October 6, 2003

Restricted Lorenz dominance of economic inequality in one and many dimensions[1]

Karl Mosler
Seminar für Wirtschafts- und Sozialstatistik
Universität zu Köln, 50923 Köln, Germany
mosler@statistik.uni-koeln.de

Abstract

The paper investigates Lorenz dominance and non-scaled Lorenz dominance to compare distributions of economic status in one and several attributes. Restrictions of these dominance relations are developed that focus on central parts of the distributions and facilitate their comparison.

Keywords: Inequality measurement, lift zonoid order, Lorenz order, majorization, multivariate inequality ordering.

1 Introduction

Economic inequality often arises from more than one attribute. E.g., households differ in income and wealth, individuals vary on earnings and education, countries on per capita income, life expectancy and mineral resources. To compare economic inequality in several, say d, attributes,

[1]Paper presented to the Workshop "Income Distribution and Welfare", Centenario Università Bocconi 2002. I thank the participants, in particular James Foster and Gleb Koshevoy for their helpful remarks. I am also indebted to two referees for their suggestions concerning terminology and interpretation.

various dominance relations among d-variate distributions have been proposed in the literature (Kolm (1977), Atkinson and Bourguignon (1982), Mosler (1994b), Koshevoy and Mosler (1997)). Each of them is related to utilitarian social welfare functions and can be characterized as unanimous preference for a certain class of utilities. But all these dominance relations appear to be rather coarse incomplete orderings. In most real data situations they are unable to compare empirical distributions. Even in one-dimensioned measurement the Lorenz curves of empirical income distributions often intersect.

Given a quasi-ordering (that is, a reflexive, transitive, but generally incomplete binary relation) of inequality among d-variate distributions, $d \geq 1$, a real-valued index is regarded as an *inequality index* if it is consistent, that is increasing, with that ordering. In contrast to the ordering, an inequality index provides a complete ordering among distributions: any two distributions of well-being can be compared by the index. But, if two distributions are not comparable by the quasi-ordering, inequality indices can be constructed which point in different directions: some indicate more inequality at the first distribution and others at the second distribution.

In most practical applications, no inequality ordering but several inequality indices are employed to compare distributions. This approach raises many questions: Which indices should be chosen? What are the consequences of combining the indices? How to decide when the indices point in opposite directions? Usually, these questions are not explicitly addressed and the decisions made on an *ad hoc* basis, which appears to be unsatisfactory.

The multivariate approach to economic inequality has been put forth by Tobin (1970) and Sen (1970, 1973). Special quasi-orderings of multi-dimensioned inequality have been proposed and investigated by Kolm (1977), Atkinson and Bourguignon (1982), Mosler (1994a,b), Koshevoy (1995), Koshevoy and Mosler (1996), Mosler (2002, Ch 9). For multivariate inequality indices and practical measurements, see Maasoumi (1986), Tsui (1995), and others. Maasoumi (1999) provides a comprehensive survey.

In the sequel a new approach to the ordering of (single-dimensioned and multi-dimensioned) inequality is proposed that focuses on properly defined central parts of the distributions. When looking at two empirical distributions of income, it is often observed that

their Lorenz curves are ordered in the middle, but cross in the upper or lower tails. Therefore we introduce the *restricted L-dominance*, which is simply the ordering of usual Lorenz functions restricted to a central subinterval $[p_0, p_1]$ of the unit interval $[0, 1]$. Comparing two income distributions by the restricted L-dominance means that we do not care about the incomes in the lower and upper parts of the two populations. A similar restriction is introduced for the univariate decreasing (resp. increasing) convex dominance, which is not scale invariant. The *restricted decreasing* (resp. increasing) *convex dominance* focuses on those incomes in the population that lie between two properly chosen income quantiles.

The idea of restricting the Lorenz dominance to a proper central part of the distributions is then extended to the multi-dimensioned case. We consider three quasi-orderings of multivariate distributions that have been proposed in the literature: the convex scaled dominance, the convex-linear scaled dominance, and the price Lorenz dominance. Each of them generalizes the usual univariate Lorenz dominance and is characterized by a class of utilitarian social welfare functions. The second quasi-ordering is finer than the first, and the third is finer than the second. Besides that, the three quasi-orderings are very similar.

Recently (see Mosler (2002)), a different characterization has been given for the convex-linear dominance: The ordering is equivalent to the inclusion of sets from a properly defined family of central regions of the two given distributions in d-space.

In this paper, we employ the characterization to define a new dominance relation among distributions in d-space: the restricted convex-linear dominance. This is done by restricting the set inclusion to a 'relevant' subfamily of central regions. Analogously, a restricted version of the price Lorenz dominance is introduced. Finally, we develop similar restricted dominance relations which are also sensitive to changes in scale. The new restricted quasi-orderings provide flexible tools to compare multivariate distributions.

Politically speaking, the restricted L-dominance orderings focus on the 'middle class'. On empirical grounds, our approach is also justified by the fact that the data on incomes (and other attributes) in the lower and upper tails of the distribution is often difficult to obtain. Further, the approach is robust against outliers, which is important when dealing with such data. (Due to the widespread reluctance to revealing income related information, the data often has low quality and contains 'bad' outliers.)

Section 2 starts with the measurement of single-dimensioned inequality by restrictions of the Lorenz dominance and the decreasing (resp. increasing) convex dominance. Section 3 introduces the three multivariate extensions of the Lorenz dominance. Next, in Section 4, characterization theorems in terms of set inclusions are given for two of them and these characterizations are used to restrict the orderings to a 'relevant' family of central regions. Section 5 presents restricted multivariate dominance relations which are not scale invariant, and Section 6 concludes with remarks on limitations and possible extensions of our approach.

2 Comparing single-dimensioned inequality

Consider a vector $X = [x_1, \dots, x_n]$ of data x_i in $\mathbb{R}$ which have positive mean, $\bar{x} > 0$. We regard the data as an income distribution in a population of n households. Let $x_{(1)} \le ... \le x_{(n)}$ denote the ordered data and $Q_X :]0,1] \to \mathbb{R}$,

$$Q_X(p) = x_{(k)} \quad \text{if} \quad \frac{k-1}{n} < p \le \frac{k}{n}, \quad k = 1, 2, \dots, n,$$

the quantile function. The integral of the quantile function,

$$GL_X(p) = \int_0^p Q_X(p')dp', \quad p \in [0,1],$$

is the *generalized Lorenz (GL) function*, its graph the *GL-curve.*

Let $Y = [y_1, \dots, y_n]$ be another vector of data in $\mathbb{R}$ with mean $\bar{y} > 0$, say, an alternative income distribution in the population. By definition, X is larger than Y in the *decreasing convex dominance*, $X \succeq_{dcx} Y$, if

$$GL_X(p) \le GL_Y(p) \quad \text{for all} \quad 0 \le p \le 1. \tag{1}$$

In other words (Marshall and Olkin, 1979), $X \succeq_{dcx} Y$ means that X *weakly supermajorizes* Y. X is larger than Y in the *increasing convex dominance* (or: X *weakly submajorizes* Y), $X \succeq_{icx} Y$, if

$$\overline{x} - GL_X(1-p) \geq \overline{y} - GL_Y(1-p) \quad \text{for all} \quad 0 \leq p \leq 1\,. \tag{2}$$

Notate $X/\bar{x} = \{x_1/\bar{x}, \ldots, x_n/\bar{x}\}$ and let

$$L_X(p) = GL_{X/\bar{x}}(p) = \frac{1}{\bar{x}} GL_X(p)\,, \quad p \in [0,1]\,.$$

The function L_X is the usual *L-function*, its graph the *L-curve*. Then, by definition, X is larger than Y in the *L-dominance*, $X \succeq_L Y$, if

$$L_X(p) \leq L_Y(p) \quad \text{for all} \quad 0 \leq p \leq 1\,. \tag{3}$$

It is seen from this definition that the Lorenz dominance is the decreasing convex dominance (and, as well, the increasing convex dominance) of the two distributions of data scaled down by their means.

Many indices have been proposed in the literature that are consistent (equivalently, increasing) with the Lorenz dominance, among them the Gini ratio, the coefficient of variation, the Pietra (see also Kuznets) index, the coefficients by Theil, Atkinson and Kolm, the coefficient of equal shares, and the coefficient of minimal majority; see e.g. Piesch (1975) and Cowell (1995). All these indices are (at least, weakly) increasing with the Lorenz dominance but each of them relates to a particular concept of distance between the two distributions and is sensible to changes in different parts of their L-curves.

As mentioned in the introduction, in order to focus on the central parts of the two distributions and to neglect eventual crossings of the L-curves in their outer tails, we introduce the following restricted quasi-orderings.

Definition 1 **(Univariate restricted dominance)** Let $0 \leq p_0 < p_1 \leq 1$ and define

(i) the *restricted decreasing convex dominance*
$X \succeq_{dcx}^{p_0,p_1} Y$ if $GL_X(p) \leq GL_Y(p)$ for $p_0 \leq p \leq p_1$.

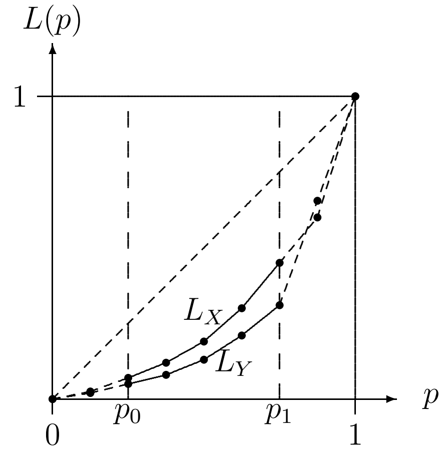


Figure 1: Restricted L-dominance, $d = 1$.

(ii) the *restricted L-dominance*
$X \succeq_L^{p_0,p_1} Y \quad \text{if} \quad L_X(p) \leq L_Y(p) \ \text{ for } \ p_0 \leq p \leq p_1\,.$

Figure 1 illustrates the restricted L-dominance.

Properties of the restricted orderings: Given a set of distributions and p_0, p_1 being fixed, both the restricted L-dominance and the restricted decreasing convex dominance are reflexive and transitive, but in general not complete, that is, only quasi-orderings. Each is monotone on dilations and implied by the respective non-restricted version.

The restricted orderings abstract from the distributions of X and Y in the population's poorest $p_0 \cdot 100$ percent and richest $(1 - p_1) \cdot 100$ percent. If two income distributions have the same mean income, the restricted decreasing convex dominance is equivalent to saying that the distributions are ordered in the usual, unrestricted decreasing convex dominance under the assumption that all 'poor people' below the p_0-quantile have an egalitarian distribution, and all 'rich people' above the p_1-quantile have an egalitarian distribution as well. (However, the same claim is not true for the restricted L-dominance.)

The question remains how the bounds p_0 and p_1 should be chosen. In view of the coefficient of minimal majority, which is given by $MM(X) =$

$L_X^{-1}(0.5)$, it makes sense to choose p_0 and p_1 in $\succeq_L^{p_0,p_1}$ such that

$$p_0 \leq L_X^{-1}(0.5) \leq p_1 \tag{4}$$

holds for any data X under consideration. Then the coefficient of minimal majority is consistent with the restricted Lorenz dominance on these X. Similarly note that the Pietra index of X, i.e. the maximum vertical distance of the Lorenz curve from the diagonal, is attained at $p = F_X(\bar{x})$, where F_X is the empirical distribution function of X. If we choose p_0 and p_1 such that

$$p_0 \leq p = F_X(\bar{x}) \leq p_1 \tag{5}$$

holds for any data set X at hand, the Pietra index is consistent with the restricted L-dominance on these data sets.

Beyond these conditions (4) and (5), the actual choice of p_0 and p_1 is to be based on the subject of the application and the quality of the data. E.g., when the personal income distribution in a developing country are studied, the extremely poor and the extremely rich (say, all people living from less than 2 Euro per day and all earning more than 1000 Euro per day) may be incidently omitted from consideration. Concerning data, e.g. tax statistics disregard the poorer part of the population which pays no income tax, while many censuses do not cover the rich.[2]

Also, like calculating and reporting several indices of inequality, one may check the restricted orderings for several values of the parameters p_0 and p_1 and report these results.

3 Multivariate extensions of L-dominance

In this section we survey three orderings of multivariate distributions each of which extends the usual, univariate L-dominance. The definition (3) of the L-dominance in dimension $d = 1$, which is based on ordered data and quantile functions, does not extend to dimensions two and more,

[2]The German EVS (Einkommens- und Verbrauchsstichprobe 1998) is restricted to monthly net incomes below 17900 Euro.

since there exists no natural order of data in $\mathbb{R}^d, d \geq 2$. However, the L-dominance in dimension one is equivalently characterized by convex evaluations of the scaled data,

$$\begin{aligned} & X \succeq_L Y \iff \\ & \sum_{i=1}^{n} \varphi\left(\frac{x_i}{\bar{x}}\right) \geq \sum_{i=n}^{n} \varphi\left(\frac{y_i}{\bar{y}}\right) \quad \text{for all convex } \varphi : \mathbb{R} \to \mathbb{R}\,. \end{aligned} \tag{6}$$

In other words, Y is unanimously preferred over X on the basis of a utilitarian social welfare function with arbitrary concave utility. This characterization of the L-dominance as a 'convex scaled dominance' generalizes in a straightforward way.

Consider a $d \times n$ data matrix $X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$ with $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{id})'$ for $i = 1, \ldots, n$. We notate

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^{n} x_{ij}\,, \quad j = 1, \ldots, d\,,$$

Then, for $i = 1, \ldots, n$,

$$\left(\frac{x_{i1}}{\bar{x}_1}, \ldots, \frac{x_{id}}{\bar{x}_d}\right)$$

is the i-th data point, componentwise scaled down by the component mean.

Two data matrices $X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$ and $Y = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n]$ in $\mathbb{R}_+^{d \times n}$, $d \geq 1$, may be compared by each of the following three multivariate dominance orderings.

Definition 2 **(Multivariate dominance)** Let $d \geq 1$ and define

(i) the *convex scaled dominance* $X \succeq_{cx}^{s} Y$ if

$$\sum_{i=1}^{n} \varphi\left(\frac{x_{i1}}{\bar{x}_1}, \frac{x_{i2}}{\bar{x}_2}, \ldots, \frac{x_{id}}{\bar{x}_d}\right) \geq \sum_{i=1}^{n} \varphi\left(\frac{y_{i1}}{\bar{y}_1}, \frac{y_{i2}}{\bar{y}_2}, \ldots, \frac{y_{id}}{\bar{y}_d}\right) \tag{7}$$

for all convex functions $\varphi : \mathbb{R}^d \to \mathbb{R}\,,$

(ii) the *convex-linear scaled dominance* $X \succeq^s_{lz} Y$ if
(7) holds for all convex-linear functions $\varphi : \mathbb{R}^d \to \mathbb{R}$, that is, for all φ that have the form

$$\varphi(z_1, \ldots, z_d) = \psi(\alpha_1 z_1 + \cdots + \alpha_d z_d) \tag{8}$$

with $\psi : \mathbb{R} \to \mathbb{R}$ convex and $\alpha_1, \ldots, \alpha_d \in \mathbb{R}$,

(iii) the *price Lorenz dominance* $X \succeq_{PL} Y$ if
(7) holds for all convex-linear φ of form (8) with ψ convex and $\alpha_1, \ldots, \alpha_d \geq 0$.

$X \succeq^s_{lz} Y$ says that, for all coefficients $\alpha_1, \ldots, \alpha_d \in \mathbb{R}$, the linear combinations of the attributes are ordered in the univariate L-dominance,

$$\left(\sum_{j=1}^{d} \alpha_j \frac{x_{1j}}{\bar{x}_j}, \ldots, \sum_{j=1}^{d} \alpha_j \frac{x_{nj}}{\bar{x}_j} \right) \succeq_L \left(\sum_{j=1}^{d} \alpha_j \frac{y_{1j}}{\bar{y}_j}, \ldots, \sum_{j=1}^{d} \alpha_j \frac{y_{nj}}{\bar{y}_j} \right) .$$

$X \succeq^s_{PL} Y$ means the same for all nonnegative coefficients. The coefficients can be interpreted as 'prices', the linear combinations as 'expenditures'. So, $X \succeq^s_{PL} Y$ is tantamount saying that for all nonnegative prices the values (that is, expenditures) of individual endowments are less equal in the X-population than in the Y-population.

The multivariate convex scaled dominance $\succeq^s_{cx}$ and the price Lorenz dominance $\succeq_{PL}$ have been introduced to inequality measurement by Kolm (1977). The convex-linear scaled dominance $\succeq^s_{lz}$ is also mentioned as the *lift zonoid scaled dominance* (Koshevoy and Mosler, 1998).

It is obvious from the definitions that

- $X \succeq^s_{cx} Y \quad \Rightarrow \quad X \succeq^s_{lz} Y \quad \Rightarrow \quad X \succeq_{PL} Y$,
- each of the three relations is reflexive and transitive, but in general not complete (i.e., a quasi-ordering),
- each is scale invariant and extends the univariate L-dominance,
- each implies the usual L-dominance for every single attribute $j = 1, \ldots, d$.

A simple example may illustrate the three quasi-orderings: Consider two distributions of income and wealth among three households,

$$X = \begin{bmatrix} 0 & 2 & 4 \\ 0 & 40 & 80 \end{bmatrix}, \qquad Y = \begin{bmatrix} 1 & 2 & 3 \\ 20 & 40 & 60 \end{bmatrix}.$$

In both distributions, the mean income equals 4, and the mean wealth equals 20. Obviously, the restriction (7) is satisfied for any convex function $\varphi : \mathbb{R}^2 \to \mathbb{R}$, hence $X \succeq_{cx}^s Y$ holds and, consequently, $X \succeq_{lz}^s Y$ and $X \succeq_{PL} Y$. For income and wealth alone, $X_1 \succeq_L Y_1$ and $X_2 \succeq_L Y_2$ is satisfied. Now we exchange the wealth of the first and the third household in X,

$$\widetilde{X} = \begin{bmatrix} 0 & 2 & 4 \\ 80 & 40 & 0 \end{bmatrix}$$

and obtain $\widetilde{X}_1 \succeq_L Y_1$ and $\widetilde{X}_2 \succeq_L Y_2$ as before. But, e.g., with $\varphi_*(z_1, z_2) = (z_1 + z_2)^2$ the l.h.s. of (7) amounts to 12, while the r.h.s. is $14 > 12$, hence $\widetilde{X} \not\succeq_{PL} Y$. Moreover, a small calculation shows that the inequality (7) with $\varphi(z_1, z_2) = \psi(a_1 z_1 + a_2 z_2)$ holds if either $a_2/a_1 \geq 3$ or $a_2/a_1 \leq 1/3$ and ψ is arbitrary convex, and the reverse inequality holds if $1/3 \leq a_2/a_1 \leq 3$ and ψ is arbitrary convex.

The price Lorenz dominance $\succeq_{PL}$ relates to a dominance relation proposed by Atkinson and Bourguignon (1982) as an ordering of multivariate well-being. For $d = 2$, $\succeq_{PL}$ is implied by the first dominance relation in Atkinson and Bourguignon (1982), which is defined by postulating (7) for all differentiable functions φ whose mixed partial second derivatives are nonnegative. Obviously, φ in (8) with nonnegative $\alpha_1, \ldots, \alpha_d$ is such a function. Note that, with $\succeq_{PL}$, every correlation increasing transfer increases inequality. As a consequence, $\succeq_{PL}$ appears to be a sensible measure of inequality only if every pair of attributes is substitutional rather than complementary.

Further, it can be demonstrated that each of the quasi-orderings $\succeq_{cx}^s$, $\succeq_{lz}^s$ and $\succeq_{PL}$ is

- antisymmetric among distributions,
- increasing on dilation. In particular, the quasi-orderings are minimal at an egalitarian distribution (which is constant at the mean).

The three quasi-orderings are different, but the differences are small. For example, compare the following two distributions of income and wealth [in Thsd. Euro] among $n = 6$ households,

$$X = \begin{bmatrix} 0 & 4 & 8 & 4 & 4 & 4 \\ 20 & 0 & 40 & 20 & 20 & 20 \end{bmatrix},$$
$$Y = \begin{bmatrix} 2 & 2 & 4 & 4 & 6 & 6 \\ 10 & 10 & 30 & 30 & 20 & 20 \end{bmatrix}.$$

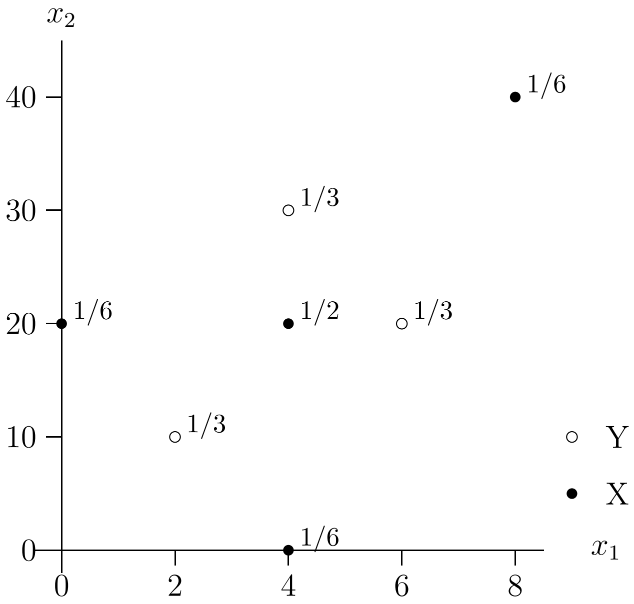


Figure 2: Example for $X \succeq_{lz} Y$, but $X \not\succeq_{cx} Y$.

It can be shown (see a similar example in Elton and Hill (1992)) that $Y \succeq_{lz}^{s} X$ and thus $Y \succeq_{PL} X$ holds, but with $\varphi(z_1, z_2) = \max\{z_1, z_2, 1\}$, which is a convex function, we obtain

$$\sum_{i=1}^{n} \varphi\left(\frac{x_{i1}}{\bar{x}_1}, \frac{x_{i2}}{\bar{x}_2}\right) = 7 < 8 = \sum_{i=1}^{n} \varphi\left(\frac{y_{i1}}{\bar{y}_1}, \frac{y_{i2}}{\bar{y}_2}\right),$$

i.e., $X \not\succeq_{cx}^{s} Y$.

For details and many more properties of the three orderings, see Mosler (2002, Ch 9-10).

4 Restricted multivariate dominance

Now we turn to the task of restricting the multivariate orderings to the 'essential' or 'central' parts of the distributions. The idea is that only those statistical units should be considered which have values not too far from the mean. In the univariate case an interval between two given quantiles has been used above. As income distributions (as well as distributions of many other attributes of well-being) are asymmetric, this interval is generally not symmetric to the mean.

In the multivariate case we rely on recent notions of interquantile regions, that is central sets, and adapt them to our setting of restricted inequality comparisons. As the joint distribution of income and other attributes of well-being is mostly dependent and asymmetric, we restrict the comparison to central regions of the distributions that are not necessarily balls or ellipsoids. We employ a theorem which characterizes the convex-linear scaled ordering in terms of certain central sets.

Let $X = [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n]$, $\quad Y = [\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n]$ be given data matrices, $\boldsymbol{x}_i = (x_{i1}, ..., x_{id})'$,
$\boldsymbol{y}_i = (y_{i1}, ..., y_{id})' \in \mathbb{R}^d_+$,

$$\boldsymbol{r}_i = \left(\frac{x_{i1}}{\bar{x}_1}, \frac{x_{i2}}{\bar{x}_2}, \ldots, \frac{x_{id}}{\bar{x}_d} \right), \qquad \boldsymbol{s}_i = \left(\frac{y_{i1}}{\bar{y}_1}, \frac{y_{i2}}{\bar{y}_2}, \ldots, \frac{y_{id}}{\bar{y}_d} \right).$$

Proposition 1:

$$X \succeq^s_{lz} Y \quad \Leftrightarrow \quad B_t(X) \supset B_t(Y) \quad \text{for all} \quad 0 \le t \le 1\,,$$

where

$$B_t(X) = \left\{ \boldsymbol{z} \in \mathbb{R}^d_+ : \boldsymbol{z} = \sum_{i=1}^n \lambda_i \boldsymbol{r}_i\,,\ \sum_{i=1}^n \lambda_i = 1\,,\ 0 \le t\lambda_i \le \frac{1}{n} \right\}.$$

For proof, see Def. 8.4 and Prop. 8.1 in Mosler (2002).

The set $B_t(X)$ is a central region of the scaled distribution of X, called *zonoid central region* (Koshevoy and Mosler, 1997). Note that the sets are convex, nested and decreasing with t. For $t = 0$, $B_0(X)$ equals the convex

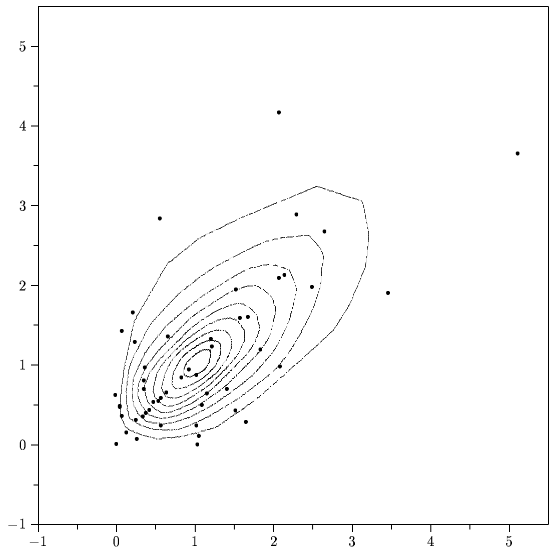


Figure 3: Zoonoid trimmed regions $B_t(X)$, $t = 0.1, \ldots, 0.9$, for $n = 50$ observations.

hull of the points $\boldsymbol{r}_1, \ldots, \boldsymbol{r}_n$. For $t = 1$, we have $B_1(X) = \{(1, \ldots, 1)\}$. Given the data X, the set $B_t(X)$ may be regarded as a region that includes a central part of the (scaled) data, and t indicates its degree of centrality.

Figure 3 exhibits the zonoid central regions of 50 simulated data points in $\mathbb{R}^2$. It illustrates how the regions reflect the asymmetric shape of the data distribution. For the computation of zonoid central regions, see Dyckerhoff (2000).

Similarly, the PL-dominance is characterized by the inclusion of certain regions.

PROPOSITION 2:

$$X \succeq_{PL} Y \quad \Leftrightarrow \quad C_t(X) \supset C_t(Y) \quad \text{for all} \quad 0 \leq t \leq 1\,,$$

where

$$C_t(X) = \left\{ \boldsymbol{z} \in \mathbb{R}^d_+ : \boldsymbol{z} \geq \sum_{i=1}^n \lambda_i \boldsymbol{r}_i\,,\ \sum \lambda_i = 1,\ 0 \leq t\lambda_i \leq \frac{1}{n} \right\}.$$

This proposition can, with some calculations, be derived from Theorem 9.7 in Mosler (2002). Note that the region $C_t(X)$ includes the zonoid central region $B_t(X)$ and, in addition, all points that are larger with respect to the componentwise order of $\mathbb{R}^d$.

When $d = 1$, for any $t \in [0, 1]$ holds

$$\begin{aligned} B_t(X) &= \left[\frac{1}{t}L_X(t),\, \frac{1}{t}(1 - L_X(1-t))\right], \\ C_t(X) &= \left[\frac{1}{t}L_X(t),\, \infty\right], \end{aligned}$$

hence

$$B_t(X) \supset B_t(Y) \;\Leftrightarrow\; L_X(t) \le L_Y(t) \quad \text{and} \quad L_X(1-t) \le L_Y(1-t)\,, \tag{9}$$

$$C_t(X) \supset C_t(Y) \;\Leftrightarrow\; L_X(t) \le L_Y(t)\,. \tag{10}$$

Consequently, in the univariate case, the convex-linear scaled dominance as well as the the price Lorenz dominance coincide with the usual Lorenz dominance.

As we have explained above, in comparing inequality by dominance relations, it makes sense to restrict the comparison to the central parts of the distributions. In view of the Propositions 1 and 2, this is readily done with the multivariate dominance relations $\succeq_{lz}^{s}$ and $\succeq_{PL}$.

Definition 3 **(Multivariate restricted dominance)** Let $d \ge 1$. For given $t_0 \ge 0$ define

(i) the *restricted convex-linear scaled dominance*
$X \succeq_{lz}^{s;t_0} Y \quad \text{if} \quad B_t(X) \supset B_t(Y) \quad \text{for} \quad t \ge t_0\,,$

(ii) the *restricted price Lorenz dominance*
$X \succeq_{PL}^{t_0} Y \quad \text{if} \quad C_t(X) \supset C_t(Y) \quad \text{for} \quad t \ge t_0\,.$

When $d = 1$, by (9) and (10) these two orderings specialize to univariate restricted dominance relations as follows:

$$\begin{aligned} X \succeq_{lz}^{s;t_0} Y &\;\Leftrightarrow\; X \succeq_{L}^{t_0,1-t_0} Y\,, \\ X \succeq_{PL}^{t_0} Y &\;\Leftrightarrow\; X \succeq_{L}^{t_0} Y\,. \end{aligned}$$

5 Restricted multivariate non-scaled dominance

The multivariate orderings considered so far are invariant to transformations of scale. It may be argued that a measure of multi-dimensioned economic status should, beyond changes of the units of measurement in the attributes, not be scale invariant, since substitutions are possible. Once the units of measurement of each attribute are fixed, an increase or decrease in one attribute can affect the valuation of the other attributes.

The dominance relations $\succeq^s_{cx}$, $\succeq^s_{lz}$ and $\succeq_{PL}$ have been defined on the basis of data scaled down by their means. Now, in a similar way, we introduce two dominance relations which are not scale invariant and based on the non-scaled data instead.

The univariate orderings of generalized L-functions are characterized through monotone convex evaluation functions (e.g. Marshall and Olkin (1979, p 109)):

$$X \succeq_{dcx} Y \quad \Longleftrightarrow \tag{11}$$

$$\sum_{i=1}^{n} \varphi(x_i) \geq \sum_{i=n}^{n} \varphi(y_i) \quad \text{for all decreasing convex } \varphi : \mathbb{R} \to \mathbb{R}\,,$$

$$X \succeq_{icx} Y \quad \Longleftrightarrow \tag{12}$$

$$\sum_{i=1}^{n} \varphi(x_i) \geq \sum_{i=n}^{n} \varphi(y_i) \quad \text{for all increasing convex } \varphi : \mathbb{R} \to \mathbb{R}\,,$$

$$\tag{13}$$

Definition 4 **(Multivariate non-scaled dominance)** Define, for $d \geq 1$,

(i) the *decreasing (increasing) convex dominance* $X \succeq_{dcx} Y$ ($X \succeq_{icx} Y$) if

$$\sum_{i=1}^{n} \varphi(x_{i1}, x_{i2}, \ldots, x_{id}) \geq \sum_{i=1}^{n} \varphi(y_{i1}, y_{i2}, \ldots, y_{id}) \tag{14}$$

for all decreasing (increasing) convex functions $\varphi : \mathbb{R}^d \to \mathbb{R}$,

(ii) the *decreasing (increasing) convex-linear dominance* $X \succeq_{dlz} Y$ if (14) holds for all convex-linear φ of form (8) with ψ decreasing (increasing) convex and $\alpha_1, \dots, \alpha_d \geq 0$.

Again, these relations are reflexive, transitive, and antisymmetric. $X \succeq_{dcx}$ implies $X \succeq_{dlz} Y$ and $X \succeq_{icx}$ implies $X \succeq_{ilz} Y$. The relation $\succeq_{dlz}$ is mentioned as the *decreasing convex-linear dominance*, while $\succeq_{ilz}$ is named the *increasing convex-linear dominance*. These dominance relations are also characterized by the inclusion of certain regions as follows.

PROPOSITION 3:

$$X \succeq_{dlz} Y \quad \Leftrightarrow \quad E_t(X) \supset E_t(Y) \quad \text{for all} \quad 0 \leq t \leq 1\,,$$

where

$$E_t(X) = \left\{ \boldsymbol{z} \in \mathbb{R}^d_+ : \boldsymbol{z} \geq \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i\,,\ \sum \lambda_i = 1,\ 0 \leq t\lambda_i \leq \frac{1}{n} \right\}.$$

Proof: Note that the ordering $X \succeq_{dlz} Y$ is the same as price supermajorization, which is characterized by the ordering of the inverse generalized Lorenz functions, $gl_X(\boldsymbol{z}) \geq gl_Y(\boldsymbol{z})$ for all $\boldsymbol{z} \in \mathbb{R}^d_+$; see Mosler (2002, pp 243f). A short calculation shows that $E_t(X) = \frac{1}{t}\{\boldsymbol{z} : gl_X(\boldsymbol{z}) \geq t\}$. This proves the proposition. □

Similarly, there holds:

PROPOSITION 4:

$$X \succeq_{ilz} Y \quad \Leftrightarrow \quad F_t(X) \supset F_t(Y) \quad \text{for all} \quad 0 \leq t \leq 1\,,$$

where

$$F_t(X) = \left\{ \boldsymbol{z} \in \mathbb{R}^d_+ : \boldsymbol{z} \leq \sum_{i=1}^{n} \lambda_i \boldsymbol{x}_i\,,\ \sum \lambda_i = 1,\ 0 \leq t\lambda_i \leq \frac{1}{n} \right\}.$$

In the univariate case ($d = 1$) we obtain

$$E_t(X) = \left[\frac{1}{t} GL_X(t),\, \infty\right], \quad F_t(X) = \left[-\infty,\, \frac{1}{t}(\overline{x} - GL_X(1-t)\right]$$

and, therefore,

$$\begin{aligned} E_t(X) \supset E_t(Y) \quad &\Leftrightarrow \quad GL_X(t) \le GL_Y(t)\,, \\ F_t(X) \supset F_t(Y) \quad &\Leftrightarrow \quad \overline{x} - GL_X(1-t) \ge \overline{y} - GL_Y(1-t)\,. \end{aligned}$$

In order to focus the inequality comparison on the central parts of the distributions, these set inclusions, again, are restricted to parameters t in some subinterval of $[0, 1]$. The restricted version of the multivariate decreasing convex-linear dominance is obtained in the following straightforward way.

Definition 5 **(Multivariate restricted non-scaled dominance)** Let $d \ge 1$. For given $0 \le t_0 < t_1 \le 1$ define the *restricted decreasing convex-linear dominance,*

$$X \succeq_{dlz}^{t_0,t_1} Y \quad \text{if} \quad E_t(X) \supset E_t(Y) \quad \text{for} \quad t_0 \le t \le t_1\,,$$

and the *restricted increasing convex-linear dominance,*

$$X \succeq_{ilz}^{t_0,t_1} Y \quad \text{if} \quad F_t(X) \supset F_t(Y) \quad \text{for} \quad t_0 \le t \le t_1\,.$$

In the univariate case, these orderings coincide with the restricted decreasing (resp. increasing) convex dominance, that is, the ordering of generalized Lorenz functions: If $d = 1$,

$$\begin{aligned} X \succeq_{dlz}^{t_0,t_1} Y \quad &\Leftrightarrow \quad GL_X(t) \le GL_Y(t) \quad \text{for } t_0 \le t \le t_1\,. \\ X \succeq_{ilz}^{t_0,t_1} Y \quad &\Leftrightarrow \quad \overline{x} - GL_X(t) \le \overline{y} - GL_Y(t) \quad \text{for } t_0 \le t \le t_1\,. \end{aligned}$$

Note that, with these non-scaled restricted orderings, the lower and the upper bound can be independently chosen. The actual choice is to be based on the specific economic question under inquiry and on the availability and quality of the data.

6 Concluding remarks

To measure multi-dimensioned inequality, four new notions of restricted Lorenz dominance have been introduced: two of them, the convex-linear dominance and the price Lorenz dominance compare distributions of relative well-being (ratio to the mean), while the other two, the decreasing and the increasing convex-linear dominances, compare distributions of absolute well-being and reflect also differences in the mean values. All four are quasi-orderings; the first dominance relation implies the second one, and the second implies the remaining two relations. Each unrestricted dominance relation is based on set inclusions of a certain family of central regions. A *restricted* dominance relation is obtained by restricting the set inclusions to a properly chosen inner part of the family. Thus, the restricted ordering focusses on a 'central' aspect of the distributions. By reducing the inner part, the ordering can be successively made finer.

The restricted orderings are also useful in comparing one-dimensioned distributions. Then they amount to the ordering of a central part of two Lorenz curves (resp. two generalized Lorenz curves) and neglect eventual crossings in their lower or upper ends.

The ordering of convex-linear dominance is rather close to that of convex dominance (which is also known as multivariate majorization). In dimension one, the two orderings coincide. But for $d \geq 2$ the convex scaled dominance $\succ^s_{cx}$, as well as the non-scaled relations $\succ_{dcx}$ and $\succ_{icx}$, cannot be restricted that way.

As the restricted multi-dimensioned orderings are stronger than the unrestricted ones (in particular, stronger than multivariate majorization), many real data sets which are not comparable by multivariate majorization may become comparable by properly restricted dominance relations.

It should be emphasized that the size and shape of a given central region depends on the whole distribution, and not only on the data included in that region. Therefore, our restricted dominance notions depend on *all data*, and not just on some central part of them. In particular, they should not be interpreted as unanimous utilitarian social welfare preference restricted to subsets of the populations.

In restricting ($d = 1$) the univariate Lorenz dominance we fix a proportion of poor households whose within-group inequality is not considered; similarly a proportion of rich households is fixed. In dimension $d \geq 2$, with the restricted convex-linear order we fix just one proportion of very poor and very rich people together whose within-group inequality we neglect.

As the multi-dimensioned distribution has no separate 'upper' and 'lower tails' there is only one parameter t_0 to choose.

We close with a few remarks on limitations and possible extensions of our approach.

- The above definitions extend to probability distributions and data with different sample sizes, and the same results are valid.
- Further multivariate inequality orderings can be based on other families of central regions, like Mahalanobis or halfspace central regions (Mosler, 2002), and others.
- The above multivariate inequality orderings imply that the attributes are substitutional. Moreover they treat the different dimensions in a symmetric way. No dimension is regarded as 'more important' than another with respect to inequality. Asymmetric orderings which are finer than the above convex-linear orderings can be developed by imposing further conditions on the coefficients α_j in (8). For multivariate indices, including those which give different weights to the dimensions, see e.g. Kolm (1977), Maasoumi (1986) and Koshevoy and Mosler (1997).

References

Atkinson, A. and Bourguignon, F. (1982). The comparison of multidimensioned distributions of economic status. *Review of Economic Studies* **49**, 183–201.

Cowell, F. (1995). *Measuring Inequality.* Prentice Hall, London, 2nd ed.

Dyckerhoff, R. (2000). Computing zonoid trimmed regions of bivariate data sets. In J. Bethlehem and P. van der Heijden, eds., *COMPSTAT 2000. Proceedings in Computational Statistics*, 295–300. Physica-Verlag, Heidelberg.

Elton, J. and Hill, T. (1992). Fusions of probability distributions. *Annals of Probability* **20**, 421–454.

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

Koshevoy, G. (1995). Multivariate Lorenz majorization. *Social Choice and Welfare* **12**, 93–102.

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

Koshevoy, G. and Mosler, K. (1997). Multivariate Gini indices. *Journal of Multivariate Analysis* **60**, 252–276.

Koshevoy, G. and Mosler, K. (1998). Lift zonoids, random convex hulls and the variability of random vectors. *Bernoulli* **4**, 377–399.

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

Maasoumi, E. (1999). Multidimensioned approaches to welfare analysis. In J. Silber, ed., *Handbook of Income Inequality Measurement*. Kluwer, Dordrecht.

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

Mosler, K. (1994a). Majorization in economic disparity measures. *Linear Algebra and its Applications* **199**, 91–114.

Mosler, K. (1994b). Multidimensional welfarisms. In W. Eichhorn, ed., *Models and Measurement of Welfare and Inequality*, 808–820. Springer, Berlin.

Mosler, K. (2002). *Multivariate Dispersion, Central Regions and Depth: The Lift Zonoid Approach.* Springer, New York.

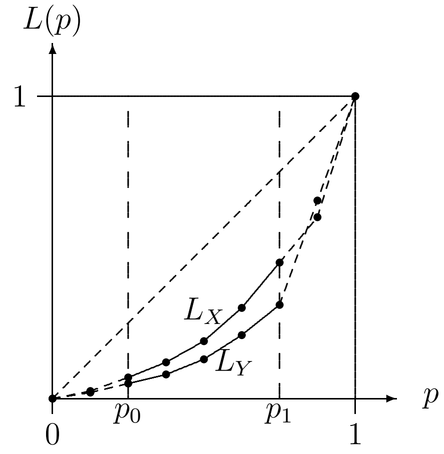
Piesch, W. (1975). *Statistische Konzentrationsmaße.* J.C.B. Mohr (Paul Siebeck), Tübingen.

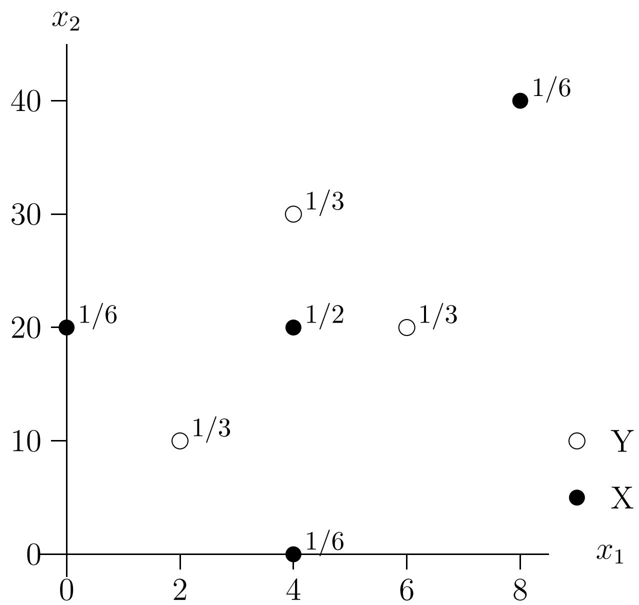
Sen, A. K. (1970). *Collective Choice and Social Welfare.* Norton, New York.

Sen, A. K. (1973). *On Economic Inequality.* Oxford University Press, Oxford.

Tobin, J. (1970). On limiting the domain of inequality. *Journal of Law and Economics* **13**, 263–277.

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


L(p)
1
L_X
L_Y
p
p_0
p_1
0
1



x_2
40
30
20
10
0
1/6
1/3
1/6
1/2
1/3
1/3
1/6
○ Y
● X
0
2
4
6
8
x_1