Majorization in Economic Disparity Measures

Karl Mosler

Fachbereich Wirtschafts- und Organisationswissenschaften
Universität der Bundeswehr Hamburg
D-22039 Hamburg, Germany

Submitted by Friedrich Pukelsheim

Dedicated to Ingram Olkin

ABSTRACT

This survey presents an account of univariate and multivariate majorization orderings and their characterization by various classes of economic disparity indices. First, a concise treatment of classical univariate results is given, including majorization with different means and different population sizes, as well as Lorenz orderings of relative and absolute disparity. Second, alternatives to the Pigou-Dalton principle of transfers are discussed which are based on transfers about a given threshold. Third, disparity in several attributes and multivariate majorization are investigated, and a multivariate version of the Lorenz curve is introduced.

1. INTRODUCTION

Consider a population of n economic units, $i = 1, \dots, n$, each of which is endowed with a quantity a_i of affluence. We will speak of i as a household and of a_i as its annual income. $(a_1, \dots, a_n)$ is called an income vector or an *income distribution*. However, i may also denote another economic unit like an individual person or a country, and a_i another attribute of economic status like endowment with some commodity. When all a_i are equal, obviously, the disparity of income is minimum, say 0, in the population. But when some a_i are different, the problem arises of measuring the degree of disparity. There are two basic questions: First, find real-valued functions which are meaningful indices of the disparity of $(a_1, \dots, a_n)$. Second, given two income vectors $(a_1, \dots, a_n)$ and $(b_1, \dots, b_n)$, decide whether one of them contains more disparity than the other.

655 Avenue of the Americas, New York, NY 10010 0024-3795/94/$7.00

Since the beginning of this century, economists have been interested in the quantitative description and statistical estimation of economic disparity. The ideas of Lorenz (1905), Gini (1912), Pigou (1912), and Dalton (1920) are closely connected with the concept of majorization introduced by Hardy, Littlewood, and Polya (1929, 1934). But it took until the late sixties for the results of Hardy, Littlewood, and Polya to enter the economic literature. Kolm (1969), Atkinson (1970), Das Gupta et al. (1973), and Fields and Fei (1978) introduced the mathematical notions of majorization and Lorenz order to economic theory; see also Kakwani (1977) and Blackorby and Donaldson (1984).

Many other authors have contributed to these topics whom we cannot mention here. We refer the reader to several monographs which cover the developments on the mathematics and the economics side. First of all, there is Marshall and Olkin's famous book on majorization (Marshall and Olkin, 1979), which, besides many new results, includes a comprehensive treatment of the mathematical literature before 1979. It also contains a history of the field. Arnold (1987) provides a nice introduction to majorization and the Lorenz order. The economic theory of disparity indices and orderings is exhibited in the classical books by Sen (1973) and Cowell (1977) and, more recently, in Chakravarty (1990a). A history of economic-disparity measurement can also be found in Arnold (1983, Chapter 1).

This survey presents an account of univariate and multivariate majorization as far as they are relevant to the analysis of economic disparity. First, a concise treatment of the classical univariate results is given. Second, some departures from the Pigou-Dalton principle of transfers are discussed. Third, multiattribute disparity is investigated, and a multivariate version of the Lorenz curve is introduced.

Section 2 starts with the notions of Pigou-Dalton transfers and majorization and their characterization by various classes of disparity indices. Ordinary majorization between vectors in $\mathbb{R}^n$ implies that the vectors have equal means. This corresponds to a transfer of incomes in a fixed population. But in many economic applications, income vectors are compared which have different means and different population sizes. In Section 2.3 we discuss growing and shrinking transfers and weak majorization. Section 3 is about relative and absolute disparity measurement. Two Lorenz orderings are given which compare arbitrary income distributions with respect to their relative and absolute disparity. In Section 4 strict and semistrict notions of disparity indices are considered as well as a class of indices which is larger than the S-convex functions. The notions are based on transfers about a given threshold θ, so-called transfers about θ and star-shaped transfers at θ.

Economic disparity does not arise from the distribution of income alone. There are attributes of affluence and well-being besides annual household income: housing equity, financial assets, free time, education, and many others. In modern theories of social choice the specific distributional inequality of attributes like these is considered. In Section 5 we will investigate disparity in several attributes and its relation to multivariate majorization. An account of the mathematical and economic literature on the multidimensioned case will be given there. Section 6 surveys multivariate versions of the Lorenz curve, including a new notion which is based on an idea of Koshevoy (1992). Section 7 concludes the paper.

Most of the material in this survey is known, and many proofs are already contained in Marshall and Olkin (1979). Other material, especially on the economics side, is found rather dispersed in the literature. New results mainly concern modified principles of transfer, multiattribute economic disparity, and multivariate Lorenz order. Some proofs of known results are provided for expository reasons.

There are important aspects of our topic which we do not cover. Some of them have been the subject of recent publications. The preservation and, even more, the attenuation of majorization and Lorenz order has applications in social choice theory, especially in the design of taxing systems (Fellman, 1976; Jakobson, 1976; Eichhorn et al., 1984). They are surveyed in Moyes (1989) and Arnold (1991). Stochastic orders other than classical majorization and their applications to welfare economics are treated in Le Breton (1991) and Mosler (1993). Chakravarty (1990a) includes a comprehensive treatment of disparity and welfare indices and of their axiomatizations. For treatments of special indices such as those due to Gini, Theil, Atkinson, and others, we refer the reader to Piesch (1975), Cowell (1977), and Nygård and Sandström (1981).

Some notation: $\mathbb{R}^n$ denotes the n-space of column vectors, $\mathbb{R}_n$ that of row vectors, $\mathbb{R}^n_+$ and $\mathbb{R}_{n+}$ the subsets of vectors having nonnegative components only. S_n is the unit simplex in $\mathbb{R}^n$, and $\mathbb{R}^{m\times n}$ is the set of (m,n) matrices. x^T denotes the transpose of x. A matrix $A = (a_{ik}) \in \mathbb{R}^{m\times n}$ is called *column stochastic* iff $\sum_{i=1}^m a_{ik} = 1$ holds for every k. It is called *doubly stochastic* iff, in addition, $\sum_{k=1}^n a_{ik} = 1$ holds for all i. The set of column stochastic matrices is denoted by $\mathcal{C}_{m,n}$, and the set of doubly stochastic matrices by $\mathcal{D}_{m,n}$. $\mathcal{P}_n$ is the set of (n,n) permutation matrices. *Increasing* means nondecreasing, and *decreasing* means nonincreasing. For $a \in \mathbb{R}^n$, let $a_{(\cdot)} = (a_{(1)}, \dots, a_{(n)})^T$ be the ordered vector where the components have been rearranged in increasing order.

2. MAJORIZATION IN THE UNIVARIATE CASE

In this section we give a short account of transfer principles and majorization, and of their economic interpretations. Classes of disparity indices are given which characterize the orderings of majorization and weak majorization. Most of the proofs can be already found in Marshall and Olkin (1979). We compare income vectors which have the same number of components. Majorization between vectors having different dimensions is investigated in Sections 3 and 5.3.

2.1. *Transfers*

Consider $\mathcal{T}_0^n = \{(a,b) : a,b \in \mathbb{R}^n, a_{(\cdot)} \neq b_{(\cdot)}$, and $Pa = a_{(\cdot)}, Pb = b_{(\cdot)}$ for some $P \in \mathcal{P}_n\}$. $\mathcal{T}_0^n$ contains all pairs of vectors which are not equal and have the same order of components. For $(a,b) \in \mathcal{T}_0^n$ define $h \in \mathbb{R}^n, h_i = b_{(i)} - a_{(i)}$. A pair $(a,b) \in \mathcal{T}_0^n$ is called a *transfer* from a to b iff a and b have the same total, i.e., iff $\sum_{i=1}^n h_i = 0$.

Given a set $\mathcal{T} \subset \mathcal{T}_0^n$, we say that a function $\phi : \mathbb{R}^n \to \mathbb{R}$ *satisfies the $\mathcal{T}$-principle of transfers* iff

$$\phi(a) \geq \phi(b) \qquad \text{whenever} \qquad (a,b) \in \mathcal{T}. \tag{2.1}$$

A transfer (a,b) is called a *Pigou-Dalton transfer* (briefly, a *PD transfer*) iff the first k elements, $1 \leq k \leq n-1$, of h are nonnegative and the remaining $n-k$ elements are nonpositive. Roughly speaking, a Pigou-Dalton transfer is a transfer from some households which are "relatively rich" to some which are "relatively poor" such that both the total income and the order among the household incomes remain unchanged. We denote the set of all PD transfers by $\mathcal{T}_{\text{PD}}^n$.

A real-valued function ϕ defined on $\mathbb{R}^n$ is a disparity index satisfying the *Pigou-Dalton principle of transfers* iff $\phi(a) \geq \phi(b)$ whenever $(a,b) \in \mathcal{T}_{\text{PD}}^n$. See Pigou (1912) and Dalton (1920).

An *elementary PD transfer* is a transfer with h having just two nonzero elements. The set of elementary PD transfers is denoted by $\mathcal{T}_{\text{ePD}}^n$. It is obvious that the condition (2.1) with $\mathcal{T} = \mathcal{T}_{\text{PD}}^n$ implies the same with $\mathcal{T} = \mathcal{T}_{\text{ePD}}^n$. But the converse is also true, since every PD transfer can be decomposed into a finite number of elementary PD transfers.

2.2. *Majorization*

Let $a, b \in \mathbb{R}^n$. We say that *a majorizes b*, $a \succ b$, iff one of the following four equivalent conditions is fulfilled:

$$(Pa, b) \in \mathcal{T}_{\text{PD}}^n \quad \text{for some} \quad P \in \mathcal{P}_n, \tag{2.2}$$

$$b = Ta \quad \text{for some} \quad T \in \mathcal{D}_{n,n}, \tag{2.3}$$

$$b \in \text{conv}\{Pa : P \in \mathcal{P}_n\}, \tag{2.4}$$

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i \quad \text{and} \quad \sum_{i=1}^k a_{(i)} \leq \sum_{i=1}^k b_{(i)} \quad \text{for } k = 1, \ldots, n-1. \tag{2.5}$$

According to (2.2), b is majorized by a iff it is the result of a PD transfer from a permutation of a; according to (2.3), iff it is the result of a doubly stochastic transformation of a. T in (2.3) is not unique. Brualdi (1984) investigates the polytope of all such T and determines its dimension. (2.3) implies that b is an *average* of a, i.e., $b = Ta$ with a row-stochastic T. When we think of a transfer between households leading from income distribution a to distribution b, t_{ij} represents the share of j which goes to i. (2.4) says that b is a convex combination of permutations of a. (Note that the permutations of a are equally ordered under majorization.) When a and b are nonnegative vectors, (2.5) means that the Lorenz curve of a lies below that of b.

A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is called *S-convex* iff $a \succ b$ implies $\phi(a) \geq \phi(b)$. In particular, an S-convex function is symmetric in its arguments. Obviously,

PROPOSITION 2.1. *$\phi : \mathbb{R}^n \to \mathbb{R}$ is S-convex if and only if ϕ is symmetric and satisfies the PD principle of transfers.*

Thus, the set of S-convex functions is the natural class of disparity indices which respect anonymity, i.e., do not distinguish between the households, and satisfy the PD principle of transfers. Moreover, if a is more dispersed than b in terms of every S-convex index ϕ, it follows that $a \succ b$. There are other classes of indices which do the same:

PROPOSITION 2.2. *Let $a, b \in \mathbb{R}^n$. Then $a \succ b$ is equivalent to each of the following conditions:*
(i) *$\phi(a) \geq \phi(b)$ for all ϕ which are S-convex,*
(ii) *$\phi(a) \geq \phi(b)$ for all ϕ which are symmetric and quasiconvex,*
(iii) *$\phi(a) \geq \phi(b)$ for all ϕ which are symmetric and convex,*
(iv) *$\sum_{i=1}^n g(a_i) \geq \sum_{i=1}^n g(b_i)$ for all $g : \mathbb{R} \to \mathbb{R}$ which are convex.*

The proposition is well known. In Section 5, Proposition 5.1, it is extended to multivariate majorization. Further equivalent conditions in terms of index classes are obtained from Propositions 2.3 to 2.5 below.

In Proposition 2.2, five classes of disparity indices ϕ are given, each of which induces the preorder $\succ$. By definition, the set of S-convex functions is the biggest one, viz. the set of all functions which are $\succ$-increasing. If majorization is regarded as the basic notion of being more unequal, every meaningful disparity index ϕ has to be S-convex and, in particular, row symmetric. Usually, properties of disparity indices are traced back to axioms. See, e.g., Fields and Fei (1978), Chakravarty (1990a). Symmetry of ϕ is based on the axiom of anonymity, which says that a permutation of households does not change inequality. Quasiconvexity of ϕ means that, if a and b have the same disparity, $\lambda a + (1 - \lambda)b$ has no more, $0 \leq \lambda \leq 1$.

The additive decomposition of ϕ, $\phi(x) = \sum_{i=1}^n g(x_i)$, is based on anonymity and either a utilitarian axiom or an axiom of nonaltruism; see Mosler (1993). Convexity of g can be interpreted in a framework of decision making under risk: Social states are evaluated by a subject who considers himself to occupy each position in the given population with equal probability and who orders the states according to their expected value of individual disutility g. Then, convexity of g is tantamount to risk aversion of the subject.

For further discussions of these classes of disparity indices, the reader is referred to Das Gupta et al. (1973) and Rothschild and Stiglitz (1973).

2.3. Growing and Shrinking Transfers

When two income distributions are compared, their total incomes may be different. For example, the comparison may involve a time interval during which the cake grows. Or pre-tax and after-tax distributions are compared, and the taxation causes the total cake to shrink. For the rest of the section we restrict ourselves to nonnegative vectors. All results besides (2.9) and (2.12) hold also for vectors of arbitrary signs.

Assume $a, b \in \mathbb{R}^n_+$, $(a, b) \in \mathcal{T}_0^n$. We call (a, b) a *growing transfer* iff

$$\sum_{i=1}^{k} h_i \geq 0 \quad \text{for} \quad k = 1, \ldots, n, \tag{2.6}$$

where again $h_i = b_{(i)} - a_{(i)}$. Similarly, $(a, b) \in \mathcal{T}_0^n$ is a *shrinking transfer* iff

$$\sum_{i=k}^{n} h_i \leq 0 \quad \text{for} \quad k = 1, \ldots, n. \tag{2.7}$$

Let $\mathcal{T}_{\text{grow}}$ and $\mathcal{T}_{\text{shoi}}$ denote the respective sets of transfers. We may think of a poverty line which separates the population into a "poorer" and a "richer" part. With a growing transfer, the poorer part always betters itself by a positive total amount, while with a shrinking transfer, the richer part always has to pay, wherever the poverty line is drawn.

We say that *a weakly supermajorizes b*, $a \succ^w b$, iff one of the following three equivalent conditions is fulfilled:

$$(Pa, b) \in \mathcal{T}_{\text{grow}} \quad \text{for some} \quad P \in \mathcal{P}_n, \tag{2.8}$$

$$b = Ta \quad \text{for some doubly superstochastic} \quad T, \tag{2.9}$$

$$\sum_{i=1}^{k} a_{(i)} \leq \sum_{i=1}^{k} b_{(i)} \quad \text{for} \quad k = 1, \ldots, n. \tag{2.10}$$

We say that *a weakly submajorizes b*, $a \succ_w b$, iff any of the following three holds:

$$(Pa, b) \in \mathcal{T}_{\text{shoi}} \quad \text{for some} \quad P \in \mathcal{P}_n, \tag{2.11}$$

$$b = Ta \quad \text{for some doubly substochastic} \quad T, \tag{2.12}$$

$$\sum_{i=k}^{n} a_{(i)} \geq \sum_{i=k}^{n} b_{(i)} \quad \text{for} \quad k = 1, \ldots, n. \tag{2.13}$$

(2.8) [(2.11)] means that b is a growing [shrinking] transfer of a permutation of the a_i's. Equation (2.9) [(2.12)] implies that b is an average of a using weights which add up to more [less] than unity. (2.12) says that the generalized Lorenz function (see Section 3.2) of a is bounded above by that of b, and (2.13) says the same for a dual notion of the generalized Lorenz function. Obviously, when total incomes are equal, both notions of weak majorization and that of majorization coincide:

PROPOSITION 2.3. *$a \succ^w b$ and $\sum_i a_i = \sum_i b_i \;\Leftrightarrow$ $a \succ_w b$ and $\sum_i a_i = \sum_i b_i \;\Leftrightarrow\; a \succ b$.*

Classes of disparity indices which induce the relations $\succ^w$ and $\succ_w$ on $\mathbb{R}^n_+$, respectively, are given by the following two propositions. They largely parallel Proposition 2.2.

PROPOSITION 2.4. *Let $a, b \in \mathbb{R}^n_+$. Then $a \succ^w b$ is equivalent to each of the following conditions:*

(i) $\phi(a) \geq \phi(b)$ *for all ϕ which are decreasing and S-convex,*

(ii) $\phi(a) \geq \phi(b)$ *for all ϕ which are decreasing, symmetric, and quasiconvex,*

(iii) $\phi(a) \geq \phi(b)$ *for all ϕ which are decreasing, symmetric, and convex,*

(iv) $\sum_{i=1}^n g(a_i) \geq \sum_{i=1}^n g(b_i)$ *for all $g : \mathbb{R} \to \mathbb{R}$ which are decreasing and convex,*

(v) $\sum_{i=1}^n \min\{a_i, \gamma\} \leq \sum_{i=1}^n \min\{b_i, \gamma\}$ *for all $\gamma \in \mathbb{R}$.*

Proof. If $a \succ^w b$, there exists $c \in \mathbb{R}^n$ such that $a \succ c$ and $c \leq b$, which is easily seen by induction on n (Marshall and Olkin, 1979, p. 123). For every decreasing and S-convex ϕ it follows that $\phi(a) \geq \phi(c) \geq \phi(b)$; hence (i). Observe that every symmetric and quasiconvex ϕ is S-convex, every convex ϕ is quasiconvex, and $\phi(x) = \sum_i g(x_i)$ is decreasing, symmetric, and convex when g is decreasing and convex. Therefore, (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) holds. Further, since for every $\gamma \in \mathbb{R}$, $x \mapsto -\min\{x_i, \gamma\}$ is a decreasing convex function, (iv) implies (v). On the other hand, (v) says the same as (2.10); hence (v) is equivalent to $a \succ^w b$. ■

PROPOSITION 2.5. *Let $a, b \in \mathbb{R}^n_+$. Then $a \succ_w b$ is equivalent to each of the following conditions:*

(i) $\phi(a) \geq \phi(b)$ *for all ϕ which are increasing and S-convex,*

(ii) $\phi(a) \geq \phi(b)$ *for all ϕ which are increasing, symmetric, and quasiconvex,*

(iii) $\phi(a) \geq \phi(b)$ *for all ϕ which are increasing, symmetric, and convex,*

(iv) $\sum_{i=1}^n g(a_i) \geq \sum_{i=1}^n g(b_i)$ *for all $g : \mathbb{R} \to \mathbb{R}$ which are increasing and convex,*

(v) $\sum_{i=1}^n \max\{a_i, \gamma\} \geq \sum_{i=1}^n \max\{b_i, \gamma\}$ *for all $\gamma \in \mathbb{R}$.*

The proof is similar to that of Proposition 2.4. Observe that Propositions 2.4 and 2.5 together with Proposition 2.3 yield another set of characterizations of $a \succ b$ in terms of classes of disparity indices ϕ.

3. DISPARITY INDICES AND LORENZ ORDERINGS

So far, we have compared vectors having the same number n of components. However, in many economic applications, nonnegative income vectors are compared which have different population sizes. In this section, we seek for disparity orderings and disparity indices by which vectors of any dimension can be compared and which have meaningful properties when different n's are involved.

3.1. *Relative and Absolute Disparity*

Let $Q_n = \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i > 0\}$, and $Q = \bigcup_{n\in\mathbb{N}} Q_n$. A disparity index on Q is a function $\phi : Q \to \mathbb{R}$ which follows some axioms. Axioms (A1) and (A2) seem rather natural.

(A1) For every $n \in \mathbb{N}$, the restriction $\phi|_{Q_n}$ is symmetric and satisfies the PD principle of transfers.
(A2) $\phi(x^{(k)}) = \phi(x)$ holds if $x \in Q_n$ and $n, k \in \mathbb{N}$, where $x^{(k)} = (x^T, x^T, \ldots, x^T)^T \in Q_{n\cdot k}$.

(A2) is called *population invariance.* It says that if the income vector is cloned k times, the disparity remains the same. We introduce

$$\Phi_0 = \{\phi : Q \to \mathbb{R} \ : \ \phi \text{ satisfies (A1) and (A2)}\}.$$

$\phi \in \Phi_0$ is called an *index of relative disparity* iff

(A3) $\phi(\beta x) = \phi(x)$ holds if $\beta > 0$, $x \in Q_n$, $n \in \mathbb{N}$.

ϕ is an *index of absolute disparity* iff

(A4) $\phi(x + \gamma \cdot \mathbf{1}) = \phi(x)$ holds if $\gamma \in \mathbb{R}$, $x \in Q_n$, $n \in \mathbb{N}$, where $\mathbf{1} = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$.

Let Φ^{rel} and Φ^{abs} denote the respective classes of indices. A nontrivial ϕ cannot be in both Φ^{rel} and Φ^{abs}. Dalton (1920) postulates that equal additions ($\gamma > 0$) should diminish economic inequality while equal subtractions ($\gamma < 0$) should increase it; thus, he argues against (A4). Kolm (1969) considers both kinds of indices and names them rightist and leftist indices, respectively. Bossert and Pfingsten (see Pfingsten, 1986) propose a continuum of axioms between (A3) and (A4):

(A5 λ) $\phi(x + \gamma[\lambda x + (1 - \lambda) \cdot \mathbf{1}]) = \varphi(x)$ holds if $\gamma \in \mathbb{R}$, $x \in Q_n$, $n \in \mathbb{N}$.

For every $\lambda \in [0, 1]$, this yields a class of λ-translation-scale invariant indices.

3.2. Lorenz Orderings

Let F be a probability distribution function on $\mathbb{R}_+$ having positive first moment μ_F, and define

$$\mathrm{GL}(F,t) = \int_0^\xi x\,dF(x), \quad \text{where} \quad t = \int_0^\xi dF(x), \;\; t \in [0,1],$$

or equivalently,

$$\mathrm{GL}(F,t) = \int_0^t F^{-1}(s)\,ds, \quad t \in [0,1],$$

where $F^{-1}(s) = \inf\{x : F(x) \geq s\}$. Then $L(F,\cdot) = \mathrm{GL}(F,\cdot)/\mu_F$ is the usual Lorenz function. $\mathrm{GL}(F,\cdot)$ is called the *generalized Lorenz function*. Observe that $\mathrm{GL}(F,\cdot)$ determines F in a unique way, while $L(F,\cdot)$ determines F up to a scale only. Consider $F^{\mathrm{rel}}(x) = F(\mu_F \cdot x)$, the distribution scaled down by μ_F, and $F^{\mathrm{abs}}(x) = F(x + \mu_F)$, the distribution shifted by μ_F.

Given two probability distribution functions F and G on $\mathbb{R}_+$, we say that F *majorizes* G, $F \succ G$, iff

$$\mathrm{GL}(F,t) \leq \mathrm{GL}(G,t) \quad \text{for all} \quad t \in [0,1]$$

and $\mathrm{GL}(F,t) = \mathrm{GL}(G,t)$ for $t = 1$. We define two Lorenz orderings for the comparison of relative and absolute disparity: the *relative Lorenz ordering*, where $F \succ_{\mathrm{LR}} G$ iff $F^{\mathrm{rel}} \succ G^{\mathrm{rel}}$, and the *absolute Lorenz ordering*, where $F \succ_{\mathrm{LA}} G$ iff $F^{\mathrm{abs}} \succ G^{\mathrm{abs}}$.

The relative Lorenz ordering is the usual Lorenz order; it corresponds to the (scale invariant) indices of relative disparity. The absolute Lorenz ordering corresponds to the (translation invariant) indices of absolute disparity. It follows readily from the definitions that in case $\mu_F = \mu_G$ the three orderings coincide:

PROPOSITION 3.1. *Let F and G be probability distribution functions on $\mathbb{R}_+$ with $\mu_F = \mu_G > 0$. Then $F \succ_{\mathrm{LR}} G$ if and only if $F \succ_{\mathrm{LA}} G$ if and only if $F \succ G$.*

Majorization between nonnegative vectors of different dimensions can be defined as follows. Given vectors $a \in \mathbb{R}_+^n$ and $b \in \mathbb{R}_+^m$, we define probability distribution functions F_a and F_b which put weight $1/n$ on each a_i and weight $1/m$ on each b_i, respectively. Then *majorization* is defined by $a \succ b$ iff $F_a \succ F_b$. In the same way, relative and absolute Lorenz ordering of vectors is defined: $a \succ_{\mathrm{LR}} b$ and $a \succ_{\mathrm{LA}} b$. The latter notions have been used in disparity measurement by Das Gupta, Sen, and Starrett (1973), Shorrocks (1983), and Moyes (1987).

PROPOSITION 3.2. [Das Gupta, Sen, and Starrett, 1973] *Let* $a \in \mathbb{R}_+^n$ *and* $b \in \mathbb{R}_+^m$. *Then*

(i) $a \succ b$ *if and only if* $\phi(a) \geq \phi(b)$ *for all* $\phi \in \Phi_0$,

(ii) $a \succ_{\text{LR}} b$ *if and only if* $\phi(a) \geq \phi(b)$ *for all* $\phi \in \Phi^{\text{rel}}$,

(iii) $a \succ_{\text{LA}} b$ *if and only if* $\phi(a) \geq \phi(b)$ *for all* $\phi \in \Phi^{\text{abs}}$.

Multivariate majorization with $n \neq m$ is investigated below in Section 5. For extensions of the Lorenz curve to the multivariate case, see Section 6.

4. OTHER PRINCIPLES OF TRANSFERS

In this section we discuss several departures from the PD principle of transfers.

4.1. Strict Principles

Let $\mathcal{T} \subset \mathcal{T}_0 \equiv \bigcup_{n \in \mathbb{N}} \mathcal{T}_0^n$. A function $\phi : Q \to \mathbb{R}$ satisfies the $\mathcal{T}$-principle of transfers iff (2.1) holds, i.e., $\phi(a) \geq \phi(b)$ whenever $(a, b) \in \mathcal{T}$. ϕ satisfies the *strict* $\mathcal{T}$*-principle of transfers* iff

$$(a, b) \in \mathcal{T} \quad \Rightarrow \quad \phi(a) > \phi(b). \tag{4.1}$$

ϕ is called *strictly S-convex* iff ϕ is symmetric and $\phi(a) > \phi(b)$ holds whenever $a, b \in \mathbb{R}^n$, $a \succ b$, b not in $\{Pa : P \in \mathcal{P}_n\}$, and $n \in \mathbb{N}$. It can be shown that ϕ is strictly S-convex if and only if ϕ is symmetric and satisfies the *strict PD principle of transfers*, i.e. (4.1) with $\mathcal{T} = \mathcal{T}_{\text{PD}} \equiv \bigcup_{n \in \mathbb{N}} \mathcal{T}_{\text{PD}}^n$ where $\mathcal{T}_{\text{PD}}^n$ denotes the set of PD transfers in $\mathbb{R}^n$. There holds a proposition analogous to Proposition 2.2 which relates the set of strictly S-convex functions to strict majorization and to other sets of strict disparity indices. We omit the details. Many common disparity indices are strictly S-convex, e.g., the variance, the coefficient of variation, and the indices of Gini, Theil, Atkinson, and others; see Piesch (1975) and Cowell (1977). An example of an index which is S-convex, but not in the strict sense, is the mean deviation about the mean.

Indices which are strictly increasing at some PD transfers and just increasing at the remaining ones have been investigated recently by Castagnoli and Muliere (1990); see Section 4.2 below.

4.2. Transfers about θ

Let $\theta \in \mathbb{R}$ be fixed, and define

$$\mathcal{T}_\theta = \{(a,b) \in \mathcal{T}_{\text{PD}} : a_{(i)} + h_i \leq \theta \text{ if } h_i > 0, a_{(i)} + h_i \geq \theta \text{ if } h_i < 0\}.$$

$\mathcal{T}_\theta$ consists of PD transfers which give some positive amount to individuals below θ and take it from individuals above θ. Every such transfer is called a *transfer about θ*.

The principle of transfers about θ can be read in the following way. θ may be interpreted as a line which separates two social classes, and the transfer about θ as an action taken by the government. Every transfer from a household above the line to a household under the line is considered as decreasing inequality (in the weak or strict sense). Note that a transfer about θ affects neither the relative order of households nor their positions above or below the line. That is, no household crosses the line: the poor remain in the lower class, and the rich in the upper class.

Disparity indices which are symmetric and satisfy the principle of transfers about θ are proposed by Mosler and Muliere (1993). These indices include the S-convex functions, but others as well. The idea is that transfers between rich people only and transfers between poor people only should not affect the index of disparity at all.

It can be shown that a differentiable function ϕ satisfies the principle of transfers about θ if and only if

$$\max_{x_i<\theta} \frac{\partial}{\partial x_i}\phi(x) \leq \min_{x_i>\theta} \frac{\partial}{\partial x_i}\phi(x) \quad \text{for all} \quad x. \tag{4.2}$$

Let $\mathcal{G}$ be the set of continuous functions $g : \mathbb{R}_+ \to \mathbb{R}$ which are differentiable everywhere besides a finite set of points where one-sided derivatives exist. Let g' denote the derivative when it exists. If

$$\phi(x_1, \ldots, x_n) = \sum_{i=1}^{n} g(x_i) \quad \text{with some} \quad g \in \mathcal{G}, \tag{4.3}$$

the condition (4.2) reads

$$g'(s) \leq g'(t) \quad \text{whenever} \quad s < \theta < t. \tag{4.4}$$

For (4.4) we say that ϕ has *increasing disparity weight about θ*.

Mosler and Muliere (1993) investigate a second set of transfers. With these transfers, certain households may cross the line from poor to rich or vice versa.

The crossings are restricted to the income interval of those households that before the transfer were situated next to the line:

$$\mathcal{T}_{\text{next}\,\theta} = \{(a,b) \in \mathcal{T}_{\text{PD}} : h_1, \dots, h_k \geq 0, h_{k+1}, \dots, h_n \leq 0,$$
$$a_{(k)} \leq \theta \leq a_{(k+1)} \quad \text{and} \quad a_{(k)} \leq a_{(i)} + h_i \leq a_{(k+1)} \text{ if } h_i \neq 0\},$$
$$\mathcal{T}_{\text{star}\,\theta} = \mathcal{T}_{\text{next}\,\theta} \cup \mathcal{T}_\theta.$$

$\mathcal{T}_{\text{next}\,\theta}$ is called the set of *transfers next to* θ, $\mathcal{T}_{\text{star}\,\theta}$ the set of *star-shaped transfers at* θ. The latter name stems from the fact that the $\mathcal{T}_{\text{star}\,\theta}$-principle of transfers corresponds to the class of disparity indices which are additively separable (4.3) with g star-shaped above at θ; see below.

Let I be an interval in $\mathbb{R}_+, \theta \in \mathbb{R}$. A function $f : I \to \mathbb{R}$ is said to be star-shaped above at θ and supported iff

$$\frac{f(s) - f(\theta)}{s - \theta} \quad \text{is increasing at all} \quad s \in I - \{\theta\}.$$

Briefly, we say that such an f is *star-shaped above at* θ. If f is differentiable, equivalently,

$$\frac{f(s) - f(\theta)}{s - \theta} \begin{cases} \geq f'(s) & \text{when} \quad s < \theta, \\ \leq f'(s) & \text{when} \quad s > \theta. \end{cases}$$

The graph of a function which is star-shaped above at θ is easily visualized: It lies above a straight line through the point $(\theta, f(\theta))$, and a spectator located at this point has a line of sight to all other points of the graph. Two obvious facts should be noted: If a function $f : I \to \mathbb{R}$ is convex, it is star-shaped above at every $\theta \in I$, and if f is star-shaped above at θ, it has nondecreasing disparity weight about θ.

PROPOSITION 4.1 (Mosler and Muliere, 1993). *Let ϕ be additive (4.3) with g in $\mathcal{G}$. Then ϕ satisfies the star-shaped principle of transfers at θ for all n if and only if g is star-shaped above at θ.*

Castagnoli and Muliere (1990) introduce a *strengthened PD principle of transfers* with respect to a given threshold θ. The principle says that an index should follow the PD principle of transfers and, in addition, the strict principle of transfers about θ. Such an index is sensitive to a transfer from a rich household to a poor one, but possibly insensitive (though not decreasing) when income is transferred either between two rich households or between two poor ones. Castagnoli and Muliere (1990) show that

$$\max_{x_i < \theta} \frac{\partial}{\partial x_i} \phi(x) < \min_{x_i > \theta} \frac{\partial}{\partial x_i} \phi(x) \tag{4.5}$$

is sufficient for ϕ to satisfy the strict principle of transfers about ϕ. Therefore

$$\Psi_1 = \{\phi : \phi \text{ is Schur convex and (4.5) holds}\}$$

is a class of disparity indices satisfying their strengthened PD principle of transfers.

5. MULTIDIMENSIONAL ECONOMIC DISPARITY AND MAJORIZATION

Economic disparity does not arise from the distribution of income alone. Other attributes of affluence and well-being appear to be of similar interest in economic analysis. Households vary in income and assets, individuals in earnings and education, countries in per capita income and mineral resources, etc. In modern theories of social choice the specific distributional inequality of attributes like these is considered; see Fisher (1956), Tobin (1970), Sen (1970, 1973). If inequality in two or more attributes is treated simultaneously, we face the problem of modeling and measuring multidimensional economic disparity.

Consider a population of economic units $i \in \{1, \ldots n\}$ and a set of attributes $k \in \{1, \ldots d\}$. We will speak of households i and commodities k. Let $a_{ik} \geq 0$ be the endowment of household i with attribute k, $A = (a_{ik}) \in \mathbb{R}_+^{n \times d}$. By a_i we denote the ith row of A (the endowment vector of household i), by a^k the kth column of A (the distribution of attribute k in the population). In what follows we assume that $d \geq 1$. Hence, all results hold as well for the univariate case.

5.1. Multivariate Majorization

Given $A, B \in \mathbb{R}^{n \times d}$, we say that A *majorizes* B, $A \succ B$, if there exists a doubly stochastic matrix $T \in \mathcal{D}_{n,n}$ such that $B = TA$ holds. This corresponds to (2.3). The relation $\succ$ is a preorder on $\mathbb{R}_+^{n \times d}$, i.e., reflexive and transitive. $A \succ B$ implies that

$$\sum_{i=1}^{n} a_i = \sum_{i=1}^{n} b_i, \tag{5.1}$$

and that

$$b_i = \sum_{j=1}^{n} t_{ij} a_j \quad \text{where} \quad t_{ij} \geq 0 \ \forall i, j \quad \text{and} \quad \sum_{j=1}^{n} t_{ij} = 1 \ \forall i. \tag{5.2}$$

Thus, when $A \succ B$, the total of each commodity stays the same with A and B, and B is obtained from A by averaging the endowment vectors of households (with weights t_{ij}).

Multivariate majorization has been investigated by Rinott (1973), Marshall and Olkin (1974, 1979), Karlin and Rinott (1981, 1983), Arnold (1987), Bigard (1987), Bhandari (1988), Das Gupta and Bhandari (1989), Tong (1989), and Strasser (1992). For majorization on general state spaces of C^*- and W^*-algebras, see Alberti and Uhlmann (1982). In the economic literature, the seminal paper on majorization and the comparison of multidimensioned disparity is Kolm (1977).

A function $\phi : \mathbb{R}^{n\times d} \rightarrow \mathbb{R}$ is called *S-convex* iff $A \succ B$ implies $\phi(A) \geq \phi(B)$. ϕ is called *row symmetric* iff, for every permutation matrix $P \in \mathcal{P}_n$, $B = PA$ implies $\phi(A) = \phi(B)$. Every S-convex ϕ is row symmetric. ϕ is called *quasiconvex* iff, for every $A_1, \ldots A_m \in \mathbb{R}^{n\times d}$ and $(\lambda_1, \ldots \lambda_m) \in S_m$, $\phi(\sum_{l=1}^m \lambda_l A_l) \leq \max_l \phi(A_l)$. The notions are related by the following lemma, which is well known when $d = 1$; see Marshall and Olkin (1979, p. 69).

LEMMA 1. *If ϕ is row symmetric and quasiconvex, then ϕ is S-convex.*

Proof. Let $A \succ B$. By Birkhoff's theorem there exist $(\lambda_1, \ldots \lambda_m) \in S_m$ and $P_1, \ldots P_m \in \mathcal{P}_n$ such that $B = \sum_{l=1}^m \lambda_l P_l A$. As ϕ is quasiconvex, $\phi(B) \leq \max_l \phi(P_l A)$, and as ϕ is row symmetric, $\phi(P_l A) = \phi(A)$ for all l; hence $\phi(B) \leq \phi(A)$. Therefore, ϕ is S-convex. ∎

PROPOSITION 5.1. *Let $A, B \in \mathbb{R}^{n\times d}, d \geq 1$. Then $A \succ B$ is equivalent to each of the following conditions:*

(i) $B \in \text{conv}\{PA : P \in \mathcal{P}_n\}$,

(ii) $\phi(A) \geq \phi(B)$ *for all ϕ which are S-convex,*

(iii) $\phi(A) \geq \phi(B)$ *for all ϕ which are row symmetric and quasiconvex,*

(iv) $\phi(A) \geq \phi(B)$ *for all ϕ which are row symmetric and convex,*

(v) $\sum_{i=1}^n g(a_i) \geq \sum_{i=1}^n g(b_i)$ *for all $g : \mathbb{R}_d \rightarrow \mathbb{R}$ which are convex,*

(vi) (5.1) *and* $\sum_{i=1}^n g(a_i) \geq \sum_{i=1}^n g(b_i)$ *for all $g : \mathbb{R}_d \rightarrow \mathbb{R}$ which are increasing and convex.*

Proof. $A \succ B \Leftrightarrow$ (i) is an immediate consequence of Birkhoff's theorem. $A \succ B \Rightarrow$ (ii) holds by definition. (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (vi) is derived from inclusions of the respective sets of functions ϕ. Finally, we have to show (vi) $\Rightarrow A \succ B$: Let P_A be a probability measure in $\mathbb{R}_d$ giving mass n^{-1} to each a_i. Then $\int g(x)\, dP_A(x) = n^{-1} \sum g(a_i)$. Similarly, P_B is considered. A well-known result on dilations (e.g. Mosler and Scarsini, 1991) says that P_A is a dilation of P_B if and only if $\int x\, dP_A(x) = \int x\, dP_B(x)$ and $\int g(x)\, dP_A(x) \geq \int g(x)\, dP_B(x)$ for all g which are increasing and convex, i.e., if (vi) holds. [Equivalently, P_A is a dilation of P_B if and only if $\int g(x)\, dP_A(x) \geq \int g(x)\, dP_B(x)$ for all g which are convex, i.e., if (v) holds.] As $A \succ B$ is equivalent to saying that P_A is a dilation of P_B, there follows $A \succ B \Leftrightarrow$ (v) $\Leftrightarrow$ (vi). ∎

The economic interpretation of Proposition 5.1 is as follows. Given a permutation matrix P, PA is the matrix where all households have interchanged their endowments according to P. In terms of majorization, PA bears the same amount of inequality as A. Proposition 5.1(i) says that B is a convex combination of such permuted endowments. There are five classes of disparity indices ϕ, each of which induces the preorder $\succ$. They are interpreted and justified as in the univariate case. By the axiom of anonymity, a permutation of households should not affect inequality; hence ϕ should be row symmetric. The axiom of nonaltruism or a utilitarian axiom yields the additive decomposition of ϕ. The S-convex functions $\phi : \mathbb{R}^{n \times d} \rightarrow \mathbb{R}$ build the largest class of disparity indices which respect multivariate majorization. Further, quasiconvexity of ϕ implies that the conditional evaluation of a household's endowment (the other endowments held fixed) is quasiconvex, which is a standard assumption in the consumption theory of the household.

Special multivariate disparity indices have been proposed and applied to real data by Maasoumi (1986), Slottje (1987), Maasoumi and Nickelsburg (1988), Slesnick (1989), and others.

5.2. *Weakening Multivariate Majorization*

Majorization appears to be a rather strong notion of multivariate disparity. The reason is that for every attribute k averaging is done with the same weights t_{ij}. Two notions weaker than $\succ$ are of special interest.

PROPOSITION 5.2. *Let*
(i) $Ap \succ Bp$ *for all* $p \in \mathbb{R}^d$,
(ii) $a^k \succ b^k$ *for all* k.
Then $A \succ B \;\Rightarrow\;$ (i) $\;\Rightarrow\;$ (ii).

The reverse implications do not hold, in general. The proof of Proposition 5.2 is obvious. (i) is named *directional majorization*, (ii) *marginal majorization*. The latter means that every attribute k is more dispersed with A than with B in terms of ordinary majorization, i.e., averaging is done using different weights for different attributes. When the k's are commodities and p is a vector of prices for them, $a_i p$ amounts to the expenditure of household i. Then, $Ap \succ Bp$ says that with A the expenditures of households are more dispersed than with B. For this reason, (i) is also called *price majorization*.

Equivalent characterizations of directional and of marginal majorization are easily found along the lines of the univariate results. There exists another equivalent to directional majorization, which we will present in the next section on multivariate Lorenz ordering. Bhandari (1988) provides geometric conditions under which directional majorization implies multivariate majorization.

Foster et al. (1990) propose a ranking of social inequality which is related to price majorization. A stochastic version of price majorization — also with subsets of prices — is discussed in Muliere and Scarsini (1989). Rietveld (1990) assumes that individual welfare is the sum of welfare components arising from different attributes of well-being. He concludes that individual welfare cannot be more unequal (in terms of the Lorenz curve) than any of its components.

5.3. Different Population Sizes

When population sizes differ, say $A \in \mathbb{R}^{n\times d}$ and $B \in \mathbb{R}^{m\times d}$, majorization is similarly defined: *A majorizes B*, $A \succ B$, iff there exists a doubly stochastic matrix $T \in \mathcal{D}_{m,n}$ such that $(1/m)B = (1/n)TA$. For $d = 1$ and nonnegative vectors, it can be shown that the definition is the same as that given in Section 3.2. Then Equation (5.2) holds and

$$\frac{1}{n}\sum_{i=1}^{n} a_i = \frac{1}{m}\sum_{j=1}^{m} b_j. \tag{5.3}$$

Obviously, with the generalized definition, Proposition 5.2 remains true. The following analog of Proposition 5.1(v) and (vi) is obtained.

PROPOSITION 5.3. *Let $A \in \mathbb{R}^{n\times d}$ and $B \in \mathbb{R}^{m\times d}$. Then $A \succ B$ is equivalent to each of the following conditions:*

(i) $\sum_{i=1}^{n} g(a_i) \geq \sum_{j=1}^{m} g(b_j)$ *for all* $g : \mathbb{R}_d \to \mathbb{R}$ *which are convex,*

(ii) (5.3), *and* $\sum_{i=1}^{n} g(a_i) \geq \sum_{j=1}^{m} g(b_j)$ *for all* $g : \mathbb{R}_d \to \mathbb{R}$ *which are increasing and convex.*

Proof. P_A is a dilation of P_B. The proposition then follows from the well known result on dilations as in the proof of Proposition 5.1. ■

In case $m \leq n$, related results are found in Fischer and Holbrook (1980) and Karlin and Rinott (1983).

Another notion which is weaker than multivariate majorization has been introduced and investigated recently by Strasser (1992). He compares concentration tables, i.e. column stochastic matrices having $d+1$ columns. We present Strasser's results in our setting, which is slightly different.

DEFINITION 5.1 (Strasser, 1992). Let $A \in \mathbb{R}_+^{n\times d}$, $B \in \mathbb{R}_+^{m\times d}$ with $(1/n)\sum_{i=1}^{n} a_{ik} = (1/m)\sum_{j=1}^{m} b_{jk} > 0$ for all $k = 1, 2, \ldots, d$. A is called *less concentrated* than B iff for every column stochastic $S \in \mathcal{C}_{d+1,n}$ there is some column stochastic $R \in \mathcal{C}_{d+1,m}$ such that

$$\begin{aligned}\frac{1}{m}\sum_{j=1}^{m} r_{kj}b_{jk} \quad &\geq \quad \frac{1}{n}\sum_{i=1}^{n} s_{ki}a_{ik} \quad \text{for } k = 1, \ldots, d\\ \frac{1}{m}\sum_{j=1}^{m} r_{d+1,j} \quad &\geq \quad \frac{1}{n}\sum_{i=1}^{n} s_{d+1,i}.\end{aligned}$$

The *concentration function* K_A of A is defined by

$$K_A(z) = \sum_{i=1}^{n} \left(\sum_{k=1}^{d+1} z_k \tilde{a}_{ik} - \max_{1 \le k \le d+1} z_k \tilde{a}_{ik} \right), \quad z \in S_{d+1},$$

where $\tilde{a}_{ik} = a_{ik} / \sum_i a_{ik}$ if $k = 1, \dots, d$, and $\tilde{a}_{i,d+1} = 1/n$.

PROPOSITION 5.4 (Strasser, 1992). *Let $A \in \mathbb{R}^{n \times d}$ and $B \in \mathbb{R}^{m \times d}$ with $(1/n) \sum_{i=1}^{n} a_{ik} = (1/m) \sum_{j=1}^{m} b_{jk} > 0$ for all $k = 1, 2, \dots, d$.*
Then $A \succ B \Rightarrow K_A(z) \ge K_B(z)$ for all $z \in S_{d+1} \Leftrightarrow$ A is less concentrated than B.

When $d = 1$, the two conditions are also sufficient for $A \succ B$. For proof of Proposition 5.4, see Strasser (1992). Strasser's result reduces the comparison of multivariate disparity (in the sense of being less concentrated) to the comparison of real-valued functions. In the univariate case, with $A = a \in \mathbb{R}^n$, K_a is closely connected to the Lorenz function L_a:

$$K_a(\lambda, 1 - \lambda) = 1 - \lambda - \lambda L_a^* \left(\frac{1 - \lambda}{\lambda} \right).$$

Here L_a^* denotes the conjugate function of L_a, $L_a^*(t) = \sup_r (rt - L_a(r))$.

The idea behind the notion of being less concentrated is the following. Let $m = n$ and $d = 1$, and assume that a is less concentrated than b. Restricted to 0–1 matrices S and $R \in C_{2,n}$, the definition says that for every subset M_S of households there is another subset M_R having as many or more members such that the total endowment of the M_R-households under b is not larger than the total endowment of the M_S-households under a. This means that under b a smaller number of households (those not in M_R) obtains a larger share. The definition, more generally, uses weighted partitions and compares properly weighted sums. Here, given $S = (s_{ki}) \in C_{2,n}$, a household i is considered to be in M_S with weight $s_{2,i}$ and to be not in M_S with the remaining weight $s_{1,i} = 1 - s_{2,i}$. However, when $d \ge 2$, the remaining weight is split between the attributes, and interpretation becomes difficult.

The relation between Strasser's notion and directional majorization has still to be explored.

6. LORENZ ORDER IN THE MULTIVARIATE CASE

Extending the Lorenz curve to several attributes is not obvious. A natural postulate is that the multiattribute notion of the Lorenz curve should be symmetric in the attributes. For $d = 2$ attributes, Taguchi (1972a, b) and Arnold (1983) have introduced Lorenz surfaces in three-space. While Taguchi's definition is neither symmetric in the attributes nor easy to handle, Arnold's is both. His definition can be written as follows.

DEFINITION 6.1 (Arnold, 1983). Let F be a probability distribution function on $\mathbb{R}_{2+}$ having finite second and positive first moments. The *Lorenz surface* of F is the graph of the function

$$L(F, s, t) = \frac{\int_0^\xi \int_0^\eta xy \, dF(x, y)}{\int_0^\infty \int_0^\infty xy \, dF(x, y)},$$

where

$$s = \int_0^\xi dF_1(x), \quad t = \int_0^\eta dF_2(y), \quad 0 \le s, t \le 1,$$

F_1 and F_2 being the marginals of F.

If F is a product distribution function, $F(x, y) = F_1(x)F_2(y)$, then $L(F, s, t)$ is just the product of the marginal Lorenz functions. Let F_c denote the one-point distribution at $c \in \mathbb{R}_{d+} \setminus \{0\}$. F_c is called the *egalitarian distribution at* c. It follows that an egalitarian distribution has Lorenz function $L(F_c, s, t) = st$ when $d = 2$. The *two-attribute Gini-Arnold index* $\mathrm{GA}(F)$ is defined as four times the volume between the Lorenz surface of F and the Lorenz surface of an egalitarian distribution. In case of a product distribution function, $1 - \mathrm{GA}(F) = [1 - G(F_1)][1 - G(F_2)]$ holds, where $G(F_i)$ is the ordinary univariate Gini index. Arnold's definitions can be used as well for $d > 2$. But even when $d = 2$, to our knowledge there are no other simple relations to majorization or economic interpretations of the above. Instead, we present another notion in $\mathbb{R}_{d+1}$, the Lorenz zonoid.

DEFINITION 6.2. Let F be a probability distribution function on $\mathbb{R}_{d+}$, and $\int_{\mathbb{R}_d} x_j \, dF(x) > 0$ for all j. Define $\tilde{x}_j = x_j / \int_{\mathbb{R}_d} x_j \, dF(x)$ for $j = 1, \dots, d$, and $T(x) = (\tilde{x}_1, \dots, \tilde{x}_d)$. Then

$$\mathrm{LZ}(F) = \left\{ z \in \mathbb{R}_{d+1} : z = \left(\int_{\mathbb{R}_d} g(x) dF(x), \int_{\mathbb{R}_d} g(x) \cdot T(x) dF(x) \right), \right.$$
$$\left. g : \mathbb{R}_d \to [0, 1] \text{ continuous} \right\}$$

is called the *Lorenz zonoid*. Then for the egalitarian distribution F_c we get $T(c) = (1, \dots, 1)$, and

$$\mathrm{LZ}(F_c) = \{z \in \mathbb{R}_{d+1} : z = \gamma(1, \dots, 1),\ \ 0 \leq \gamma \leq 1\},$$

which is the main diagonal of the unit cube in $\mathbb{R}_{d+1}$.

LEMMA 2. *Assume $d = 1$. Let $L(F)$ be the ordinary Lorenz curve, and $\bar{L}(F)$ the dual Lorenz curve given by $\bar{L}(F, t) = 1 - L(F, 1 - t)$. Then* $\mathrm{LZ}(F)$ *is the area between $L(F)$ and $\bar{L}(F)$.*

Proof. $L(F)$ is given by $L(F, t) = \int_0^\xi T(x)\, dF(x)$ with $t = \int_0^\xi dF(x)$. On the other hand, at $z_1 = t$ the lower border of the Lorenz zonoid is the infimum of $z_2 = \int_0^\infty g(x)T(x)\, dF(x)$ subject to $\int_0^\infty g(x)\, dF(x) = t$ and $g : \mathbb{R}_d \to [0, 1]$ continuous. Since $T : x \mapsto \tilde{x}$ is nonnegative and increasing, the infimum is reached in the limit when g approaches the indicator function of the interval $[0, \xi]$; hence inf $z_2 = L(F, t)$. Similarly, it is shown that sup $z_2 = 1 - L(F, 1 - t)$ at $z_1 = t$. ■

Thus the ordinary Gini index $G(F)$ equals the area of $\mathrm{LZ}(F)$. We define the d-variate *Gini zonoid index* $\mathrm{GZ}(F)$ as the $(d+1)$-dimensional volume of the Lorenz zonoid.

Now, let $A \in \mathcal{C}_{n,d}$, and F be a discrete distribution function on $\mathbb{R}_{d+}$ giving equal mass to the rows of A. Then

$$\mathrm{LZ}(A) \equiv \mathrm{LZ}(F) = \left\{ z \in \mathbb{R}_{d+1} : z = \left(\frac{1}{n} \sum_{i=1}^{n} g(i),\quad \sum_{i=1}^{n} g(i) \cdot a_i \right), \right. \\ \left. 0 \leq g(i) \leq 1 \quad \text{for all } \ i \right\},$$

and $\mathrm{GZ}(A)$ is as above. Koshevoy (1992) has introduced the definitions of Lorenz zonoid and Gini zonotope index for this case. He calls $\mathrm{LZ}(A)$ the *Lorenz zonotope* of A.

Let F and G be probability distribution functions on $\mathbb{R}_{d+}, d \geq 1$. The *multivariate Lorenz order* $\succ_{\mathrm{L}}$ between F and G is defined as follows:

$$F \succ_{\mathrm{L}} G \quad \text{iff} \quad \mathrm{LZ}(F) \supset \mathrm{LZ}(G).$$

Similarly, when A and B are in $\mathcal{C}_{n,d}$,

$$A \succ_{\mathrm{L}} B \quad \text{iff} \quad \mathrm{LZ}(A) \supset \mathrm{LZ}(B).$$

For every F and $c \in \mathbb{R}_d$, obviously, $\mathrm{LZ}(F_c) \subset \mathrm{LZ}(F)$ holds; hence $F \succ_{\mathrm{L}} F_c$. The egalitarian distribution at some c is dominated by every other distribution. Similarly, if we define $E = (e_{ik}), e_{ik} = 1/n$ for all i and k, we get $\mathrm{LZ}(E) = \{z \in \mathbb{R}_{d+1} : z = \gamma(1, \ldots 1),\ 0 \leq \gamma \leq 1\}$. For every $A \in \mathcal{C}_{n,d}$ we have $\mathrm{LZ}(E) \subset \mathrm{LZ}(A)$; hence $A \succ_{\mathrm{L}} E$.

For $d = 1$, the definition is equivalent to that of ordinary Lorenz order; see Lemma 2.

PROPOSITION 6.1 (Koshevoy, 1992). *Let* $A, B \in \mathcal{C}_{n,d}$. *Then* $A \succ_{\mathrm{L}} B$ *if and only if* $Ap \succ Bp$ *for all* $p \in \mathbb{R}^d$.

For proof, see Koshevoy (1992). By the proposition, multivariate Lorenz order of matrices in $\mathcal{C}_{n,d}$ is the same as price majorization and is therefore a necessary (but in general not sufficient) condition for multivariate majorization. Moreover, it follows that the Gini zonoid index is not only consistent with multivariate Lorenz order but also with multivariate majorization.

7. CONCLUSIONS

We have presented various disparity orderings for distributions of single attribute and multiattribute well-being, and we have given classes of disparity indices which are compatible with the orderings and induce them. The orderings have been based on majorization or variants thereof, and the index classes usually have been subsets of S-convex functions.

In principle, every set Φ of symmetric functions $\phi : \mathbb{R}^n \to \mathbb{R}$ induces a majorization ordering $\succ_\Phi$ on $\mathbb{R}^n$: $a \succ_\Phi b$ iff $\phi(a) \geq \phi(b)$ for all $\phi \in \Phi$. If Φ is a singleton the distributions become completely comparable. If Φ is the set of symmetric and S-convex functions, $\succ_\Phi$ is ordinary majorization.

Ordinary majorization and the related orderings introduced above are preorderings only. They are rather coarse orderings under which (in the univariate case) distributions are only comparable if two Lorenz curves do not intersect. For theoretical and practical reasons there is some need for disparity comparisons when Lorenz curves intersect, i.e., for orderings finer than majorization.

Majorization between distributions having equal means is equivalent to convex stochastic ordering. There exist many other stochastic orderings in the literature, (see Mosler and Scarsini 1991) some of which have a meaning in terms of economic disparity. Shorrocks and Foster (1987) discuss a subset of strictly S-convex functions which induces an ordering of third-degree stochastic dominance (being finer than majorization, which corresponds to second-degree stochastic dominance). For further stochastic orderings in the measurement of economic disparity and welfare, see Alzaid (1990), Le Breton (1991), and — with multiple attributes — Atkinson and Bourguignon (1982, 1989) and Mosler (1993).

Economic disparity is one aspect of economic welfare, and every disparity index can be seen as (the negative of) an index of welfare. Authors stressing this view are Sen (1973), Cowell (1977), and Chakravarty (1990a). Recent new approaches include Chakravarty (1990b) and Bossert (1990).

A final remark on multivariate disparity measurement. In a certain sense, the notion of majorization is clarified when we consider the multiattribute case. Strictly speaking, majorization has nothing to do with the existence of "rich" people and "poor" people, but rather with the existence of people who are "different." When looking at endowments in just one attribute, the tails of their distributions are the most striking feature. This effect disappears when we have more than one attribute. Also, it is known that PD transfers in the multivariate yield a coarser ordering than multivariate majorization. Thus, the core of majorization does not consist in PD transfers but rather in the averaging of endowments (by doubly stochastic matrices) or the mixing of permuted endowments.

Future research may concentrate on the following topics: disparity orderings finer than majorization (especially in the multivariate), multivariate Lorenz orders, transfer principles other than the Pigou-Dalton, and decomposition of disparity.

Above all, Marshall and Olkin's book is a treasury of further variants of majorization, some of which still have to be exploited for economic-disparity measurement.

REFERENCES

Alberti, P. M. and Uhlmann, A. 1982. *Stochasticity and Partial Order*, Reidel, Boston.

Alzaid, A. A. 1990. Lorenz ranking of income distributions, *Statist. Papers* 31:209–224.

Arnold, B. C. 1983. *Pareto Distributions*, International Co-op. Publishing House, Fairland, Md.

Arnold, B. C. 1987. *Majorization and the Lorenz Order: A Brief Introduction*. Springer-Verlag, Berlin.

Arnold, B. C. 1991. Preservation and attenuation of inequality as measured by the Lorenz order, in *Stochastic Orders and Decision under Risk* (K. Mosler and M. Scarsini, Eds.), Inst. of Mathematical Statistics, Hayward, Calif., pp. 25–37.

Atkinson, A. B. 1970. On the measurement of inequality, *J. Econom. Theory* 2:244–263.

Atkinson, A. B. and Bourguignon, F. 1982. The comparison of multidimensioned distributions of economic status, *Rev. Econom. Stud.* 49:183–201.

Atkinson, A. B. and Bourguignon, F. 1989. The design of direct taxation and family benefits, *J. Public Econom.* 41:3–29.

Bhandari, S. K. 1988. Multivariate majorization and directional majorization; positive results, *Sankhyā Ser. A* 50:199–204.

Bigard, A. 1987. Analyse de l'inegalité multicritère, *Math. Sci. Hum.* 97:47–55.

Blackorby, C. and Donaldson, D. 1984. Social criteria for evaluating population change, *J. Public Econom.* 25:13–33.

Bossert, W. 1990. Maximin welfare orderings with variable population size, *Soc. Choice Welf.* 7:39–45.

Brualdi, R. A. 1984. The doubly stochastic matrices of a vector majorization, *Linear Algebra Appl.* 61:141–154.

Castagnoli, E. and Muliere, P. 1990. A note on inequality measures and the Pigou-Dalton principle of transfers, in *Income and Wealth Distribution, Inequality and Poverty* C. Dagum and M. Zenga, Eds.), Springer-Verlag, Berlin, pp. 171–182.

Chakravarty, S. R. 1990a. *Ethical Social Index Numbers*, Springer-Verlag, Berlin.

Chakravarty, S. R. 1990b. On quasi-orderings of income profiles, *Methods Oper. Res.* 60:455–473.

Cowell, F. A. 1977. *Measuring Inequality*, Philip Allan, Oxford.

Dalton, H. 1920. The measurement of the inequality of incomes, *Econom. J.* 30:348–361.

Das Gupta, P., Sen, A., and Starrett, D. 1973. Notes on the measurement of inequality, *J. of Econom. Theory* 6:180–187.

Das Gupta, S. and Bhandari, S. K. 1989. Multivariate majorization, in *Contributions to Probability and Statistics* (L. J. Gleser et al., Eds.), Springer-Verlag, New York, pp. 63–74.

Eichhorn, W., Funke, H., and Richter, W. F. 1984. Tax progression and inequality of income distribution, *J. Math. Econom.* 13:127–131.

Fellman, J. 1976. The effect of transformation on Lorenz curves, *Econometrica* 44:823–824.

Fields, G.S. and Fei, J. C. H. 1978. On inequality comparisons, *Econometrica* 46:303–316.

Fischer, P. and Holbrook, J. A. R. 1980. Balayage defined by the nonnegative convex functions, *Proc. Amer. Math. Soc.* 79:445–448.

Fisher, F. M. 1956. Income distribution, value judgements and welfare, *Quart. J. Econom.* 70:380–424.

Foster, J. E., Majumdar, M. K., and Mitra, T. 1990. Inequality and welfare in market economics, *J. Public Econom.* 41:351–367.

Gini, C. 1912. *Variabilità e Mutabilità: Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche* (Variability and Changeability: Contribution to the Study of Distributions and Statistical Relations), Cuppini, Bologna.

Hardy, G. H., Littlewood, J. E., and Polya, G. 1929. Some simple inequalities satisfied by convex functions, *Messenger Math.* 58:145–152.

Hardy, G. H., Littlewood, J. E., and Polya, G. 1934. *Inequalities*, Cambridge U. P., London.

Jakobson, U. 1976. On the measurement of degree of progression, *J. Public Econom.* 5:161–168.

Kakwani, N. C. 1977. Applications of Lorenz curve in economic analysis, *Econometrica* 45:719–727.

Karlin, S. and Rinott, Y. 1981. Entropy inequalities for classes of probability distributions II. The multivariate case, *Adv. Appl. Probab.* 13:325–351.

Karlin, S. and Rinott, Y. 1983. Comparison of measures, multivariate majorization, and application to statistics, in *Studies in Econometrics, Time Series, and Multivariate Statistics* (S. Karlin, T. Amemiya, and L. A. Goodman, Eds.), Academic, New York, pp. 465–489.

Kolm, S. C. 1969. The optimal production of social justice, in *Public Economics* (J. Marjolis and H. Guitton, Eds.), Macmillan, New York, pp. 145–200.

Kolm, S. C. 1977. Multidimensional egalitarianisms, *Quart. J. Econom.* 91:1–13.

Koshevoy, G. 1992. An Equivalence Theorem and Multidimensional Inequality, Mimeo, Russian Academy of Science, Moscow.

Le Breton, M. 1991. Stochastic orders in welfare economics, in *Stochastic Orders and Decision under Risk* (K. Mosler and M. Scarsini, Eds.), Inst. of Mathematical Statistics, Hayward, Calif., pp. 190–206.

Lorenz, M. O. 1905. Methods of measuring the concentration of wealth, *Publ. Amer. Statist. Assoc.* 9:209–219.

Maasoumi, E. 1986. The measurement and decomposition of multi-dimensional inequality, *Econometrica* 54:991–997.

Maasoumi, E. and Nickelsburg, G. 1988. Multivariate measures of well-being and an analysis of inequality in the Michigan data, *J. Business and Econom. Statist.* 6:327–334.

Marshall, A. W. and Olkin, I. 1974. Majorization in multivariate distributions, *Ann. Statist.* 2:1189–1200.

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

Mosler, K. 1993. Multidimensional welfarisms, in *Models and Measurement of Inequality and Welfare* (W. Eichhorn, Ed.), Springer-Verlag, Berlin.

Mosler, K. and Muliere, P. 1993. Robin Hood operations, absolutely speaking, in *Statistics and Quantitative Economoics* No. 60, Hamburg.

Mosler, K. and Scarsini, M. 1991. Some theory of stochastic dominance, in *Stochastic Orders and Decision under Risk* (K. Mosler and M. Scarsini, Eds.), Inst. of Mathematical Statistics, Hayward, Calif., pp. 261–284.

Moyes, P. 1987. A new concept of Lorenz domination, *Econom. Lett.* 23:203–207.

Moyes, P. 1989. Some classes of functions that preserve the inequality and welfare orderings of income distributions, *J. Econom. Theory* 49:347–359.

Muliere, P. and Scarsini, M. 1989. Multivariate decisions with unknown price vector, *Econom. Lett.* 29:13–19.

Nygård, F. and Sandström, A. 1981. *Measuring Income Inequality*, Almqvist and Wiksell International, Stockholm.

Pfingsten, A. 1986. Distributionally neutral tax changes for different inequality concepts, *J. Public Econom.* 30:385–393.

Piesch, W. 1975. *Statistische Konzentrationsmaße*, J. C. B. Mohr, Tübingen.

Pigou, A. C. 1912. *Wealth and Welfare*, Macmillan, New York.

Rietveld, P. 1990. Multidimensional inequality comparisons, *Econom. Lett.* 32:187–192.

Rinott, Y. 1973. Multivariate majorization and rearrangement inequalities with some applications to probability and statistics, *Israel J. Math.* 15:60–77.

Rothschild, M. and Stiglite, J. E. 1973. Some further results on the measurement of inequality, *J. Econom. Theory* 6:188–204.

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

Sen, A. K. 1973. *On Economic Inequality*, Oxford U.P., Oxford.

Shorrocks, A. F. 1983. Ranking income distributions, *Economica* 50:3–17.

Shorrocks, A. F. and Foster, J. E. 1987. Transfer sensitive inequality measures, *Rev. Econom. Stud.* 54:485–497.

Slesnick, D. T. 1989. Specific egalitarianism and total welfare inequality: A decompositional analysis, *Rev. Econom. and Statist.* 71:116–127.

Slottje, D. J. 1987. Relative price changes and inequality in the size distribution of various components of income: A multidimensional approach, *J. Business and Econom. Statist.* 5:19–26.

Strasser, H. 1992. Concentration of multivariate statistical tables, *Statist. Papers* 33:95–117.

Taguchi, T. 1972a. On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case—I, *Ann. Inst. Statist. Math.* 24:355–382.

Taguchi, T. 1972b. On the two-dimensional concentration surface and extensions of concentration coefficient and Pareto distribution to the two dimensional case—II, *Ann. Inst. Statist. Math.* 24:599–619.

Tobin, J. 1970. On limiting the domain of inequality, *J. Law and Econom.* 13:263–277.

Tong, Y. L. 1989. Probability inequalities for n-dimensional rectangles via multivariate majorization, in *Contributions to Probability and Statistics* (L. J. Gleser et al., Eds.), Springer-Verlag, New York, pp. 146–159.