparisons of multidimensional distributions (iv) a brief overview of some results on multidimensional stochastic orders.

*Claudio Zoli, Dipartimento di Scienze Economiche, Università di Verona, Viale dell'Università 4, 37129-Verona, Italy. E-MAIL: claudio.zoli@univr.it

Contents

1 Introduction

When comparing income distributions in terms of inequality, the most common tool adopted in applied works in the Lorenz curve. Comparisons made in terms of Lorenz curves account for heterogeneity of views of policy makers or evaluators about the degree of inequality aversion to adopt in the evaluation. It is indeed well known that if a Lorenz curve of a distribution is above the one of another distribution then inequality will be evaluated as lower for the former compared to the latter by a set of inequality indices. A source of "discomfort" in applying the concept of dominance in terms of the Lorenz curve relies instead on the fact that it induces a partial order, that is, not all distributions can be unanimously compared since it is possible that for some distributions the Lorenz curves intersect. The analysis of the dual aspect of "lack of completeness" and "unanimity of evaluations" is one of the issues explored in the first part of this lecture. We review results on the literature on "stochastic orders" i.e. orders of distributions based on unanimous dominance in terms of family of evaluation functions and investigate their properties and their application to welfare, inequality and poverty measurement.[1]

The first part of the lecture will focus on the analysis of unidimensional distributions, e.g. income or consumption vectors of a given population. However the most interesting questions arise when facing comparisons based on multidimensional distributions, where for instance also education, health, or endowment of bundles of goods are taken into account. There exist therefore two different ways to approach the multidimensional problem, the first one makes use of appropriate procedures to embody all information within an unique money metric indicator (e.g. equivalence scales or calculating budget income). Within this framework the different opinions concerning the aggregating procedures could be taken into account and partial ranking criteria could be specified in order to be consistent with a given range of admissible aggregating functions.

From the second point of view, inequality, as well as poverty and welfare, are *explicitly considered as multidimensional phenomena*. Therefore the domain of the evaluation process is left multidimensional and ranking criteria are specified over the set of all variables. Within this second framework it becomes natural to first investigate whether it is more informative to focus first on the distribution of each attributes across the individuals as done by the Human Development Index or consider as a starting point the aggregation of each attribute per individual identifying the a distribution of utilities or individual indicators of well being. In the second part of the lecture we explore results on consistent aggregation obtained following these two

[1] Note that in what follows:
$\succsim$ denotes a binary relation over a set of alternative,
$\geqslant$ denotes a stochastic order over distribution or random variables,
$\succcurlyeq$ denotes a stochastic dominance condition on distributions or random variables, while
$\geq$ denotes the usual inequality symbol.

procedures highlighting the implications for many indices currently adopted in the literature. Then we move to the more demanding quest of focussing on comparisons that use information on the whole matrix without filtering it through any intermediate aggregation. This second approach is based on the concept of majorization [see Marshall and Olkin, 1974] that extends to the multivariate case the machinery applied in the univariate case to obtain the Lorenz dominance criterion. A number of critical issues arise in the multidimensional context and various criteria of dominance can be adopted. We will explore some of then and then we will get back to the stochastic orders framework illustrating some multidimensional results deriver in the statistics/probability and decision theory literature.

The structure of the lecture can be divide into 4 parts: (i) an introduction to basic results concerning unidimensional evaluations of inequality, welfare and poverty (ii) an illustration of the problems of aggregation of evaluations when applied in the multidimensional context where individuals exhibit various attributes, (iii) the discussion of the potentials and limits of the application of generalizations of the majorization approach to comparisons of multidimensional distributions (iv) a brief overview of some results on multidimensional stochastic orders. Each one of these parts can be addressed almost separately, this is particularly the case for the distinction between the unidimensional analysis (part (i)) and the multidimensional case (parts (ii), (iii), (iv)).

1.0.1 Aims of the lecture

Overview of selected issues underlying the theory of measurement of inequality, welfare, poverty and well being.

Two broad perspectives:

1. **Unidimensional**: Individuals/households are homogeneous in all ethically relevant characteristics except consumption or income.

2. **Multidimensional**: heterogeneous individuals exhibiting differences in a number of "characteristics" (transferable and non transferable) e.g. income, health, housing, bundles of goods, education, household size, level of needs.....

A number of interrelated perspectives of evaluation can be taken into account:

1. *Inequality* (focuses on dispersions across agents, i.e. considers how the cake (e.g. GDP) is shared in the population),

2. *Welfare* (takes into account also the size of the cake, for instance each agent can improve her situation in a world where the per-capita income is larger even though is more unequally distributed)

3. *Poverty* (focuses on deprived agents, size and dispersion matters but the concern is only for those deprived, e.g. with income below the poverty line)

4. *Well being*: multidimensional perspective focussing on size and dispersion. A more unequal distribution in one attribute across the population might compensate for the unequal distributions in other attribute, i.e. it is not always the case that an increase in inequality in one attribute reduces society well being. This is clearly the case if agents that are better of in the distribution of some attributes continue to be advantaged by the distribution of other attributes, that is if the correlation between the distribution of the attributes is positive.

Evaluations can be formalized through:

1. *Complete rankings*, i.e. indices or welfare functions, or
2. *Partial rankings*, i.e. identification of dominance conditions that may not rank any pair of distribution, this is the case for instance for the Lorenz dominance condition based on comparisons of Lorenz curves. Dominance occurs in this case if the Lorenz curve of one distribution is above the one of another distribution, if these curve intersect the dominance test is not conclusive.

Our Concern is

- To provide some intuitions on the interrelations between the various concepts in the unidimensional case and then move to the MORE INTERESTING multidimensional case.........

1.1 Unidimensional/Multidimensional set up.

Consider:

- n homogeneous individuals $i = 1, 2, ..., n \geq 2$
- $\mathbb{R}^n_+$: $n-$dimensional (^{n}) vector of non-negative ($_+$) real numbers ($\mathbb{R}$)
- $d \geq 1$ characteristics, goods, attributes, attainments (e.g. income) $j = 1, 2, 3, ..., d$
- Distribution $X \in \mathbb{R}^{n \times d}_+$

$$X = \begin{bmatrix} x_{11} & ... & \cdots & x_{1(d-1)} & x_{1d} \\ x_{21} & x_{22} & \cdots & \cdots & x_{2d} \\ \vdots & . & & & \vdots \\ x_{n1} & x_{n2} & \cdots & \cdots & x_{nd} \end{bmatrix}$$

Rows identify individuals/households and columns are associated with attributes or characteristics

- $\mathbf{x}_{.j}$ distribution of attribute j across all individuals, $\mathbf{x}_{i.}$ distribution of all the attributes for individual i, $\mathbf{x} \in \mathbb{R}^d_+$ generic vector of values for the attributes, or $\mathbf{x} \in \mathbb{R}^n_+$ generic vector of distribution for n individuals.

- $F_X(x)$ or $F(x)$cumulative distribution function: percentage of individuals with income not higher than x (unidimensional case) or in general $F_X(\mathbf{x})$ percentage of individuals whose vector of attributes is not larger than $\mathbf{x} \in \mathbb{R}^d_+$ (multidimensional case)

- $\mu(\mathbf{x}_{.j})$ average of distribution $\mathbf{x}_{.j}$ of attribute j (e.g. income) :

$$\mu(\mathbf{x}_{.j}) = \sum\nolimits_{i=1}^{n} x_{ij}/n.$$

- $\hat{\mathbf{x}}_{.j}$ ordered distribution of attribute j: $\hat{x}_{(1)j} \leq \hat{x}_{(2)j} \leq ..\hat{x}_{(i)j} ... \leq \hat{x}_{(n)j}$

- $I(X)$ Inequality index, $I : \mathbb{R}^{n \times d}_+ \to \mathbb{R}$ [function from the set of all admissible distributions $\mathbb{R}^{n \times d}_+$ to ($\to$) the set of real numbers ($\mathbb{R}$)]

- $W(X)$ Social Evaluation Function (SEF) it is defined over incomes $W : \mathbb{R}^{n \times d}_+ \to \mathbb{R}$

- $z_j > 0$ poverty line for attribute j, $z \in \mathbb{R}^d_+$ vector of poverty lines for each attribute.

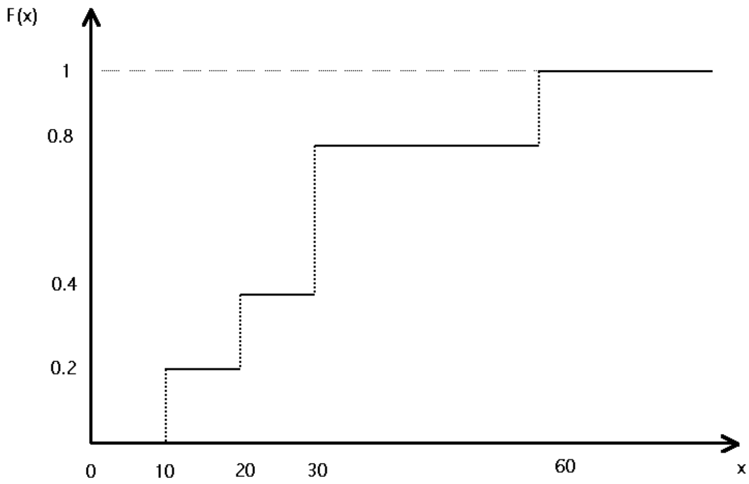
- $P(X,z)$ Poverty index, $P : \mathbb{R}^{n \times d}_+ \times \mathbb{R}^d_{++} \to \mathbb{R}$ [function from the set of all admissible distributions ($\mathbb{R}^{n \times d}_+$) and from the set of positive values ($\mathbb{R}_{++}$) of any poverty line z_j, to ($\to$) the set of real numbers ($\mathbb{R}$)]

2 The Unidimensional case

2.0.1 Making use of distribution functions

As a starting point one might analyse distributions making use of statistical graphical tools. For instance the primitive tool both for continuous or discrete distributions can be the Cumulative Distribution Function. It can also be specified [later on] for multidimensional distributions.

Definition 2.1 (Cumulative Distribution Function) $F : \mathbb{R}_+ \to [0, 1]$ *Function* $F(x)$ *plotting the proportion of income units within the population with income at most* x.

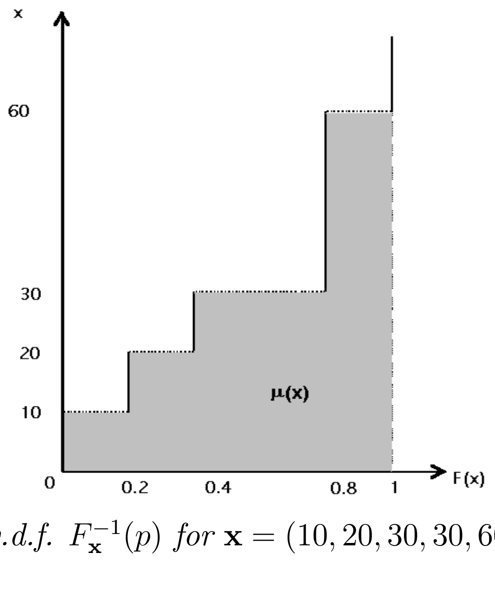


C.d.f. $F_{\mathbf{x}}(x)$ *for* $\mathbf{x} = (10, 20, 30, 30, 60)$

It is possible to use the information in the C.d.f. in order to construct a dual representation. Reversing the graphs we get the *Pen's Parade (or inverse distribution function)*: The parade of all the incomes in increasing order weighted by the percentage of individuals to whom these incomes belong (sample weight).

Definition 2.2 (Inverse Distribution Function) $F^{-1} : [0, 1] \rightarrow \mathbb{R}_+$. *Function* $F^{-1}(p)$ *plotting the income level corresponding to the* p^{th} *quantile of the population once incomes are ranked in ascending order, i.e. left continuous inverse distribution:*

$$F^{-1}(p) = \inf\{x \in \mathbb{R}_+ : F(x) \geq p\}.$$



Inv.d.f. $F_{\mathbf{x}}^{-1}(p)$ *for* $\mathbf{x} = (10, 20, 30, 30, 60)$

2.1 How to rank distributions?

2.1.1 Stochastic orders

We consider the issue of unanimous dominance for a set of indices or evaluation functions. We focus on two broad families of functionals. commonly adopted in inequality, welfare and risk analysis.

Definition 2.3 *Additively decomposable order* $\geqslant_{\mathcal{U}}$

$$\begin{aligned} X &\geqslant {}_{\mathcal{U}} Y \\ &\Leftrightarrow \\ W_u(X) &= \int_{\mathbb{R}} u(x)dF_X(x) \geq \int_{\mathbb{R}} u(x)dF_Y(x) = W_u(Y), \quad \forall u \in \mathcal{U} \\ &\quad \textit{alternatively} \\ \frac{1}{n}\sum_{i=1}^{n} u(x_i) &\geq \frac{1}{n}\sum_{i=1}^{n} u(y_i) \quad \forall u \in \mathcal{U} \end{aligned}$$

The key property of the utilitarian/expected utility representation is Independence:

Definition 2.4 (Independence) *Joining (or mixing) two distributions (or populations of individuals) with a third distribution (another group of individuals) the ranking of the new distributions obtained is consistent with the ranking of the former two distributions:* $W_u(X, Z) \geq W_u(Y, Z)$ *if and only if* $W_u(X) \geq W_u(Y)$.

The realization of each individual is transformed according to the function u while his/her weight in the final formula enters linearly.

Next family is dual w.r.t. the previous and is based on weighted averages of the realizations, where weight depends on the ranking of each realization.

Definition 2.5 *Rank dependent (dual) order* $\geqslant_{\mathcal{V}}$

$$\begin{aligned} X &\geqslant {}_{\mathcal{V}} Y \\ &\Leftrightarrow \\ W_v(X) &= \int_0^1 v(p) \cdot F_X^{-1}(p)dp \geq \int_0^1 v(p) \cdot F_Y^{-1}(p)dp = W_v(Y) \ \forall v \in \mathcal{V} \\ &\quad \textit{alternatively} \\ \frac{1}{n}\sum_{i=1}^{n} v_i \cdot \hat{x}_{(i)} &\geq \frac{1}{n}\sum_{i=1}^{n} v_i \cdot \hat{y}_{(i)} \ \forall v \in \mathcal{V} \\ \textit{where } v_i &\geq 0;\ \hat{x}_{(1)} \leq \hat{x}_{(2)} \leq ... \leq \hat{x}_{(i)} \leq ... \leq \hat{x}_{(n)} \end{aligned}$$

The key property this rank dependent/generalized Gini representation is Comonotonic Independence:

Definition 2.6 (Comonotonic Independence) *When adding to two distributions (i.e. distributions of labour income) a third distribution (e.g. a distribution of capital income) that is comonotonic w.r.t. the former two (i.e. capital income is ranked in the same order as the labour incomes in the former distributions) then the two new distributions obtained is consistent with the ranking of the former two distributions:* $W_v(X+Z) \geq W_v(Y+Z)$ *if and only if* $W_v(X) \geq W_v(Y)$.

The realization of each individual is considered linearly in the final evaluation while the individual weight in the final formula enters through v_i according to his/her position in the ranking of the attribute.

Remark 2.1 *These criteria are partial orders i.e. they do not necessarily provide a clear-cut ranking. For some distributions the answer may not be conclusive this is the case when for instance* $W_v(X) > W_v(Y)$ *for some* $v \in \mathcal{V}$ *but* $W_{v'}(X) < W_{v'}(Y)$ *for some others* $v' \in \mathcal{V}$

2.2 Implementing stochastic orders:

2.2.1 Comparison Tests

Consider X income distribution of finite mean $\mu(X)$ [or μ_X], defined on a bounded support in $\mathbb{R}_+$.

- Is it possible to device tools that we can apply directly in order to test dominance of one distribution over another?

The most common tools applied in inequality analysis to compare income distributions are indeed the partial orders induced by the stochastic dominance conditions (direct and inverse).

For instance typical dominance conditions are:

Definition 2.7 (Lorenz Dominance) *Define the* ***Lorenz curve*** *for* X*:*

$$L_X(p) := \int_0^p \frac{F_X^{-1}(t)}{\mu(X)} dt.$$

Income profile X *Lorenz dominates income profile* Y, $X \succcurlyeq_L Y$, *if and only if*

$$L_X(p) \geq L_Y(p) \text{ for all } p \in [0,1].$$

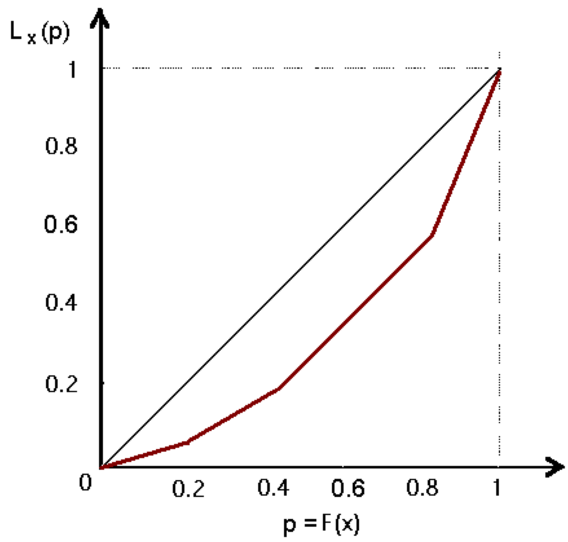
That is the Lorenz curve is derived for distributions of incomes ranked in increasing order and plots the proportion of total income belonging to the poorer $100p^{th}$ percentile of the population, in the case of discrete distributions for $p = i/n$ we have

$$L_x(i/n) = \frac{\sum_{j=1}^{i} \hat{x}_{(j)}}{\sum_{j=1}^{n} x_j} \qquad \text{where } \hat{x}_{(i)} \leq \hat{x}_{(i+1)}$$

for all the points in between i/n and $(i+1)/n$ the Lorenz curve is a straight line joining $L_x(i/n)$ and $L_x(i+1/n)$. We denote $L_x(p)$, for $p \in [0,1]$, the curve obtained joining the points $L_x(i/n)$

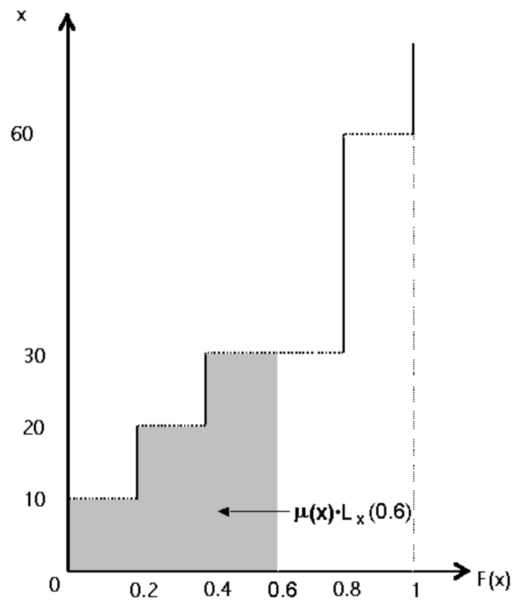
Example 2.1 *For distribution* $\mathbf{x} \in \mathbb{R}_+^n = (10, 20, 30, 30, 60)$ *we get*

$$[(i/n, L_x(i/n))] = \left[(0,0), (\frac{1}{5}, \frac{1}{15}), (\frac{2}{5}, \frac{1}{5}), (\frac{3}{5}, \frac{2}{5}), (\frac{4}{5}, \frac{3}{5}), (1,1)\right]$$



Lorenz curve for $\mathbf{x} = (10, 20, 30, 30, 60)$

Note that the Lorenz curve is obtained integrating the graph of the inverse distribution function an dividing by the average income.



Lorenz curve derived from inverse distribution.

Definition 2.8 (Generalized Lorenz Dominance) *Define the* ***generalized Lorenz curve*** *for X:*

$$GL_X(p) := \mu(X) \cdot L_X(p).$$

Income profile X generalize Lorenz dominates income profile Y, $X \succcurlyeq_{GL} Y$, if and only if

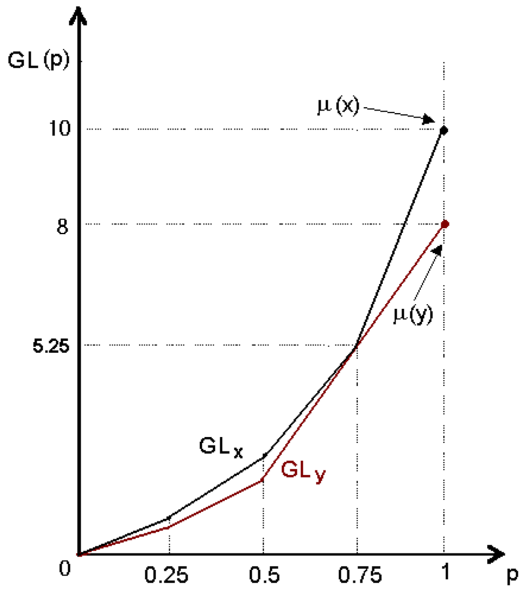
$$GL_X(p) \geq GL_Y(p) \text{ for all } p \in [0,1]\,.$$

See Kolm (1969), Shorrocks (1983). The notion of generalized Lorenz dominance takes into account not only the distribution of the shares of income but also the "size of the cake" i.e. the average income. Note that when $p = 1$, the top part of the GL curve focusses on the average income: $GL_X(1) = \mu(X)$. At the other extreme the GL dominance for low values of p requires lexicographic dominance in terms of the income of the poorest individuals.

Remark 2.2 *If $\mu(X) = \mu(Y)$, $\succcurlyeq_{GL} \iff \succcurlyeq_L$;*
or let X/μ_X denote the income distribution X normalized by its mean μ_X then $X/\mu_X \succcurlyeq_{GL} Y/\mu_Y \iff X \succcurlyeq_L Y$.

Example 2.2 *Here is the plot the Generalized Lorenz curves for the distributions in $\mathbb{R}^n_+$*

$$\mathbf{y} = (3, 7, 11, 11)\,, \;\; \mathbf{x} = (4, 8, 9, 19)$$



Generalized Lorenz Curves for $\mathbf{y} = (3, 7, 11, 11)$, $\mathbf{x} = (4, 8, 9, 19)$

The GL curves intersect therefore there is no-dominance.

Next results introduce the various notions of stochastic dominance most common in the risk/inequality literature and illustrate their connections and differences with the Lorenz type dominance conditions, moreover we will explore their connections with stochastic orders of the two families presented above.

2.2.2 Stochastic dominance conditions

"Direct" Stochastic Dominance (SD) is based on comparisons of distribution functions or survival functions and their integrals.

$\Delta F(x) = \Delta^1 F(x) := F_X(x) - F_Y(x)$ and $\Delta^i F(x) := \int_0^p \Delta^{i-1} F(t)dt$ for for $i = 2, 3, ...$The *SD* condition of order $i > 1$ ($\succcurlyeq_{SD[i]}$) is obtained comparing the integral of the inverse distribution functions derived recursively.

$\succcurlyeq_{SD[i]}$: Stochastic dominance condition of order i

Definition 2.9 (Stochastic Dominance) $X \succcurlyeq_{SD[i]} Y$ *iff* $\Delta^i F(x) \leq 0$ *for all* $x \in [0, \bar{x}]$.

If instead of comparing distribution functions we follow the dual approach of comparing inverse distributions we have the family of *Inverse Stochastic Dominance* (ISD) conditions introduced in Muliere and Scarsini (1989).

Let $\Delta(p) = \Delta^1(p) := F_X^{-1}(p) - F_Y^{-1}(p)$ and $\Delta^i(p) := \int_0^p \Delta^{i-1}(t)dt$ for $i = 2, 3, ...$ The *ISD* condition of order $i > 1$ ($\succcurlyeq_{ISD[i]}$) is obtained comparing the integral of the inverse distribution functions derived recursively.

Definition 2.10 (Inverse Stochastic Dominance) *$X \succcurlyeq_{ISD[i]} Y$ iff $\Delta^i(p) \geq 0$ for all $p \in [0,1]$.*

Remark 2.3 *Note that $X \succcurlyeq_{ISD[1]} Y$ denotes rank dominance (Saposnik, 1981) while $X \succcurlyeq_{ISD[2]} Y$ denotes generalized Lorenz dominance.*

Remark 2.4 *$X \succcurlyeq_{SD[i]} Y \Longrightarrow X \succcurlyeq_{SD[i+1]} Y$; $X \succcurlyeq_{ISD[i]} Y \Longrightarrow X \succcurlyeq_{ISD[i+1]} Y$;*

Remark 2.5 *$X \succcurlyeq_{SD[i]} Y \Longleftrightarrow X \succcurlyeq_{ISD[i+1]} Y$ for $i = 1, 2$.*
$X \succcurlyeq_{SD[i]} Y \not\Longrightarrow \not\Longleftarrow X \succcurlyeq_{ISD[i+1]} Y$ for $i = 3, 4, 5....$

2.2.3 Stochastic Orders and Stochastic Dominance: Integral Stochastic orders

Consider random variables (income profiles) X defined on $[0, \bar{x}]$ the Integral Stochastic Order (Additively decomposable) can be specified as

$$X \geqslant_{\mathcal{U}} Y \Leftrightarrow \int_0^{\bar{x}} u(x) d[\Delta F(x)] \geq 0 \;\; \forall u \in \mathcal{U}. \tag{1}$$

In order to investigate the relationships between stochastic orders $\geqslant_{\mathcal{U}}$ and stochastic dominance $\succcurlyeq_{SD[i]}$ we focus on "nested" classes of utility functions $u : [0, \bar{x}] \to \mathbb{R}$.

$$\mathcal{U}^1 := \{u \text{ continuous and non-decreasing}\}.$$
$$\mathcal{U}^2 := \{u \in \mathcal{U}^1 : \text{concave}\}.$$
$$\mathcal{U}^3 := \{u \in \mathcal{U}^2 : u' \text{ convex}\}.$$

Remark 2.6 *$P_\alpha(X, z) = \int_0^z (z - x)^\alpha dF_X$ is the absolute version of Foster, Greer and Thorbecke (1984) poverty index [FGT] for $\alpha \geq 0$, where $z > 0$ denotes the poverty line.*

For discrete distributions ranked in increasing order $\hat{\mathbf{x}}$, with q individuals not above the poverty line z the relative version of FGT index is $P_\alpha^r = z^{-\alpha} \cdot P_\alpha$:

$$P_\alpha^r(\mathbf{x}, z) = \frac{1}{n} \sum\nolimits_{i=1}^q \left(\frac{z - \hat{x}_{(i)}}{z} \right)^\alpha \;\; \textit{for } \alpha \geq 0,$$

Thus

$$\begin{aligned} P_0(\mathbf{x}, z) &= \frac{q}{n}, \;\; \textit{Head count ratio} \\ P_1(\mathbf{x}, z) &= \frac{1}{n} \sum\nolimits_{i=1}^q \left(z - \hat{x}_{(i)} \right) \\ P_2(\mathbf{x}, z) &= \frac{1}{n} \sum\nolimits_{i=1}^q \left(z - \hat{x}_{(i)} \right)^2. \end{aligned}$$

Dominance in terms of poverty indices for any poverty line can be related to stochastic dominance conditions.

Theorem 2.1 (Foster-Shorrocks (1988); Fishburn (1976)) *Let $n \in \{1,2,3\}$. The following statements are equivalent:*

(i) $P_{k-1}(X,z) \leq P_{k-1}(X,z)$ for all poverty lines $z \leq \bar{x}$.

(ii) $X \succcurlyeq_{SD[k]} Y$.

See also Dardanoni and Lambert (1988) for $n = 3$. Less poverty as evaluated according to the poverty index P_{n-1} and for all possible poverty lines coincides with stochastic dominance of order n.

Remark 2.7 *Note that $P_0(X,\bar{x}) = H(X,\bar{x})$: the headcount of profile X evaluated at the upper bound, i.e. the proportion of population with income not larger than $\bar{x}$, thus $P_0(X,\bar{x}) = 1 = P_0(Y,\bar{x})$.*

$P_1(X,\bar{x}) \leq P_1(Y,\bar{x}) \Longleftrightarrow \mu(X) \geq \mu(Y)$

If $\mu(X) = \mu(Y)$ then $P_2(X,\bar{x}) \leq P_2(Y,\bar{x}) \Longleftrightarrow \sigma^2(X) \leq \sigma^2(Y)$ where σ^2 denotes the variance.

Next result clarifies a further relation between poverty and evaluations in term of stochastic orders:

Theorem 2.2 *Let $k = 1,2,3$, the following statements are equivalent:*

(i) $X \geqslant_{\mathcal{U}^k} Y$

(ii) $X \succcurlyeq_{SD[k]} Y$ [and $\mu(X) \geq \mu(Y)$ for $k = 3$].

Concerning the role of variance it is also worth to point out that

Theorem 2.3 (Shorrocks-Foster (1987); Dardanoni-Lambert (1988)) *If (a) $\Delta^2 F(x)$ changes sign once, (b) $\mu(X) = \mu(Y)$ and (c) X dominates Y in terms of Leximin, then $X \succcurlyeq_{SD[3]} Y \Longleftrightarrow \sigma^2(X) \leq \sigma^2(Y)$.*

Remark 2.8 (Inequality Comparisons) *Concerning inequality comparisons based on relative inequality indices (i.e. scale invariant indices) then all previous statements [except for $\succcurlyeq_{SD[1]}$] hold provided that income profiles X/μ_X and Y/μ_Y are compared.*

Here we take the view that for distributions with equal means welfare dominance implies a reduction in inequality, more precisely $X/\mu_X \succcurlyeq_{SD[n]} Y/\mu_Y$ means that X shows less inequality than Y.

Since X/μ_X and Y/μ_Y exhibit the same mean equal to 1, then noticing that $X \succcurlyeq_{SD[1]} Y \Longrightarrow \mu(X) \geq \mu(Y)$ it follows that either X/μ_X and Y/μ_Y are not comparable according to $\succcurlyeq_{SD[1]}$ or they induce the same distribution function and therefore they are equivalent for all orders of stochastic dominance.

2.2.4 Stochastic Orders and Stochastic Dominance: Dual linear rank-dependent stochastic orders

We focus directly on the specification of the stochastic order based on weighting functions:

$$\begin{aligned} X &\geqslant {}_{\mathcal{V}}Y \Leftrightarrow \int_0^1 v(p) \cdot F_X^{-1}(p)dp \geq \int_0^1 v(p) \cdot F_Y^{-1}(p)dp \; \forall v \in \mathcal{V} \\ &\Leftrightarrow \int_0^1 v(p) \cdot \Delta(p)dp \geq 0 \; \forall v \in \mathcal{V} \end{aligned} \tag{2}$$

Before moving to the selection of the classes of weighting functions $\mathcal{V}$ we recall some parametric classes of SEFs, the class of S-Gini [single parameter] SEFs $\Xi(\delta; .)$ introduced in Donaldson and Weymark (1980, 1983) and Yitzhaki (1983). It is parameterized by $\delta \geq 1$ and is obtained letting $v(p) = \delta(1-p)^{\delta-1}$ that is

$$\Xi(\delta; X) \; : \; = \delta \int_0^1 (1-p)^{\delta-1} F_X^{-1}(p)dp \tag{3}$$

$$= \int_{\mathbb{R}_+} [1 - F_X(t)]^{\delta} \, dt. \tag{4}$$

Note that $\Xi(0; X_p) = \lim_{\delta \to 0} \Xi(\delta; X) = F_X^{-1}(p)$; $\Xi(1; X) = \mu(X)$ while for $\delta = 2$ we obtain the SEF associated with the Gini index $G(.)$ i.e. $\Xi(2; X) = \mu(X) \cdot [1 - G(X)]$. That is

$$G(X) := 1 - \frac{\Xi(2; X)}{\mu(X)} = \int_0^1 (2p-1) \cdot \frac{F_X^{-1}(p)}{\mu(X)} dp.$$

We investigate the relationships between stochastic orders $\geqslant_{\mathcal{V}}$ and inverse stochastic dominance $\succcurlyeq_{ISD[i]}$.

We focus on "nested" classes of weighting functions $v : [0,1] \to \mathbb{R}_+$ integrating to 1, where

$$\mathcal{V}_1 := \{v \in \mathcal{L}_1([0,1]) : v \geq 0, \text{ and } \int_0^1 v(t)dt = 1\}$$

$$\mathcal{V}_2 := \{v \in \mathcal{V}_1 : v \text{ non-increasing}\}$$

$$\mathcal{V}_3 := \{v \in \mathcal{V}_2 : v \text{ convex}\}.$$

Definition 2.11 (Truncated income profiles X_p) *X_p denotes the income profile X truncated at quantile p such that $F_{X_p}^{-1}(t) = F_X^{-1}(t \cdot p)$.*

Then $\mu(X_p)$ is the incomplete mean of distribution X evaluated for the poorest p fraction of individuals.

Theorem 2.4 (Maccheroni, Muliere, Zoli (2005)) *Let $k \in \{1, 2, 3\}$. The following statements are equivalent:*
(i) $\Xi(k-1; X_p) \geq \Xi(k-1; Y_p)$ for all $p \in [0,1]$
(ii) $X \succcurlyeq_{ISD[k]} Y$.

Inverse stochastic dominance of order n is equivalent to dominance for S-Gini indices where $\delta = n - 1$ for distributions truncated at any p.

Remark 2.9 (Zoli (1999)) *An interesting case:* $k = 3$;
$X \succcurlyeq_{ISD[3]} Y \Longleftrightarrow \mu(X_p)[1 - G(X_p)] \geq \mu(Y_p)[1 - G(Y_p)]$ *for all* $p \in [0, 1]$.
Note that if $\mu(X) = \mu(Y)$ *then* $X \succcurlyeq_{ISD[3]} Y \Longrightarrow G(X) \leq G(Y)$.

As a result we obtain as derived in Zoli (1999), Wang and Young (1998), and Aaberge (2004)

Theorem 2.5 *Let* $k = 1, 2, 3$ *the following statements are equivalent:*
(i) $X \geqslant_{\mathcal{V}_k} Y$
(ii) $X \succcurlyeq_{ISD[k]} Y$ [*and* $\mu(X) \geq \mu(Y)$ *for* $k = 3$].

Concerning the role of the Gini index it is also worth to point out that

Theorem 2.6 (Zoli (1999)) *If (a)* $\Delta^2(p)$ *changes sign once, (b)* $\mu(X) = \mu(Y)$ *and (c)* X *dominates* Y *in terms of Leximin, then* $X \succcurlyeq_{ISD[3]} Y \Longleftrightarrow G(X) \leq G(Y)$.

2.3 Relation with more general results on unidimensional inequality and welfare

We start with inequality indices that satisfy the following conditions:

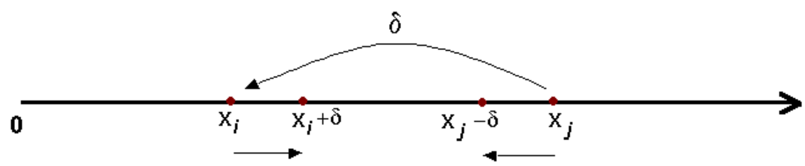
- $I(\mathbf{x})$ is continuous in x_i : small changes in incomes do not lead to big changes in the value of the index.

- $I(\mathbf{x})$ is normalized: that is $I(\mu, \mu, .., \mu) = 0$.

2.3.1 Basic properties:

Axiom 2.1 (Symmetry (S)) $I(\mathbf{x})$ *is invariant with respect to permutation of the incomes.*

Axiom 2.2 (Pigou-Dalton Principle of Transfers (PT)) *A transfer (of* $\delta > 0$*) from a rich person* (j) *to a poor person* (i) : $x_j > x_i$, *which leaves their relative positions unchanged* $(x_j - \delta > x_i + \delta)$ *reduces inequality:* $I(\mathbf{y}) < I(\mathbf{x})$ *if* $x_k = y_k$ *for all* $k \neq i, j$, $x_j > x_i$, $y_i = x_i + \delta$, $y_j = x_j - \delta$.



Progressive Transfer

Such a transfer is called: PROGRESSIVE TRANSFER (the inverse type of transfer, i.e. going in the opposite direction is called *Regressive Transfer*)

Axiom 2.3 (Relative Inequality (Rel)) *Often called also Scale Invariance:* $I(\mathbf{x}) = I(\lambda\mathbf{x})$ *for* $\lambda > 0$.

Theorem 2.7 (Hardy, Littlewood & Polya 1934 (HL&P)) *Consider a fixed number of individuals* n, *let* $\mu(\mathbf{x}) = \mu(\mathbf{y})$, *the following statements are equivalent:*

(1) For all $k \leq n$, $\sum_{i=1}^{k} \hat{x}_i \geq \sum_{i=1}^{k} \hat{y}_i$ *with at least one strict inequality (>).*

(2) $\hat{\mathbf{x}}$ *can be obtained from* $\hat{\mathbf{y}}$ *through a finite sequence of progressive transfers.*

(3) Let $W_u(\mathbf{x}) = \sum_{i=1}^{n} u(x_i)$ *the Utilitarian Social Evaluation Function,* $W_u(\mathbf{x}) > W_u(\mathbf{y})$ *for all* $W_u(\mathbf{x})$ *such that* $u(.)$ *is increasing and strictly concave.*

(4) Let $I_\phi(\mathbf{x}) = \sum_{i=1}^{n} \phi(x_i)$ *the additive inequality index* $I_\phi(\mathbf{x}) < I_\phi(\mathbf{y})$ *for all* $I_\phi(\mathbf{x})$ *such that* $\phi(.)$ *is strictly convex.*

Direct relation of HL&P theorem with results in term of inequality indices.

Theorem 2.8 (Dasgupta, Sen and Starret 1973) *Let distributions* $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ *and* $\mu(\mathbf{x}) = \mu(\mathbf{y})$, *the following statements are equivalent:*
1.a) $I(\mathbf{x}) < I(\mathbf{y})$ *for all inequality indices* $I : \mathbb{R}_+^n \to \mathbb{R}$ *satisfying Symmetry and Principle of Transfers (strictly S-Convex indices)*
2.a) $\mathbf{x} \succ_L \mathbf{y}$

Extension to relative inequality comparisons of distributions with different mean incomes.

Theorem 2.9 (Foster 1985) *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ *the following statements are equivalent:*
1) $I(\mathbf{x}) < I(\mathbf{y})$ *for all inequality* ***Relative*** *indices* $I : \mathbb{R}_+^n \to \mathbb{R}$ *satisfying Symmetry and Principle of Transfers*
2) $\mathbf{x} \succ_L \mathbf{y}$.

2.3.2 Links between Inequality & Welfare

We consider Social Evaluation Function (SEF) $W(\mathbf{x})$

Definition 2.12 (Social Evaluation Function (SEF)) $W : \mathbb{R}_+^n \to \mathbb{R}$ *is a SEF if for every income distribution in* $\mathbf{X}$ *provides a* ***welfare*** *evaluation.*

SEFs differ from Social Welfare Functions because they provide a welfare evaluation based on individuals' income and not directly on individuals' utility or well-being.

Axiom 2.4 (Inequality - Welfare Consistency (IWC)) *If* $\mu(\mathbf{x}) = \mu(\mathbf{y})$ *then for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$

$$I(\mathbf{x}) \leq I(\mathbf{y}) \Leftrightarrow W(\mathbf{x}) \geq W(\mathbf{y}).$$

However the SEF should also be increasing in each individual income. As a result

Theorem 2.10 (Shorrocks (1983); Kolm (1969)) *Let* $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^n$ *the following statements are equivalent:*

(i) $\mathbf{x} \succ_{GL} \mathbf{y}$

(ii) $W(\mathbf{x}) > W(\mathbf{y})$ *for all increasing SEFs* $W(\mathbf{x})$ *satisfying Symmetry,and Principle of Transfers.*

(iii) $\frac{1}{n}\sum_{i=1}^{n} u(x_i) > \frac{1}{n}\sum_{i=1}^{n} u(y_i)$ *for all Average Utilitarian SEFs where* $u(.)$ *is increasing and strictly concave.*

2.4 Poverty Evaluations

Note that for a given poverty line $z > 0$ poverty dominance conditions can be specified as integral stochastic orders, or a rank-dependent stochastic orders, respectively when the family of indices

$$P_p(X, z) = \int_0^z p(x, z) dF_X \tag{5}$$

or

$$P_v(X, z) = \int_0^1 v(p) \cdot \left[z - F_X^{-1}(p)\right]_+ dp \tag{6}$$

are considered. The associated poverty stochastic orders $\geqslant^P$can be specified as:

$$\begin{aligned} X &\geqslant {}_{\mathcal{U}}^{P} Y \Leftrightarrow P_p(X, z) \leq P_p(Y, z) \ \ \forall - p(x, z) = u_z(x) \in \mathcal{U} \\ &\Leftrightarrow \int_0^z u_z(x) dF_X \geq \int_0^z u_z(x) dF_Y \ \ \forall u_z \in \mathcal{U} \\ &\Leftrightarrow \int_0^z u_z(x) d\Delta F \geq 0 \ \forall u_z \in \mathcal{U} \end{aligned} \tag{7}$$

Remark 2.10 *Note that:*

(i) $X \geqslant_{\mathcal{U}^1}^{P} Y \Longleftrightarrow X \succcurlyeq_{SD[1]} Y$ *on* $[0, z]$

(ii) $X \geqslant_{\mathcal{U}^2}^{P} Y \Longleftrightarrow X \succcurlyeq_{SD[2]} Y$ *on* $[0, z]$ *and* $P_0(X,z) = H(X, z) \leq H(Y, z) = P_0(Y,z)$

While for rank-dependent stochastic orders we have for $[t]_+ := \max\{t, 0\}$

$$\begin{aligned} X &\geqslant {}_{\mathcal{V}}^{P} Y \Leftrightarrow P_v(X, z) \leq P_v(Y, z) \ \ \forall v \in \mathcal{V} \\ &\Leftrightarrow \int_0^1 v(p) \cdot \left[z - F_X^{-1}(p)\right]_+ dp \leq \int_0^1 v(p) \cdot \left[z - F_Y^{-1}(p)\right]_+ dp \ \ \forall v \in \mathcal{V} \\ &\Leftrightarrow \int_0^1 v(p) \cdot \Delta_z(p) dp \geq 0 \ \forall v \in \mathcal{V} \end{aligned} \tag{8}$$

where $\Delta_z(p) = \left[z - F_Y^{-1}(p)\right]_+ - \left[z - F_X^{-1}(p)\right]_+$.

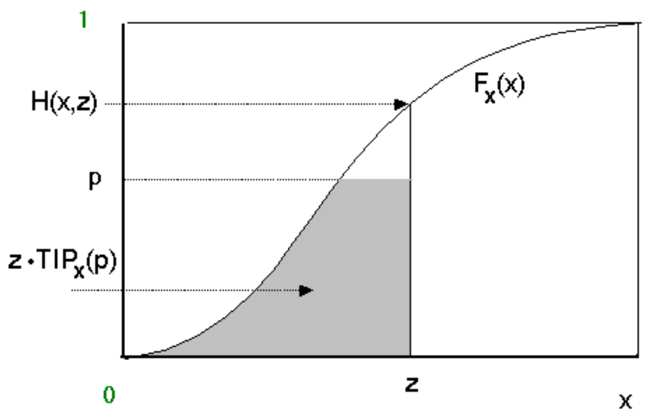
Remark 2.11 *Note that:*
(i) $X \geqslant^P_{\mathcal{V}_1} Y \Longleftrightarrow \left[z - F_Y^{-1}(p)\right]_+ \geq \left[z - F_X^{-1}(p)\right]_+ \;\forall p \in [0,1]$
(ii) $X \geqslant^P_{\mathcal{V}_2} Y \Longleftrightarrow \int_0^p \left[z - F_Y^{-1}(p)\right]_+ dp \geq \int_0^p \left[z - F_X^{-1}(p)\right]_+ dp \;\; \forall p \in [0,1]$

Note that the integral condition in (ii) denotes dominance according to the absolute version (i.e. multiplied by the value of the poverty line $z > 0$) *poverty deprivation curve* or absolute *TIP curve* (TIP stands for *Three I's of Poverty*, because it captures the Incidence, Intensity and Inequality aspects of the poverty evaluation) derived in Spencer and Fisher (1992), Jenkins and Lambert (1997) and Shorrocks (1998).

Definition 2.13 (TIP curve)

$$TIP_X(p,z) := \frac{1}{z}\int_0^p \left[z - F_X^{-1}(t)\right]_+ dt.$$

The TIP curve plots *relative income gaps* $\left[z - F_X^{-1}(p)\right]_+/z$ that are ordered from largest to smallest cumulatively against the population share (p).



Geometric derivation of TIP Curve

The TIP curve represents incidence, inequality and intensity:

- *Incidence:* the population share from which the curve becomes flat is the headcount ratio $H(X,z)$ (for continuous distributions) i.e. the proportion of population with income at most equal to z. The further in the p space the curve become flat the larger is the proportion of poor individuals within the society.

- *Intensity:* the maximum height of the curve represents the poverty gap ratio $H(X,z) \cdot I(X,z) = \frac{1}{z}\int_0^1 \left[z - F_X^{-1}(p)\right]_+ dp$. The higher is the curve the larger is

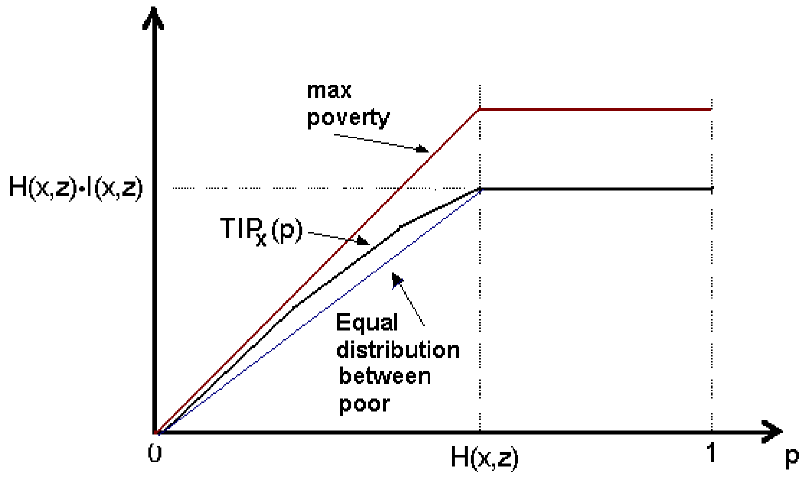


Figure 1: TIP curve

the average relative income (poverty) gap per person in the entire population. It coincides with the relative income gap perceived by any individual in a society where all individuals are poor and identical.

- *Inequality:* the curvature of the curve between the origin and the headcount ratio - summarizes the rate at which the gap decreases as income rises. The "higher" is the curvature the larger is the inequality between poor individuals.

The maximum poverty consistent with a fixed $H(\mathbf{x},z)$ is obtained when each poor individual has income 0 i.e. a poverty gap of z. In this case the TIP curve is a line of 45 degrees till $H(\mathbf{x},z)$, then it is flat. Absence of poverty leads to a TIP curve that coincides with the horizontal axis. The overall maximum poverty coincides with a 45 degree line for $p \in [0,1]$.

3 Multidimensional Case

3.1 Consistency in Aggregation

Rubinstein, Fishburn (JET 1986): Algebraic aggregation theory; Dutta, Pattanaik, Xu (Economica 2003): On Measuring Deprivation and the Standard of Living in a Multidimensional Framework on the Basis of Aggregate Data; Gajdos, Maurin (JET 2004): Unequal uncertainties of uncertain inequalities: an axiomatic approach

- **Question:** Is it possible to obtain consistent ranking across matrices aggregating......

1. first for each attribute taking the distribution across agents and then aggregating the summary Macro result across attributes (Procedure 1) (e.g. HDI)

2. for each agent obtaining an individual index of personal well being and then aggregating the distribution of these indices for all the population (Procedure 2) (e.g. additively decomposable SWFs over multiattribute distributions)......

3.1.1 Consistent iterative aggregation

.... moreover one would like to apply the same aggregator in each procedure when aggregating across individuals and another on when aggregating across attributes for instance for two attributes we have

$$X = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$$

Definition 3.1 *Procedure 1: first columns then rows*

$$\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix} \begin{matrix} \Downarrow \\ Aggr.\ \psi \end{matrix}$$
$$\downarrow$$
$$\underset{\Longrightarrow\ Aggregator\ \phi}{\begin{bmatrix} \psi(x_{11}; x_{21}) & \psi(x_{12}; x_{22}) \end{bmatrix}} \longmapsto \phi[\psi(x_{11}; x_{21}); \psi(x_{12}; x_{22})]$$

Definition 3.2 *Procedure 2: first rows then columns*

$$\overset{\Longrightarrow\ Aggregator\ \phi}{\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}} \longmapsto \begin{bmatrix} \phi(x_{11}; x_{12}) \\ \phi(x_{21}; x_{22}) \end{bmatrix} \begin{matrix} \Downarrow \\ Aggr.\ \psi \end{matrix}$$
$$\downarrow$$
$$\psi[\phi(x_{11}; x_{12}); \phi(x_{21}; x_{22})]$$

3.1.2 A result by Dutta et al. (2003):

Assumptions on ϕ and ψ:

Definition 3.3 (Consistency) $\phi \circ \psi(X) \geq \phi \circ \psi(Y)$ *iff* $\psi \circ \phi(X) \geq \psi \circ \phi(Y)$

moreover suppose that $x_{ij} \in [c_{\min}; c^{\max}]$ and

- $\phi : [c_{\min}; c^{\max}]^d \rightarrow [0, 1]$; $\phi(\mathbf{1}c_{\min}) = 0$; $\phi(\mathbf{1}c^{\max}) = 1$
- $\psi : [c_{\min}; c^{\max}]^n \rightarrow [0, 1] : \psi(\mathbf{1}c_{\min}) = 0$; $\psi(\mathbf{1}c^{\max}) = 1$
- ϕ and ψ are continuous and strictly increasing in each argument
- ψ is symmetric across agents

- ϕ exhibit **non increasing increments,** that is

$$\begin{aligned} & \phi(x_1, x_2, ., x_h, ..., x_k + t, ..x_n) - \phi(x_1, x_2, ., x_h..., x_k, ..x_n) \\ \leq \ & \phi(x_1, x_2, ., x_h - \tau, ..., x_k + t, ..x_n) - \phi(x_1, x_2, ., x_h - \tau, ..., x_k, ..x_n) \end{aligned}$$

for $\tau, t > 0$.

Theorem 3.1 *Given the assumptions on ϕ and ψ, the two procedures are consistent iff*

$$\begin{aligned} \phi(x_{i.}) &= \frac{\sum_{j=1}^{d} w_j \cdot x_{ij} - c_{\min}}{c^{\max} - c_{\min}}; \text{ where } w_j > 0, \ \sum_{j=1}^{d} w_j = 1 \\ \psi(x_{.j}) &= \frac{\frac{1}{n}\sum_{i=1}^{n} x_{ij} - c_{\min}}{c^{\max} - c_{\min}} \end{aligned}$$

We obtain essentially HDI types of indices that can be constructed as weighted averages of normalized attributes evaluated across the whole population.

Correlation between attributes is lost indeed $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ are considered socially indifferent.

Remark 3.1 *The assumption of non increasing increments (NII) per each agent across attributes is crucial.*

For instance here is an example without NII.

Example 3.1 *Check that for given increasing functions $f_j : [c_{\min}; c^{\max}] \rightarrow [0, 1] : f(c_{\min}) = 0;\ f(c^{\max}) = 1$ and the increasing function $H : [0, 1] \rightarrow [0, 1]$ the following functional forms satisfy consistency and all other properties but not necessarily NII*

$$\begin{aligned} \phi(x_{i.}) &= H^{-1}(\sum_{j=1}^{d} w_j \cdot H[f_j(x_{ij})]); \text{ where } w_j > 0, \ \sum_{j=1}^{d} w_j = 1 \\ \psi(x_{.j}) &= H^{-1}(\frac{1}{n}\sum_{i=1}^{n} H[f_j(x_{ij})]). \end{aligned}$$

Property NII imposes linearity in H and f_j.

Foster et. al. (JHD 2005) consider the case where H is isoelastic i.e. $H(t) = t^{1-\varepsilon}/(1-\varepsilon)$ for $\varepsilon \geq 0$ and $w_j = 1/d$.

If we let $s_{ij} = f_j(x_{ij})$ the index "consistent in aggregation" is obtained for

$$\begin{aligned}\phi(x_{i.}) &= \left(\frac{1}{d}\sum_{j=1}^{d}[s_{ij}]^{1-\varepsilon}\right)^{1/(1-\varepsilon)} \quad ; \ \psi(x_{.j}) = \left(\frac{1}{n}\sum_{i=1}^{n}[s_{ij}]^{1-\varepsilon}\right)^{1/(1-\varepsilon)} \\ W(X) &= \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{d}\sum_{j=1}^{d}[s_{ij}]^{1-\varepsilon}\right)^{1/(1-\varepsilon)} ,\end{aligned}$$

then H satisfies NII in terms of the distribution of y iff $\varepsilon = 0$.

- We use the notation $W(.)$ for the final aggregator because our main focus in this section is for a well-being evaluations. That is multidimensional welfare indices.

Main positive features of the index is that it is Subgroup Consistent i.e. $W(X,Y) \geq W(X,Z)$ iff $W(Y) \geq W(Z)$. Where (X,Y) denotes that the population is partitioned into two groups of individuals.

However the index still continues to consider $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ as socially indifferent.

3.1.3 Possible alternative solutions

If any concern for correlation, i.e. dependency between the distributions of the attributes is lost if consistency in aggregation is required then one may drop this assumption and can take into account an average of the results arising from the two procedures.

This has already been suggested in the literature on the measurement of inequality under uncertainty.

Gilboa and Schmeidler (JMathE 1989): Maximin expected utility with non unique prior; Ben Porath et al. (JET 1997): On the measurement of inequality under uncertainty; Gajdos and Maurin (JET 2004): Unequal uncertainties of uncertain inequalities: an axiomatic approach; Gajdos and Weymark (ET 2005): Multidimensional Generalized Gini indices

We here review some related considerations formulating first the problem then considering additive aggregation methods and finally moving to aggregation methods that are non additive and do not satisfy consistency in aggregation.

Suppose for simplicity that we normalize each agent realization in a given space with $s_{ij} := f(x_{ij}) \in [0,1]$ that is a *score function* associated with the realization of agent i on space j. Moreover one can further assume that

- this normalization makes comparable scores of the same agent in different characteristics (Symmetry between characteristics)
- and agents are all treated equally in the final evaluation (Anonymity)

Alternative sequences. Some examples Consider the following distributions:

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} ; X' = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} ;$$

if attributes carry the same weight $W(X) = W(X')$

$$Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} ; Z' = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} ;$$

if agents are equally relevant $W(Z) = W(Z')$

$$Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} ; Y' = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} ;$$

if both previous considerations hold then $W(Y) = W(Y')$

However *linear symmetric aggregation functions* as those required by consistent aggregation value indifferent all distributions.

But this should not be the case in particular comparing Y, Y' with all other matrices.

Problems with consistent additive decomposition The previous considerations *extend also to the additive decomposition of the matrices* as in Foster et al. (2005) or even in the more general result presented earlier on consistent aggregation....

The procedure was including considerations on inequality in the distribution across agents but what is left aside is the correlation between the distributions of the attributes

$$Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} ; Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} ; X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

- It is possible to regain some considerations if we give up the issue of consistency in aggregation...specifying a given order of aggregation and in an additive framework consider different parameters ε in aggregating across distribution w.r.t. those applied in aggregating across individuals. (Decancq, Decoster Schokkaert World Dev 2008)

......but what is left aside even in this case is the correlation between the distributions of the attributes in

$$Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} ; Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} ; X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} .$$

More precisely:

- Depending on the order of aggregation if we start first deriving an individual index of well being symmetric in attributes then $W(Y) = W(X)$.

- If we aggregate first across attributes deriving a Macro index of the distribution across individuals then $W(Y) = W(Z)$ This second result is more controversial and highlight some critical aspects underlying the procedure that first aggregates across attributes. For more general results in this direction see Pattanaik et al. (2008)

The general critical issue in comparing Y and Z is the increase in the correlation between attributes in Z.

From $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ to $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ we have transferred attribute 1 from agent 2 to agent 1 that now clearly dominates the latter.

Definition 3.4 (CIT) *In general a* ***Correlation Increasing Transfer*** $CIT\ (i, j)$ *is a sequence of "rearrangements" of the distribution of attributes (one attribute per step of the sequence) involving only two individuals* (i, j) *s.t. as the result of the process one individual ends up being weakly dominated by the other in any attribute.*

Epstein and Tanny (CanJEc1980), Tsui (JET 1995, SCW 1999)

Example 3.2 *A sequence of CIT*

$$Y' = \begin{bmatrix} 1 & 3 & 2 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} \underset{CIT(13)}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 2 \\ 3 & 2 & 1 \\ 2 & 3 & 3 \end{bmatrix} \underset{CIT(12)}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 1 \\ 3 & 2 & 2 \\ 2 & 3 & 3 \end{bmatrix} \underset{CIT(23)}{\Longrightarrow} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} = Z'$$

The last matrix (Z') is more unfair than the first one (Y') looking from the individuals perspective.

All symmetric indices first aggregating across attributes will consider $W(Y') = W(Z')$

Also inequality indices aggregating first across individuals may lead to $I(Y') > I(Z')$

See Dardanoni (1995) comment on Maasaouni (1986) Multidimensional Inequality Index

Using non additive measures (Gini type welfare/inequality indices) We use a functional [ψ in the previous discussion] in order to aggregate across individuals their realizations in each attribute [aggregating vertically for each matrix column] and another functional [ϕ in the previous discussion] in order to aggregate across attributes per each individual [aggregating horizontally for each matrix row]. In particular we consider Gini welfare evaluation over individuals [$\psi = G$] AND the average aggregator over attributes [$\phi = \mu$]

Definition 3.5 ***First procedure*** *(average of Gini indices of the agents distribution for each attribute): aggregating for each attribute taking the Gini index across agents AND then averaging the obtained result.*

Let $v_1 \geq ... \geq v_i \geq v_{i+1} \geq ... \geq v_n \geq 0$ where $\sum_i v_i = 1$ let $\hat{s}_{(i)j}$ denote the increased order of s_{ij} for each attribute j across agents i.e. $\hat{s}_{(1)j} \leq \hat{s}_{(2)j} \leq ... \leq \hat{s}_{(n)j}$ then

$$G_j(s_{.j}) = \sum_{i=1}^{n} v_i \cdot \hat{s}_{(i)j}$$

averaging across all attributes we get

$$G^1(S) = \frac{1}{d} \sum_{j=1}^{d} \sum_{i=1}^{n} v_i \cdot \hat{s}_{(i)j}.$$

Remark 3.2 *Note that we assume symmetry across attributes thus the weighting function v_i applied in $G_j(s_{.j})$ is independent from j.*

Second procedure with Gini type indices.

Definition 3.6 ***Second procedure*** *(Gini index of average agent score): Average over attributes for each agent AND Gini aggregation over individuals scores.*

Average score of agent i :

$$\mu_i = \mu(s_{i.}) = \frac{1}{d} \sum_{j=1}^{d} s_{ij};$$

let $\hat{\mu}_{[i]}$ denote the increased order of μ_i i.e. $\hat{\mu}_{[1]} \leq \hat{\mu}_{[2]} \leq ... \leq \hat{\mu}_{[n]}$ and let $\beta_1 \geq ... \geq \beta_i \geq \beta_{i+1} \geq ... \geq \beta_n \geq 0$ where $\sum_i \beta_i = 1$ then

$$G^2(S) = \sum_{i=1}^{n} \beta_i \cdot \hat{\mu}_{[i]} = \sum_{i=1}^{n} \beta_i \cdot \left(\frac{1}{d} \sum_{j=1}^{d} \hat{s}_{[i]j} \right) = \frac{1}{d} \sum_{j=1}^{d} \sum_{i=1}^{n} \beta_i \cdot \hat{s}_{[i]j}.$$

Remark 3.3 *Note that the order of $\hat{s}_{(i)j}$ and $\hat{s}_{[i]j}$ does not necessarily coincide. In particular for a given attribute j the elements $\hat{s}_{(i)j}$ are ranked in increasing order in terms of values of s_{ij}, while the elements $\hat{s}_{[i]j}$ are ranked according to the increasing order of the elements μ_i. Thus $\hat{s}_{(i)j} = \hat{s}_{[i]j}$ only if the order of the scoring functions s_{ij} for attribute j is the same as the order of the average scores μ_i.*

- If the rank correlation of each attribute across agent is perfectly positive then $\hat{s}_{(i)j} = \hat{s}_{[i]j}$ for any i and j, thus $G^1(S) = G^2(S)$ if $\beta_i = v_i$.
- However note that in general $G^1(S) = G^2(S)$ only if $\beta_i = v_i = 1/n$.

Remarks on Gini indices Getting back to the original matrices

$$X = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}; Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}; Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};$$

recall that X', Z', Y' are obtained permuting either rows or columns thus they obtain the same evaluation of the symmetric distributions X, Z, Y respectively. Applying the mentioned procedure one gets:

$$\begin{array}{ll} G^1(X) = 1/2 & G^2(Z) = 1 - \beta_1 \\ G^1(Z) = G^1(Y) = 1 - v_1 & G^2(X) = G^2(Y) = 1/2 \end{array}$$

The final index that considers a weighted average of $G^1(.)$ and $G^2(.)$ according to

$$W(S) := \alpha \cdot G^1(S) + (1 - \alpha) \cdot G^2(S)$$

will give:

$$\begin{aligned} W(Y) &= \alpha \cdot (1 - v_1) + (1 - \alpha) \cdot 1/2 \leq 1/2 \\ W(X) &= 1/2 \\ W(Z) &= \alpha \cdot (1 - v_1) + (1 - \alpha) \cdot (1 - \beta_1) \leq 1/2 \end{aligned}$$

recalling that $(1 - v_1) \leq 1/2$ and $(1 - \beta_1) \leq 1/2$. More precisely if $\alpha \in (0, 1)$, $1/2 < v_1$ and $1/2 < \beta_1$ then

$$W(X) > W(Y) > W(Z).$$

Remark 3.4 *Note this mixed procedure is different from Hicks (1997 World Dev.) proposal of taking the average of the Gini welfare index of the distribution of each attribute. (This measure coincides with the one obtained applying only the first procedure)*

Remark 3.5 *Gajdos and Weymark (ET 2005) characterize families of Generalized averages across attributes of Gini indices across individuals per each attribute. (Again in line with the application of the first procedure only)*

However as already pointed out for these measures

$$Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \text{ and } Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

are considered socially indifferent thereby neglecting considerations based on correlation in the distribution of the attributes.

What about applying a Gini type evaluation to the overall matrix? A problem arises due to comonotonic independence (across attributes). Consider matrices

$$H = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}; H' = \begin{bmatrix} 0 & 0 \\ 1/2 & 0 \end{bmatrix}; K = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 0 \end{bmatrix};$$

note that K is comonotonic w.r.t. H and H' because the ranking of the attributes is the same in all matrices: taken any pair of cells ij and $i'j'$ it is always true that $H_{ij} \geq H_{i'j'} \Longleftrightarrow K_{ij} \geq K_{i'j'}$ (similarly for comparisons of K and H').

- *Comonotonic Independence* is the key property characterizing Gini type (i.e. rank dependent) evaluations!

By Comonotonic independence between K and H and K and H' it follows that for a Gini index G

$$G(H+K) \geq G(H'+K) \Longleftrightarrow G(H) \geq G(H')$$

...but $G(H) = G(H')$ by anonymity (i.e. symmetry across individuals), thus

$$G(H+K) = G(H'+K)$$

where

$$H+K = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 0 \end{bmatrix}; H'+K = \begin{bmatrix} 1/2 & 1/2 \\ 1 & 0 \end{bmatrix}.$$

Note that the general indication for the multivariate Gini index $G(H+K) = G(H'+K)$ is in line with the one obtained following the first procedure of aggregation first across individuals and then averaging across attributes that will lead to $G^1(H+K) = G^1(H'+K)$. Once again *correlation is neglected* indeed $H+K$ can be obtained from $H'+K$ through a Correlation Increasing Transfer in the first attribute thus one would expect that $W(H+K) \leq W(H'+K)$ as it is the case applying the previous mixed two steps procedure.

- **Open question**...appropriate definition of multidimensional Gini functionals...

4 Multidimensional orders

In this section we move from multidimensional indices to dominance conditions.

The key component of interest in the multidimensional framework is still the dependence between attributes

Different tools can be applied:

- Multidimensional Majorization.
- Multidimensional versions of Lorenz and Generalized Lorenz curves and related dominance conditions.
- Multidimensional Stochastic dominance conditions based on stochastic orders specified in the unidimensional framework.

4.1 Multidimensional Majorization

Kolm (QJE 1977): Multidimensional egalitarianism; Koshevoy (SCW 1995): Multivariate Lorenz majorization; Koshevoy, Mosler (JASA 1996): The Lorenz zonoid of a multivariate distribution; Koshevoy, Mosler (AStA 2007): Multivariate Lorenz dominance based on zonoids; Marshall, A. W. and Olkin, I. (1979): Inequalities: Theory of Majorization and Its Applications. New York: Academic Press; Weymark (2004) The normative approach to the measurement of multidimensional inequality; Savaglio (2004) Multidimensional inequality: a survey

The *Multidimensional Majorization* criteria generalizes the majorization conditions based on the application of bistochastic matrices moving the perspective from the unidimensional case to multidimensional distributions where attributes are transferable between individuals (e.g. bundles of goods) see Marshall and Olkin (1979).

Consider matrices $X, Y \in \mathbb{R}^{n\times d}$ for n individuals, d goods (characteristics) and such that $\sum_{i=1}^{n} \mathbf{x}_i = \sum_{i=1}^{n} \mathbf{y}_i$ where $\mathbf{x}_i, \mathbf{y}_i \in \mathbb{R}^d$ denote vectors of goods belonging to individual i.

- Distribution of a fixed bundle of d goods across n individuals

Definition 4.1 *Y multidimensionally majorizes X (in the weak sense)*

$$Y >_M X \Longleftrightarrow X = \Pi Y,$$

where Π is a $n \times n$ bistochastic matrix [all elements of Π are non-negative with row sum and column sum equal to 1].

X is obtained from Y averaging the endowments vectors of the individuals, or in other terms X shows less disparity in the distributions of the bundles of goods than Y.

Example 4.1 $\Pi = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.6 & 0.2 \\ 0.7 & 0.2 & 0.1 \end{bmatrix}; Y = \begin{bmatrix} 10 & 30 \\ 20 & 30 \\ 10 & 0 \end{bmatrix} \Longrightarrow X = \Pi Y = \begin{bmatrix} 12 & 9 \\ 16 & 24 \\ 12 & 27 \end{bmatrix}$

In the example the individual 3 situation [third row] in X is substantially improved (she has overtaken individual 1) if compared to Y. This fact may make difficult to agree on a reduction of dispersion in X w.r.t. Y but we need to rely on symmetric evaluations across individuals (note that a permutation matrix is bistochastic) therefore one may find more immediate to compare $Y = \begin{bmatrix} 10 & 30 \\ 20 & 30 \\ 10 & 0 \end{bmatrix}$ and $X' = \begin{bmatrix} 12 & 27 \\ 16 & 24 \\ 12 & 9 \end{bmatrix}$.
Not only for each attribute the distribution is less disperse but also disparities between individuals bundles seem to be reduced.

Note however that *all attributes are "mixed" in the same way for each individual* (i.e. they are multiplied by the same raw of Π) i.e. $\mathbf{x}_i = \sum_{k=1}^{n} \pi_{ik} \cdot \mathbf{y}_k$ thus the averaging of every attribute is made using the same weights depending on the individuals and not on the attribute itself. This operation reduces the disparities between the individuals bundles.

A welfare interpretation Consider matrices $X, Y \in \mathbb{R}^{n\times d}$ such that $\sum_{i=1}^n \mathbf{x}_i = \sum_{i=1}^n \mathbf{y}_i$

Theorem 4.1 *$Y >_M X$ is equivalent to the following conditions*

(I) $\phi(Y) \leq \phi(X)$ for all $\phi : \mathbb{R}^{n\times d} \rightarrow \mathbb{R}$ which are S-concave;

(II) $\sum_{i=1}^n u(\mathbf{y}_i) \leq \sum_{i=1}^n u(\mathbf{x}_i)$ for all $u : \mathbb{R}^d \rightarrow \mathbb{R}$ which are concave [they can also be increasing].

S-Concave function: symmetric functions such that $\phi(Y) \leq \phi(\Pi Y)$

The distribution X of a fixed amount of resources improves welfare

Does it means that we have also less inequality in terms of the distribution of the concave and increasing utilities $u(\mathbf{y}_i)$?

Will it be possible to decompose the change from Y to X in terms of progressive transfers?

A controversial implication in terms of inequality Dardanoni (REI 1996) On multidimensional inequality measurement

Example 4.2 $\Pi = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix}; Y = \begin{bmatrix} 1 & 1 \\ 1 & 3 \\ 3 & 1 \end{bmatrix} \Longrightarrow X = \Pi Y = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 2 & 2 \end{bmatrix}$

The utility of the poorest individual [agent 1] is left unchanged, while according to any strictly concave utility agents 2 and 3 are better off.

There is more inequality as well as welfare even though the resources are fixed!!! This outcome contradicts the link between welfare and inequality evaluations in the unidimensional framework. For unidimensional distributions with fixed amount of attributes (e.g. fixed income) an increase in inequality reduces welfare.

- If we consider majorization dominance as an ethically compelling criterion then in evaluating inequality the approach that first aggregates in terms of individual utilities might not be appropriate.

4.1.1 Relation with Progressive (Pigou Dalton) transfers

Relation with Progressive (Pigou Dalton) transfers

Will it be possible to decompose the change from Y to X in terms of progressive transfers?

Definition 4.2 *T transform (Pigou Dalton transfer)*

$$\begin{aligned} \Pi_{2,3}(\lambda) &= \lambda \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (1-\lambda) \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda & (1-\lambda) \\ 0 & (1-\lambda) & \lambda \end{bmatrix} \end{aligned}$$

a convex combination of identity matrix and a permutation matrix involving a permutation of 2 individuals.

These transformations are those underlying the Pigou Dalton progressive transfer in the unidimensional case.

Will it be possible to decompose the change from Y to X in terms of progressive transfers?

Remark 4.1 *In general when $n \geq 3$ not all bistochastic matrices Π can be obtained as product of T transforms.*

This issue is particularly crucial when $d \geq 2$.

Example 4.3 *The matrix* $\Pi = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix}$*, cannot be obtained as product of T-transforms.*

The 3 entries with 0 involving all individuals cannot be replicated by a chain of T-transforms different from permutation matrices in which case we won't be able to obtain the 0.5 entries

Remark 4.2 *Note however that in the unidimensional case the above mentioned problem is not an issue because there exist always the possibility to obtain the final distribution through T transforms even though they generate a different bistochastic matrix. We check this statement in next example.*

Example 4.4 $\begin{bmatrix} 0.5 & 0 & 0.5 \\ 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix} \begin{bmatrix} 30 \\ 30 \\ 0 \end{bmatrix} = \begin{bmatrix} 15 \\ 15 \\ 30 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 30 \\ 30 \\ 0 \end{bmatrix}$ *thus*

$$\begin{bmatrix} 15 \\ 15 \\ 30 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} 30 \\ 30 \\ 0 \end{bmatrix}.$$

The above two matrices are T transforms $\Pi_{i,j}(\lambda)$: the first is a permutation matrix an coincides with $\Pi_{1,3}(0)$ while the second is $\Pi_{2,3}(0.5)$. The example clarifies that even though the transition from $\begin{bmatrix} 30 \\ 30 \\ 0 \end{bmatrix}$ *to* $\begin{bmatrix} 15 \\ 15 \\ 30 \end{bmatrix}$ *can be formalized through a bistochastic matrix that cannot be decomposed into the product of T transforms, it can also be obtained through the product of appropriately derived T transforms.*

With *two [or more] dimensions* it is essential to be able to generate the specific bistochastic matrix considered.

A further weakness Multidimensional majorization is of no help in ranking matrices capturing the notion of increase in correlation in the distributions of the attributes.

Recall the initial examples where $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$; these two distributions cannot be ranked in terms of majorization. Even if one would argue that Y is preferable to Z bot in terms of less inequality than in terms of higher welfare/well being *there is no bistochastic matrix* Π s.t. $Y = \Pi Z$. Indeed one can get that all matrices Z' that (weakly) majorize Z are:

$$Z' = \begin{bmatrix} \lambda & (1-\lambda) \\ (1-\lambda) & \lambda \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \lambda & \lambda \\ (1-\lambda) & (1-\lambda) \end{bmatrix}$$

for $\lambda \in [0, 1]$.

4.1.2 Tests for Multidimensional Majorization...

- Unfortunately SO FAR there exists no multidimensional generalization of the Lorenz curve which could be used to rank matrices according to the $>_M$ preorder.

Standard majorization could be weakened in different ways in order to be applied in different economic contexts. The most interesting device is the *price majorization* criterion suggested in Kolm (QJE 1977).

Definition 4.3 *The matrix Y is said to price majorize X that is*

$$Y >_P X \Longleftrightarrow Yp >_M Xp \ \forall p \in \mathbb{R}^d_+, \ (\textit{or } \forall p \in \mathbb{R}^d)$$

i.e. the distribution of potential incomes associated to X, and evaluated according to the vector of prices p, Lorenz dominates the one associated to Y for all possible price profiles [they can be also negative].

Price Majorization... According to $>_P$ the distribution of initial endowments X is always preferred to Y by an inequality averse policy maker which is concerned in maximizing the distribution of indirect utilities and attach to each individual the same direct utility function, this evaluation is valid no matter what will be the equilibrium price profile.

$$Y >_M X \Leftrightarrow X = \Pi Y \Longrightarrow Xp = \Pi Y p \Leftrightarrow Yp >_M Xp \Leftrightarrow Y >_P X$$

- thus $>_M \Longrightarrow >_P$ but the converse is not always true.

There exist multidimensional generalizations of the Lorenz curve that can be used to test $>_P$both when $p \in R^d_+$, and $p \in R^d$. They also work as analogous of generalized Lorenz dominance over distributions of income budgets. The are the *Lorenz Zonoid, the Lift Zonoid and their extensions*!!!

Before moving to the analysis of the construction of these dominance tests notice that the price majorization condition can be useful to rank $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Consider $\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} = p \in R^2$ then $Zp = \begin{bmatrix} p_1 + p_2 \\ 0 \end{bmatrix}$ and $Yp = \begin{bmatrix} p_2 \\ p_1 \end{bmatrix}$ then if either $p \in R^2_{++}$ or $p \in R^2_{--}$we can show that

$$\begin{aligned} Yp &= \begin{bmatrix} p_2 \\ p_1 \end{bmatrix} = \begin{bmatrix} \lambda & (1-\lambda) \\ (1-\lambda) & \lambda \end{bmatrix} \begin{bmatrix} p_1 + p_2 \\ 0 \end{bmatrix} = Zp \\ \text{when } \lambda &= \frac{p_2}{p_1 + p_2} \in (0,1) \end{aligned}$$

thus $Zp >_M Yp$.

On the other hand if for instance $p_1 < 0 < p_2$ is negative then either λ or $(1-\lambda)$ would be negative and thus the matrix above cannot be consider a bistochastic matrix anymore. As a result price majorization does not hold.

- *Price majorization with positive prices* appears an interesting candidate for a meaningful multidimensional dominance condition.

4.1.3 Lift Zonoid and Lorenz Zonoid (by Koshevoy & Mosler)

For empirical distributions the definition of the lift/Lorenz zonoid is the following:

Definition 4.4 *The Lift Zonoid $Z(X)$ is a convex compact set in the $(d+1)$ space obtained as the weighted sum of segments $\mathbf{x}_i \in \mathbb{R}^d$, for all possible sets of normalized weights, that is*

$$Z(X) = \left\{ \sum_{i=1}^{n} z_{0i}; \sum_{i=1}^{n} z_{0i}\mathbf{x}_i : 0 \le z_{0i} \le 1/n, \ i = 1, 2, ..., n \right\}.$$

The Lorenz Zonoid $LZ(X)$ is the Lift Zonoid evaluated over distribution

$$\tilde{X} := (\mathbf{x}_{.1}/\mu(\mathbf{x}_{.1}); \mathbf{x}_{.2}/\mu(\mathbf{x}_{.2}); ...; \mathbf{x}_{.d}/\mu(\mathbf{x}_{.d}))$$

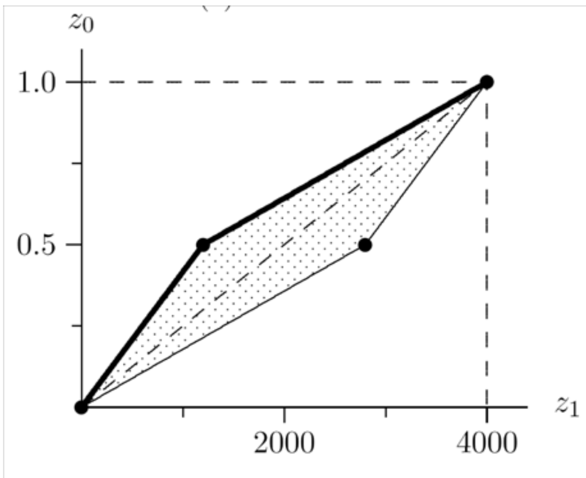
where each attribute is normalized dividing it by its average: $LZ(X) := Z(\tilde{X})$.

Zoonoids: the intuition

- Take all subsets of the population of a given relative size $z_0 \in [0,1]$ (e.g. 50%) [where $z_0 = \Sigma_i z_{0i}$] what is the aggregate (divided by n) realization in the d dimensional space of the resources of any of these subsets covering a z_0 proportion of population? Take the convex hull of all these distributions in the d dimensional space. We have obtained the **section of the Lift Zonoid for a fixed value of population share** z_0. As z_0 moves from 0 to 1 we construct the Lift Zonoid. For finite populations it is necessary to "convexify" all sections corresponding to adjacent proportions of population.

- In order to get the Lorenz Zonoid we need just to apply the same logic to the normalized distributions of each attribute. So normalize each columns in relative terms so that any aggregate amount sums to n.

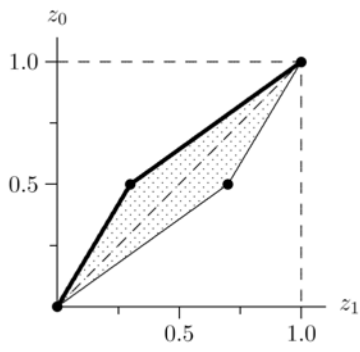
Example 4.5 *Lift zonoid of univariate distribution (2400,5600), example taken from Koshevoy and Mosler (AStA2007) see figure*



(b) Lift zonoid

Note that when considering the 50% of the population the amount of resources (divided by 2) owned by a group of that size ranges from 2400/2 to 5600/2. With a population of n *agents equally split between 50% of poors endowed with 2400 and 50% of riches endowed with 5600 then any group of* $n/2$ *agents will own average resources (divided by* n*) for* $\frac{n_p}{n} \cdot 2400 + \frac{n_r}{n} \cdot 5600$ *where* $n_p + n_r = n/2$ *and* $n_p \leq n/2$; $n_r \leq n/2$*. Once* n_p *and* n_r *move from* 0 *to* $n/2$ *we obtain realizations that range between* 1200 *and* 2800*.*

Example 4.6 *Lorenz zonoid of univariate distribution (2400,5600) [in relative terms (0.6; 1.4)] taken from Koshevoy and Mosler (AStA2007) see figure*



(a) Lorenz zonoid

- Dominance is defined in terms of *inclusion of Zonoids.*

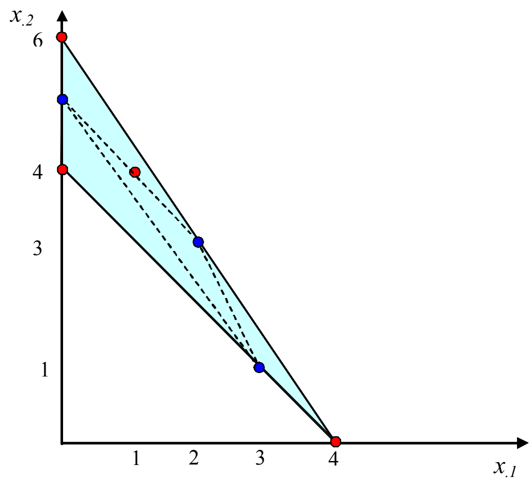
Example 4.7 *For*

$$Y = \begin{bmatrix} 0 & 6 \\ 0 & 4 \\ 4 & 0 \\ 1 & 4 \end{bmatrix} \textit{ and } X = \begin{bmatrix} 0 & 5 \\ 0 & 5 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}$$

note that $Z(X) \subseteq Z(Y)$.

Moreover note that we have (Generalized) Lorenz dominance for each attribute!!

A graphical representation for $Z(X) \subseteq Z(Y)$ evaluated for $z_0 = 1/4$ shows that for $z_0 = 1/4$ i.e. when we consider the realizations of each agent, the convex hull of these realizations for distribution X (identified by the blue dots and the dotted lines) is included in the convex hull of the realizations of Y (identified by the red dots and the continuous lines).



$Z(X) \subseteq Z(Y)$ for $z_0 = 1/4$

Clearly inclusion of the convex hull of the average realization of groups of a given size should also be checked for $z_0 = 2/4$, and $z_0 = 3/4$ (for $z_0 = 1$ the zonoids coincide given that the aggregate total amount of resources is the same in X and Y). Further calculations will show that these conditions hold thus $Z(X) \subseteq Z(Y)$.

Remark 4.3 *There is no (4×4) bistochastic matrix Π such that $X = \Pi Y$ for X and Y presented in the previous example.*

In order to accommodate for the transformation from Y to X involving the first two individuals the only admissible matrix should be

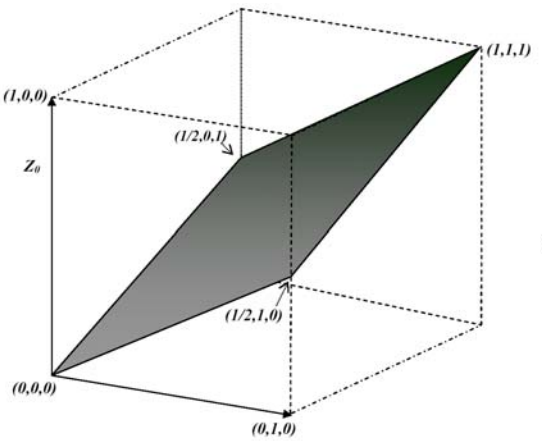
$$\Pi = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & c & d \end{bmatrix}.$$

If we set $a = 1/3; b = 2/3$ in order to accommodate for the first attribute of the third individual we cannot obtain the distribution in X of her second attribute!!!

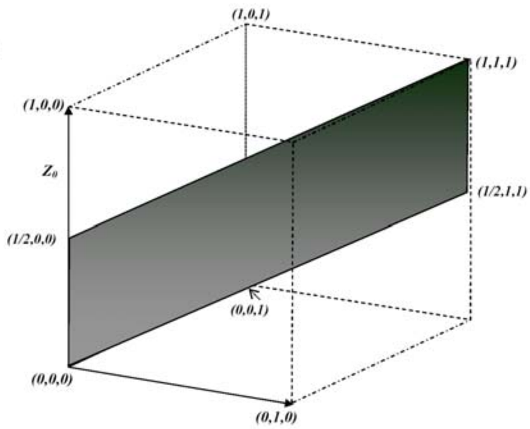
This remark clarifies that

$$Z(X) \subseteq Z(Y) \not\Longrightarrow Y >_M X.$$

A further example: Zonoids fail in taking into account the effect of correlation increasing transformations If one considers $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ there is no inclusion relation between the Lift Zonoids of the two distributions. In order to check it is sufficient to compare the graphical 3-dimensional representation of the Zonoids of the two distributions (for convenience of exposition without loss of generality we haven't normalized the realizations of each group).



Lift Zonoid of Y



Lift Zonoid of Z

This problem is the logical counterpart of the lack of comparability of the two matrices in terms of price majorization when also negative prices are taken into account as discussed earlier.

Lorenz Dominance and Price Majorization The next theorem clarifies the connection between Zonoids dominance and price majorization with positive and negative prices

Theorem 4.2 *The following conditions are equivalent:*
(I) $LZ(X) \subseteq LZ(Y)$.
(Ia) $LZ(X_S) \subseteq LZ(Y_S)$ *for the distributions of all the subsets* S *of attributes*
(II) $\tilde{Y} >_P \tilde{X}$ *(for all* $p \in \mathbb{R}^d$*).*
(III) $\tilde{X}p$ *Generalize Lorenz dominates* $\tilde{Y}p$ *for all* $p \in \mathbb{R}^d$*. (Budget dominance)*
(IV) $\psi(\tilde{Y}p) \leq \phi(\tilde{X}p)$ *for all* $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ *which are S-concave and all* $p \in \mathbb{R}^d$*.*
(V) $\sum_{i=1}^n v(\tilde{\mathbf{y}}_i \cdot p) \leq \sum_{i=1}^n v(\tilde{\mathbf{x}}_i \cdot p)$ *for all* $v : \mathbb{R} \rightarrow \mathbb{R}$ *which are concave and all* $p \in \mathbb{R}^d$*.*

- Is it possible to obtain conditions analogous to Generalized Lorenz dominance of budgets when prices are only positive?

Budget dominance with positive prices The dominance tool requires to *extend the notion of Lift Zonoid.*

Definition 4.5 *The* ***Extended Lift Zonoid*** $eZ(X)$ *is obtained extending the volume of the Lift Zonoid taking all points below it for the coordinate relative to the population share and all points above the Lift Zonoid in the d dimensional space of the attributes. Thus for instance any two dimensional section of* $Z(X)$ *for a given population share* z_0 *the extension requires to take all points north-east w.r.t. each point in* $Z(X)$ *associated with the distribution of the attributes of each groups of agents covering the share* z_0*. In general* $eZ(X) := Z(X) + (\mathbb{R}_{-} \times \mathbb{R}^d_{+})$*.*

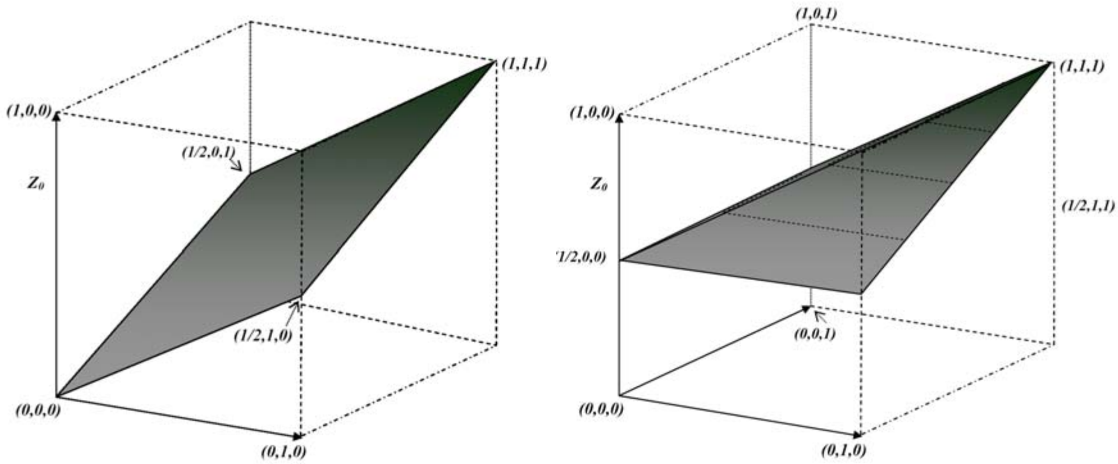
- Price dominance with positive prices can be implemented through the *Extended Lift Zonoid* also for distributions with different total amounts of attributes.

The result:

Theorem 4.3 *The following conditions are equivalent:*
(I) $eLZ(X) \subseteq eLZ(Y)$.
(II) Xp *Generalize Lorenz dominates* Yp *for all* $p \in \mathbb{R}^d_+$*. (Budget dominance)*
(III) $\psi(Yp) \leq \phi(Xp)$ *for all* $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ *which are increasing and S-concave and all* $p \in \mathbb{R}^d_+$*.*
(IV) $\sum_{i=1}^n v(\mathbf{y}_i \cdot p) \leq \sum_{i=1}^n v(\mathbf{x}_i \cdot p)$ *for all* $v : \mathbb{R}_+ \rightarrow \mathbb{R}$ *which are increasing and concave and all* $p \in \mathbb{R}^d_+$*.*

If for each share of population the upper contour set of the section of the Lift Zonoid of X in the d dimensional space is included into the same set for Y then social welfare is larger in X than in Y for SWFs that are increasing and inequality averse.

Getting back to the previous examples comparing $Z = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ and $Y = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. It has already been shown that $Zp >_M Yp$ for $p \in \mathbb{R}^2_+$ thus according to the previous theorem it is also true that $eLZ(Y) \subseteq eLZ(Z)$. One can check this last condition extending the graphical representation of the zonoids of Z and Y depicted in the previous figures.



$eLZ(Y)$: volume below the shaded area $eLZ(Z)$: volume below the shaded area

The extension to $eLZ(Y)$ simply requires to consider all points below the surface representing $LZ(Y)$, given that the extension to all points in the attributes space located north east of each point on the surface still identifies points that are below the surface. For the extension to $eLZ(Z)$ it becomes also relevant to consider the extension to all points in the attributes space that are north east w.r.t. those on the $LZ(Z)$ surface, in this case it is particularly relevant the extension of all points in $LZ(Z)$ located on the segment from $(1/2, 0, 0)$ to $(1, 1, 1)$.

4.2 Multidimensional Stochastic orders

Integral Stochastic Orders can be applied also in a multidimensional framework where for instance are considered distributions in $\mathbb{R}^d_+$.

Here we focus on very few stochastic orders, identifying some key aspects and referring to Shaked and Shanthikumar (1994) and Müller and Stoyan (2002) for a survey of the relevant literature.

Let $X = (X_{.j} : j = 1, 2, ..., d)$ denote the marginal distributions of each attribute j with generic realization $\mathbf{x} = (x_1, x_2, ..., x_d)$ identifying a d dimensional vector of realizations one for each attribute with

- Cumulative Distribution Function:

$$F_X(\mathbf{x}) := P(X \leq \mathbf{x}) := P(X_{.j} \leq x_j \text{ for all } j = 1, 2, 3, ..., d)$$

  with marginals $F_{X_{.j}}(x)$ for $j = 1, 2, 3, ..., d$.

- Survival Function:

$$\bar{F}_X(\mathbf{x}) := P(X > \mathbf{x}) := P(X_{.j} > x_j \text{ for all } j = 1, 2, 3, ..., d)$$

- In this framework probabilities logically correspond to proportions of populations in the multidimensional distribution setup.

A first result

- Let $u : \mathbb{R}^d_+ \to \mathbb{R}$ denote an "utility" [evaluation] function over d dimensional attributes realizations $\mathbf{x}$ and

$$^d\mathcal{U}_1 := \{u : u \text{ non-decreasing}\}$$

  i.e. if $\mathbf{x} \geq \mathbf{x}'$ then $u(\mathbf{x}) \geq u(\mathbf{x}')$. Thus $\geqslant_{^d\mathcal{U}_1}$ denotes the following integral stochastic order

Definition 4.6 *The multidimensional (d-dimensions) integral stochastic order for functionals in* $^d\mathcal{U}_1$ *requires that*

$$X \geqslant_{^d\mathcal{U}_1} Y \Longleftrightarrow \int_{\mathbb{R}^d_+} udF_{\mathbf{X}} \geq \int_{\mathbb{R}^d_+} udF_{\mathbf{Y}} \ \ \forall u \in {}^d\mathcal{U}_1$$

Definition 4.7 (Upper Set) *The set* $U \in \mathbb{R}^d_+$ *is an upper set iff for all* $\mathbf{x} \in \mathbb{R}^d_+$ *if* $\mathbf{x} \in U$ *then* $\mathbf{y} \in U$ *if* $\mathbf{y} \geq \mathbf{x}$.

Theorem 4.4 *The following statements are equivalent:*
(i) $X \geqslant_{^d\mathcal{U}_1} Y$
(ii) $P(X \in U) \geq P(Y \in U)$ *for all upper set* U *in* $\mathbb{R}^d_+$.

An analogous result can be obtained when focussing on comparisons of budgets where attributes are evaluated in terms of non negative "prices".

Definition 4.8 *Let* $\mathbb{P} := \{p \in \mathbb{R}^d_+ : p_1 + p_2 + ... + p_d = 1\}$

$$X \geqslant^1_{\mathbb{P}} Y \Longleftrightarrow \int_{\mathbb{R}^d_+} g(Xp)dF_{\mathbf{X}} \geq \int_{\mathbb{R}^d_+} g(Yp)dF_{\mathbf{Y}} \ \ \forall g \in {}^1\mathcal{U}_1 \ \forall p \in \mathbb{P}.$$

The following statements are equivalent (Muliere and Scarsini, 1989):

Theorem 4.5 *(i)* $X \geqslant^1_{\mathbb{P}} Y$

(ii) $P(Xp > t) \geq P((Yp > t)$ *for all* $t > 0$. *(Dominance for all upper sets whose boundary is an hyperplane)*

Problem 4.1 *Is it possible to add to the list of equivalent conditions also those on comparisons of* $F_{\mathbf{X}}$ *and* $F_{\mathbf{Y}}$ *or of* $\bar{F}_{\mathbf{X}}$ *and* $\bar{F}_{\mathbf{Y}}$*?*

Definition 4.9 (Upper Orthant order) $X \succcurlyeq_{uo} Y \iff \bar{F}_{\mathbf{X}}(\mathbf{t}) \geq \bar{F}_{\mathbf{Y}}(\mathbf{t})$ *for all* $\mathbf{t} \in \mathbb{R}^d_+$.

Definition 4.10 (Lower Orthant order) $X \succcurlyeq_{lo} Y \iff F_{\mathbf{X}}(\mathbf{t}) \leq F_{\mathbf{Y}}(\mathbf{t})$ *for all* $\mathbf{t} \in \mathbb{R}^d_+$.

- $X \succcurlyeq_{uo} Y$ and $X \succcurlyeq_{lo} Y$ are independent.
- $X \geqslant_{{}^d\mathcal{U}_1} Y \Longrightarrow [X \succcurlyeq_{uo} Y$ and $X \succcurlyeq_{lo} Y]$
- $[X \succcurlyeq_{uo} Y$ and $X \succcurlyeq_{lo} Y] \not\Longrightarrow X \geqslant_{{}^d\mathcal{U}_1} Y$.
- $[X \succcurlyeq_{uo} Y$ and $X \succcurlyeq_{lo} Y]$ give the *Concordance order* $X \succcurlyeq_c Y$, i.e. an order of association between variables

Solution 4.1 *From last set of remarks is clear that the answer to the question posed in the formation of the previous problem is NO!*

Note that by construction Upper Sets are unions of Upper Orthants, but as previously stated dominance for all Upper Orthants is not sufficient to guarantee dominance for all Upper Sets.

Definition 4.11 ($\Delta-$monotone functions) *Consider the function* $u : \mathbb{R}^d_+ \to \mathbb{R}$, *let* $\varepsilon > 0$, $\mathbf{1} := (1, 1, 1, 1, ..., 1)$ *and* $\mathbf{1}_i := (0, 0, 0, 1_i, 0, ..., 0)$

$$\Delta^\varepsilon_i u(\mathbf{x}) := u(\mathbf{x} + \varepsilon \mathbf{1}_i) - u(\mathbf{x}).$$

Function u *is* $\Delta-$*monotone if for every set* $\{i_1, i_2, .., i_k\} \subset \{1, 2, 3, ..., d\}$ *and every* $\varepsilon_i > 0$ *for* $i \in \{1, 2, 3, .., k\}$ *then*

$$\Delta^{\varepsilon_1}_{i_1} \Delta^{\varepsilon_2}_{i_2} ... \Delta^{\varepsilon_k}_{i_k} u(\mathbf{x}) \geq 0.$$

Definition 4.12 ${}^d\Delta_{\mathcal{M}}$ *is the set of all bounded* $\Delta-$*monotone functions* $u : \mathbb{R}^d_+ \to \mathbb{R}$.

Definition 4.13 ${}^d\Delta_{\mathcal{A}}$ *is the set of all bounded* $\Delta -$ *antitone functions* $u : \mathbb{R}^d_+ \to \mathbb{R}$ *i.e.* $u(\mathbf{x}) \in {}^d\Delta_{\mathcal{A}} \Leftrightarrow -u(-\mathbf{x}) \in {}^d\Delta_{\mathcal{M}}$.

Remark 4.4 *Note:* $\Delta-$ *antitone functions satisfy* ***decreasing increments.***

Theorem 4.6 *(i)* $X \succcurlyeq_{uo} Y \Longleftrightarrow X \geqslant_{^d\Delta_{\mathcal{M}}} Y$
(ii) $X \succcurlyeq_{lo} Y \Longleftrightarrow X \geqslant_{^d\Delta_{\mathcal{A}}} Y$.

Definition 4.14 (Supermodular functions) *The function* $u : \mathbb{R}_+^d \to \mathbb{R}$, *is supermodular if for all* $\mathbf{x}, \mathbf{y} \in \mathbb{R}_+^d$

$$\begin{aligned} & u\left(\max\{x_1, y_1\}; ..; \max\{x_d, y_d\}\right) + u\left(\min\{x_1, y_1\}; ..; \min\{x_d, y_d\}\right) \\ \geq \;& u\left(\mathbf{x}\right) + u\left(\mathbf{y}\right). \end{aligned}$$

alternatively if u *is twice differentiable* $\frac{\partial^2 u}{\partial x_i \partial x_j} \geq 0$ *for all* $i, j \in \{1, 2, 3, .., d\}$, $i \neq j$.

Remark 4.5 *A function* $u : \mathbb{R}_+^d \to \mathbb{R}$ *is supermodular if and only if* $\int_{\mathbb{R}_+^d} u dF_{\mathbf{X}} \geq \int_{\mathbb{R}_+^d} u dF_{\mathbf{Y}}$ *whenever* X *is obtained from* Y *through a Correlation Increasing Transformation. (It is an indicator of dependence across attributes)*

Definition 4.15 *The set of supermodular functions is* ${}^d\mathcal{U}_{SM}$.

Consider the bivariate case:

Theorem 4.7 *Let* $d = 2$, *the following statements are equivalent:*
(i) $X \geqslant_{^2\mathcal{U}_{SM} \cap\, ^2\mathcal{U}_1} Y$,
(ii) $X \succcurlyeq_{uo} Y$ *and* $X_{.j} \succcurlyeq_1 Y_{.j}$ *for* $j = 1, 2$

Theorem 4.8 *Let* $d = 2$, *the following statements are equivalent:*
(i) $X \geqslant_{^2\mathcal{U}_{SM}} Y$
(ii) $Y \succcurlyeq_{lo} X$ *and* $X_{.j} = Y_{.j}$ *for* $j = 1, 2$
(iii) $X \succcurlyeq_{uo} Y$ *and* $X_{.j} = Y_{.j}$ *for* $j = 1, 2$

Related results can be found in Atkinson and Bourguignon (1987), Bourguignon and Chakravarty (2002), Athey (2000, 2002).
For $d > 2$ some of the equivalences may break.

Theorem 4.9 *Let* $d = 2$, *and* $X_{.j} = Y_{.j}$ *for* $j = 1, 2$ *the following statements are equivalent:*
(i) $X \geqslant_{^2\mathcal{U}_{SM}} Y$;
(ii) $Y \succcurlyeq_{lo} X$
(iii) $X \succcurlyeq_{uo} Y$
(iv) $Cov[f_1(X_{.1}), f_2(X_{.2})] \geq Cov[f_1(Y_{.1}), f_2(Y_{.2})]$ *for increasing functions* f_1, f_2;
(v) $X \geqslant_{^d\Delta_{\mathcal{M}}} Y$.

One more result by Scarsini (J appl. Prob 1998), it relates to comparisons of multidimensional distributions (not necessarily bivariate) with common marginals.

Theorem 4.10 *If* $X_{.j} = Y_{.j}$ *for all* $j = 1, 2, .., d$ *then* $Y \geqslant_{^d\mathcal{U}_{SM}} X$ *implies that* Xp *Lorenz dominates* Yp *for all* $p \in \mathbb{R}_+^d$.

With common marginals stochastic dominance in terms of supermodular functions implies price dominance (with positive prices).

A closer look at dependence between variables. The dependence structure of a distribution can be represented by a Copula (a random vector uniformly distributed between $[0,1]$)

Consider for the Frechet class $\Gamma(F_{X_{.1}}, F_{X_{.2}}, ..., F_{X_{.d}})$ of d dimensional distributions with $F_{X_{.1}}, F_{X_{.2}}, ..., F_{X_{.d}}$ as marginals.

Given a $F_X \in \Gamma(F_{X_{.1}}, ..., F_{X_{.d}})$ there exist a copula $C : [0,1]^d \to [0,1]$ s.t. for all $\mathbf{x} \in \mathbb{R}^d$

$$F_{\mathbf{X}}(\mathbf{x}) = C[F_{\mathbf{X}_{.1}}(x_1), F_{\mathbf{X}_{.2}}(x_2), ..., F_{\mathbf{X}_{.d}}(x_d)]$$

and can be constructed if F_X is continuous as:

$$C[\mathbf{u}] := F_X[F_{x_{.1}}^{-1}(u_1); F_{x_{.2}}^{-1}(u_2); ..., F_{x_{.d}}^{-1}(u_d)] \quad \mathbf{u} \in [0,1]^d.$$

This result clarifies a connection between comparisons of marginals when the association between the variable represented by the copula is fixed.

Theorem 4.11 *If X and Y have a common copula then $X_{.j} \succcurlyeq_1 Y_{.j}$ for all $j = 1, 2, ..., d$ implies $X \geqslant_{d\mathcal{U}_1} Y$.*

Thus with common copula (association) then dominance in terms of the distributions of each attribute is sufficient to guarantee multivariate dominance!

5 Suggested Readings

Surveys on Stochastic Orders: Shaked, M. and Shanthikumar, J. G. (1994); Müller, A. and Stoyan, D. (2002).

References for Inequality, Welfare and Poverty measurement:

Books: Sen (1997); Lambert, (2001); Chakravarty, (1990);

Surveys: Mosler and Muliere (1998); Zheng (1997, 2000); Cowell, F. A. (2000) Chakravarty and Muliere, (2003, 2004);

References

[1] Aaberge, R. (2004): Ranking intersecting Lorenz curves. CEIS Working Paper 45.

[2] Athey, S. (2000): Characterizing properties of stochastic objective functions. Mimeo MIT.

[3] Athey, S. (2002): Monotone comparative statics under uncertainty. *Quarterly Journal of Economics*, 187-223.

[4] Atkinson, A. B. (1970): On the measurement of inequality. *Journal of Economic Theory*, **2**, 244-263.

[5] Atkinson, A. B. (1987): On the measurement of poverty. *Econometrica* **55**, 749-764.

[6] Atkinson, A.B. (2003): Multidimensional deprivation: constrasting social welfare and counting approaches. *Journal of Economic Inequality*, vol. 1, pp. 51–65.

[7] Atkinson, A. B. and Bourguignon, F. (1982): The comparison of multidimensioned distribution of economic status. *Review of Economic Studies*, **49**, 183-201.

[8] Ben Porath, E. and Gilboa, I. (1994): Linear measures, the Gini index, and the income-equality trade-off. *Journal of Economic Theory*, **18**, 59-80.

[9] Ben-Porath, E., Gilboa, I., Schmeidler, D., 1997. On the measurement of inequality under uncertainty. *Journal of Economic Theory* 75, 194–204.

[10] Bourguignon, F. and S. R. Chakravarty (2003): The Measurement of Multidimensional Poverty, *Journal of Economic Inequality*, 1, 25–49.

[11] Bourguignon, F. and Chakravarty, S. R. (2002): Multidimensional poverty orderings. Delta, Working Paper 2002-22.

[12] Chakravarty, S. R. (1990): Ethical Social Index Numbers. Berlin: Springer-Verlag.

[13] Chakravarty, S. R. (2003): A generalization of the human development index. *Review of Development Economics*, 7, 99-114.

[14] Chakravarty, S and Muliere, P. (2003): Welfare indicators: a review and new perspectives. 1. Measurement of inequality. *Metron, International Journal of Statistics,* **61**, 457-497.

[15] Chakravarty, S and Muliere, P. (2004): Welfare indicators: a review and new perspectives. 2. Measurement of poverty. *Metron, International Journal of Statistics,* **62**, 247-281.

[16] Chakravarty. S. and F. Bourguignon (2008): Multidimensional Poverty Orderings: Theory and Applications, Forthcoming in Welfare, Development, Philosophy and Social Science: Essays for Amartya Sen's 75th Birthday (Volume 2: Development Economics and Policy), K. Basu and R. Kanbur (eds.), London: Oxford University Press.

[17] Chakravarty, S. R., D'Ambrosio C. (2006) The measurement of social exclusion. *Review of Income and Wealth*, 52 (3) , 377–398.

[18] Chateauneuf, A., Gajdos, T. and Wilthien, P. H. (2002): The principle of strong diminishing transfer. *Journal of Economic Theory,* ***103***, 311-333.

[19] Cowell, F. A. (2000): Measurement of Inequality. In *Handbook of Distribution* (Atkinson, A. B. and Bourguignon, F. eds.) North Holland

[20] Dardanoni, V. and Lambert, P. J. (1988): Welfare rankings of income distributions: a role for the variance and some insights for tax reform. *Social Choice and Welfare*, **5**, 1-17.

[21] Dardanoni, V. (1996): On Multidimensional Inequality Measurement in *Research on Economic Inequality: Income Distribution, Social Welfare, Inequality and Poverty* (Vol. 6), eds. C. Dagum and A. Lemmi, JAI Press Inc., pp. 201-205.

[22] Dasgupta, P., Sen, A. K. and Starrett, D. (1973): Notes on the measurement of inequality. *Journal of Economic Theory*, **6**, 180-187.

[23] Decancq, K., Decoster, A. & Schokkaert, E. (2007). The evolution in world inequality in well-being. Discussion Paper Series 07/04, Center for Economic Studies, Katholieke Universiteit Leuven.

[24] Deutsch, J. and Silber J. (2005): Measuring multidimensional poverty: an empirical comparison of various approaches, *Review of Income and Wealth* 51 (1) , 145–174.

[25] Donaldson, D. and Weymark, J. A. (1980): A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, **22**, 67-86.

[26] Donaldson, D. and Weymark, J. A. (1983): Ethically flexible Gini indices for income distributions in the continuum. *Journal of Economic Theory*, **29**, 353-358.

[27] Duclos, P. Y., Sahn, D. and Younger, S. D. (2006). Robust multidimensional poverty comparisons. *Economic Journal*, 116, 943–68.

[28] Dutta I., Pattanaik P. K. and Xu Y. (2003): On Measuring Deprivation and the Standard of Living in a Multidimensional Framework on the Basis of Aggregate Data, *Economica*, 70, 197-221.

[29] Ebert, U. (1988): Measurement of inequality: an attempt at unification and generalization. *Social Choice and Welfare*, **5**, 147-69.

[30] Epstein L. G. and S. M. Tanny (1980): Increasing Generalized Correlation: A Definition and Some Economic Consequences. *The Canadian Journal of Economics*, **13**, 16-34.

[31] Fields, G. S. and Fei, C. H. (1978): On inequality comparisons. *Econometrica,* **46**, 303-316.

[32] Fishburn, P. C. (1976): Continua of stochastic dominance relations for bounded probability distributions. *Journal of Mathematical Economics*, **3**, 295-311.

[33] Fishburn, P. C. and Willig, R. D. (1984): Transfer principles in income redistribution. *Journal of Public Economics*, **25**, 323-328.

[34] Foster, J. (1985): Inequality measurement. In *Fair Allocation* (Young, H. P. ed.)*;* Proceeding of Symposia in Applied Mathematics **33,** 31-68. Providence: The American Mathematical Society.

[35] Foster, J.; Greer, J. and Thorbecke, D. (1984): A class of decomposable poverty measures. *Econometrica*, **52**, 761-766.

[36] Foster, J. and Shorrocks, A. F. (1988): Poverty orderings. *Econometrica*, **56**, 173-177.

[37] Foster, J.E., L. Lopez-Calva, and M. Szekely (2005): Measuring the Distribution of Human Development: Methodology and an Application to Mexico. *Journal of Human Development,* **6**, 5-29.

[38] Gajdos,T., Maurin, E. (2004): Unequal uncertainties and uncertain inequalities:An axiomatic approach. *Journal of Economic Theory* 116, 93–118.

[39] Gajdos, T., and J. A. Weymark. (2005): Multidimensional Generalized Gini Indices, *Economic Theory*, 26 (3): 471-496.

[40] Gastwirth, J. L. (1971): A general definition of the Lorenz curve. *Econometrica*, **39**, 1037-1039.

[41] Gilboa, I., Schmeidler, D. (1989): Maximin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141–153.

[42] Hadar, J. and Russell, W. (1969): Rules for ordering uncertain prospects. *American Economic Review*, **49**, 25-34.

[43] Hardy, G.H., Littlewood, J.E., and Polya, G. (1934), *Inequalities,* London, Cambridge University Press.

[44] Hicks, D.A. (1997): The inequality-adjusted Human Development Index: a constructive proposal. *World Development* , 25, 1283–1298.

[45] Jenkins, S. and Lambert, P. J. (1997): Three I's of poverty curves, with an analysis of UK poverty trends. *Oxford Economic Papers* **49**, 317-327.

[46] Kakwani, N. C. (1980): On a class of poverty measures. *Econometrica*, **48**, 437-446.

[47] Kolm, S. C. (1969): The optimal production of social justice. In *Public Economics* (Margolis, J. and Gutton, H. ed.), pp. 145-200. London: Mcmillan.

[48] Kolm, S. C. (1976): Unequal inequalities I, II. *Journal of Economic Theory*, **12**, 416-442; **13**, 82-111.

[49] Kolm, S.-C. (1977): Multidimensional Egalitarianism. *The Quarterly Journal of Economics*, 91 (1): 1-13.

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

[51] Koshevoy, G., Mosler, K. (1996) The Lorenz zonoid of a multivariate distribution. *Journal of the American Statistical Association* 91, 873–882

[52] Koshevoy, G., Mosler, K. (2007): Multivariate Lorenz dominance based on zonoids. *Advances in Statistical Analysis* 91, 57–76.

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

[54] Lambert, P. J. (2001): *The Distribution and Redistribution of Income: a Mathematical Analysis*. 2nd Edition. Manchester: Manchester University Press.

[55] LeBreton, M. (1994): Inequality, poverty measurement and welfare dominance: an attempt at unification, *Models and Measurement in of Welfare and Inequality* (Eichhorn, W ed.) 120-140. Berlin: Springer Verlag.

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

[57] Maccheroni, F., Muliere, P. and Zoli, C. (2005): Inverse stochastic orders and generalized Gini functionals. *Metron, International Journal of Statistics*, **63**, 3, 529-559.

[58] Marshall, A. W. and Olkin, I. (1979): *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press.

[59] Mehran, F. (1976): Linear measures of income inequality. *Econometrica,* **44**, 805-809.

[60] Mosler, K. and Muliere, P. (1998): Welfare means and equalizing transfers. *Metron, International Journal of Statistics,* **56**, 11-52.

[61] Muliere, P. and Scarsini, M. (1989): A note on stochastic dominance and inequality measures. *Journal of Economic Theory*, **49**, 314-323.

[62] Müller, A. (1997): Stochastic orders generated by integrals: a unified study. *Advances in Applied Probability*, **29**, 414-428.

[63] Müller, A. and Stoyan, D. (2002): *Comparisons Methods for Stochastic Models and Risks.* John Wiley and Sons, Ltd: Chichester, UK.

[64] Pattanaik, P. K., Reddy, S. G. and Xu, Y. (2008): On procedures for measuring deprivation and living standards of societies in a multi-attribute framework. Working Paper 08-02 Andrew Young School of Policy Studies.

[65] Rubinstein A. and P. C. Fishburn (1986): Algebraic aggregation theory. *Journal of Economic Theory*, 38, 1, 63-77.

[66] Rothschild, M. and J. Stiglitz (1970), "Increasing risk: I. A definition", *Journal of Economic Theory* **2**, 225-243.

[67] Rothschild, M. and Stiglitz, J. E. (1973): Some further results on the measurement of inequality. *Journal of Economic Theory*, **6**, 188-204.

[68] Saposnik, R. (1981): Rank-dominance in income distributions. *Public Choice*, **36**, 147-51.

[69] Savaglio, E., 2002. Multidimensional inequality: A survey. In: Farina, F., Savaglio, E. (Eds.), *Inequality and Economic Integration.* Routledge, London.

[70] Scarsini, M. (1998): Multivariate convex orderings, dependence, and stochastic equality. *Journal of Applied Probability*, **35**, 93-103.

[71] Sen, A. K. (1973): *On Economic Inequality.* Oxford: Claredon Press. (1997) expanded edition with the annexe "*On Economic Inequality After a Quarter Century*" by Foster, J. And Sen, A.K..

[72] Sen, A. K. (1976): Poverty: an ordinal approach to measurement. *Econometrica*, **44**, 219-231.

[73] Shaked, M. and Shanthikumar, J. G. (1994): *Stochastic Orders and Their Applications.* Academic Press: Boston.

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

[75] Shorrocks, A. F. (1998): Deprivation profiles and deprivation indices. In *The Distribution of Welfare and Household Production: International Perspectives,* Jenkins, S. A., Kaptein, S. A. and van Praag, B. eds. 250-267. London: Cambridge University Press.

[76] Shorrocks, A. F. and Foster, J. E. (1987): Transfer sensitive inequality measures. *Review of Economic Studies*, **14**, 485-497.

[77] Silber, J. (1999): *Handbook of Income Inequality Measurement.* J. Silber ed. Kluwer Academic: Boston.

[78] Spencer, B. D. and Fisher, S. (1992): On comparing distributions of poverty gaps. *Sankya,* **54B**, 114-126.

[79] Thistle, P. D. (1993): Negative moments, risk aversion and stochastic dominance. *Journal of Financial and Quantitative Analysis*, **28**, 301-311.

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

[81] Tsui, K.-Y. (1999): Multidimensional inequality and multidimensional generalied entropy measures: an axiomatic approach, *Social Choice and Welfare,* **16**, 145–158.

[82] Tsui, K.-Y.(2002): Multidimensional poverty indices, *Social Choice and Welfare,* **19**, 69–93.

[83] Wang, S. S. and Young, V. R. (1998): Ordering risks: Expected utility theory versus Yaari's dual theory of risk. *Insurance: Mathematics and Economics,* **22**, 145-161.

[84] Weymark, J. A. (1981): Generalized Gini inequality indices. *Mathematical Social Sciences*, **1**, 409-430.

[85] Weymark, J. (2004): The normative approach to the measurement of multidimensional inequality, Working Paper No. 03-W14R, Vanderbilt University. In: Farina, F., Savaglio, E. (Eds.), Inequality and Economic Integration. Routledge, London.

[86] Yaari, M. E. (1987): The dual theory of choice under risk. *Econometrica*, **55**, 95-115.

[87] Yitzhaki, S. (1983): On an extension of the Gini inequality index. *International Economic Review,* **24**, 617-628.

[88] Zheng, B. (1997): Aggregate poverty measures. *Journal of Economic Surveys,* **11**, 123-162.

[89] Zheng, B. (2000): Poverty orderings. *Journal of Economic Surveys,* **14**, 427-466.

[90] Zoli, C. (1999): Intersecting generalized Lorenz curves and the Gini index. *Social Choice and Welfare,* **16**, 183-196.

[91] Zoli, C. (2002): Inverse stochastic dominance, inequality measurement and Gini indices.*Journal of Economics,* Supplement # 9, P. Moyes, C. Seidl and A.F. Shorrocks (Eds.), Inequalities: Theory, Measurement and Applications, 119-161.