Working Paper No. 266

Extension of Generalized Lorenz Dominance Criterion to Multivariate Attributes

Kazuyuki Nakamura

February, 2012

UNIVERSITY OF TOYAMA 富山大学

FACULTY OF ECONOMICS
UNIVERSITY OF TOYAMA

Extension of Generalized Lorenz Dominance Criterion to Multivariate Attributes*

Kazuyuki Nakamura†
Faculty of Economics, University of Toyama
Gofuku 3190, Toyama 930-8555, JAPAN

February, 2012

Abstract

This paper considers an extension of generalized Lorenz dominance (GL) criterion to the case of multivariate attributes. Based on the uniform majorization of Kolm (1977), we propose an extended version of GL which we call uniform supermajorization which allows for attributes with different means. It is verified that uniform supermajorization has similar implications as uniform majorization for welfare ordering. Furthermore, we show that this criterion can be used for comparison of distributions of unequal populations. We also provide a procedure for empirical investigations which is in the standard form of linear programming problems.

key words: Generalized Lorenz dominance, Multidimensional inequality, Uniform majorization.
JEL code: D31, D63

1 Introduction

Generalized Lorenz dominance (GL) proposed by Shorrocks (1983) and Kakwani (1984) is one of the most frequently used criteria in the empirical literature of income inequality.[1] There are some important reasons why GL is appropriate for investigating income inequality and social welfare ordering. First, as with the ordinary Lorenz dominance criterion, GL brings a clear welfare implication referred to as Shorrocks Theorem.[2] Second, under the assumption of replication invariance on social evaluation function (SEF), a welfare

*This study is supported by CEAKS University of Toyama, and JSPS KAKENHI (21530309). An earlier version of this paper was presented at the workshop held at Renmin University of China. I am grateful to Guanjian Xu, Jun Cui, and Jianguo Jiao.

†Faculty of Economics, University of Toyama. Email: knakamur@eco.u-toyama.ac.jp

[1] Kakwani (1984) provided an international comparison of welfare rankings based on GL. Bishop et al. (1993) implemented GL comparison for ten western countries based on Luxembourg Income Study data. Chiou (1996) analyzed the Taiwanese distribution of income by using GL. Mukhopadhaya (2003) analyzed the change in the social welfare in Singapore using GL and other dominance criteria.

[2] See Shorrocks (1983) and Lambert (1993, ch.3).

comparison between societies with different population sizes can be performed easily.[3] Finally, drawing the generalized Lorenz curve, we can easily implement welfare comparison based on the GL criterion.

Although income and wealth are primary factors characterizing well-being of person, other variables such as health, education and environmental quality also play key roles as determinants of welfare. For example, the Human Development Index (HDI) of the United Nations is intended to evaluate well-being of society with respect to living standard, health, and education. Investigations on disparities based on multivariate attributes have been considered in many fields of economics, including development economics, public finance and regional economics.

This paper extends the GL criterion to the case of multivariate attributes. Inferences on multidimensional inequality are classified into two approaches: aggregative and non-aggregative. In the aggregative approach, a variety of inequality indices that reflect the multivariate attributes have been developed and their properties have been investigated.[4] Furthermore, by using developed indices, much literature has considered empirically multidimensional inequality in various countries and periods.[5] Within the context of the non-aggregative approach, Lorenz dominance criterion by Atkinson (1970) has been extended to various types of multidimensional criteria. For example, Atkinson and Bourguignon (1987) proposed a procedure referred to as sequential generalized Lorenz dominance (SGL) which compares the distribution of income reflecting differences in needs among households.[6]

Uniform majorization (UM) proposed by Kolm (1977) also can be classified into the non-aggregative approach as an extension of the ordinary Lorenz dominance criterion.[7] As pointed out by Kolm, UM has a clear welfare implication which is the same as that of univariate Lorenz dominance. A pair-wise comparison of distributions employing UM can be implemented in the case of equal means for each pair of compared attributes. However, this is not always the case in practice. In empirical analysis, we usually have to consider welfare ordering by comparing distributions having different mean and population sizes.

[3]It should be noted that replication invariance is a controversial property in evaluating social welfare. See Aboudi et al. (2010) and Blackorby et al. (2009).

[4]For example, see Weymark (2006), and Tsui (1999).

[5]Lugo (2007) compared the multidimensional inequality indices by applying them to Argentina data, and investigated the features and limitations of multidimensional inequality indices including Maasoumi (1986), Tsui (1995) and Bourguignon (1999). Brandolini (2008) investigated income and health inequalities in EU countries by using various inequality indices including those of Maasoumi (1986), Tsui (1995), and Bourguignon and Chakravaty (2003).

[6]Employing multi-dimensional dominance approaches based on SGL, Nilsson (2010) considered the inequality among Zambian households. Muller and Trannoy (2011) proposed a multidimensional dominance criterion having attributes consisting of income, health and education and investigated the change in the welfare distribution among countries between 2000 and 2004.

[7]Various concepts of multivariate majorization are discussed in Marshall and Olkin (2009, ch.15). In the literature of the economic inequality, by introducing the concept of Lorenz zonotope, Koshevoy (1995) investigated a multivariate generalization of the Lorenz curve. See also Koshevoy and Mosler (2007). Tsui (1999) proposed correlation increasing majorization which focuses on the transfer of resources to increase correlation among individual resource endowment. Kolm (1977) and List (1999) considered nonnegative price majorization, in which the attributes of each person are summed up using nonnegative weight.

As a result, UM itself is not employed to compare such distributions, although it has been taken as an axiom to be satisfied by multidimensional inequality indices.[8]

Herein, we focus on an extension of UM to the case where both population size and attribute mean differ for the pair of distributions. Consequently, such an extension is also a generalization of GL. It is desirable that an extended version of GL maintains the convenient properties noted above. The extension we propose has the same welfare implications as GL. In addition, this is achieved without any additional specifications on the utility and social welfare.

In order to link the theoretical results with practical investigation, we provide an empirical procedure based on linear programming. When considering well-being consisting of multivariate attributes, it is not possible to provide a graphical representation like the generalized Lorenz curve. Instead of such geometric investigation, we will show a procedure to obtain the welfare ranking by using a linear-programming model. The procedure presented in this paper has some advantages. First, when one distribution does not dominate the other distribution, we can show quantitatively each individual's shortage of attributes to achieve welfare superiority by solving the linear programming problem. This may be useful information for policy makers. Second, by solving the linear programming problem, we can know a preference on attributes that does not lead to welfare superiority if there is no relationship between distributions of attributes.

The remainder of the present paper is organized as follows. In Section 2, we provide an analytical framework. In Section 3, we consider the extension of GL to the case of multivariate attributes. In Section 4, we deal with comparisons among distribution with different population sizes. In Section 5, we present an empirical procedure. We illustrate the procedure by using Chinese provincial data in Section 6. In the last section, we conclude the analysis.

2 Analytical Framework

2.1 Individuals, Attributes and their Distributions

Let X and Y be two empirical distributions of m attributes among n_X and n_Y persons, respectively. Thus, two distributions of attributes are represented by m-by-n_J matrices as follows:

$$X \equiv [\mathbf{x}_1, ..., \mathbf{x}_{n_X}] \in \mathbb{R}^{m,n_X}, \tag{1}$$

$$Y \equiv [\mathbf{y}_1, ..., \mathbf{y}_{n_Y}] \in \mathbb{R}^{m,n_Y}, \tag{2}$$

where m-dimensional column vectors $\mathbf{x}_i$ and $\mathbf{y}_i$ represent the attributes of person i as follows:

[8] For example, see Tsui (1995, 1999), and Gajdos and Weymark (2005).

$$\mathbf{x}_i = \begin{bmatrix} x_{1i} \\ \vdots \\ x_{mi} \end{bmatrix} \in \mathbb{R}^{n_X}, \quad \mathbf{y}_i = \begin{bmatrix} y_{1i} \\ \vdots \\ y_{mi} \end{bmatrix} \in \mathbb{R}^{n_Y}.$$

Hereinafter, a society whose distribution of attributes is characterized by $J = X, Y$ is referred to as society J. In a later section when we focus on a distribution of one of m attributes, row vectors $\mathbf{x}^j$ and $\mathbf{y}^j$ representing the marginal distribution of attribute $j = 1, \ldots, m$ will be used. That is,

$$\mathbf{x}^j = [x_{j1}, ..., x_{jn_X}],$$
$$\mathbf{y}^j = [y_{j1}, ..., y_{jn_Y}].$$

Example 1 (Empirical distribution of attributes). Consider a society consisting of three persons. Suppose that income and health status affect their well-being. In this case, the distribution of attributes can be written by the 2-by-3 matrix X as follows:

$$X = \begin{array}{r c} & \begin{array}{ccc} \text{person1} & \text{person2} & \text{person3} \end{array} \\ \begin{array}{r} \text{income}\{ \\ \text{health}\{ \end{array} & \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \end{bmatrix} \end{array}.$$

2.2 Utility Index

Suppose that utility of each person depends on her/his attributes and is represented by a utility index. The utility index is not necessarily a utility function. We can imagine a situation in which a policy maker or a researcher uses the utility index to evaluate individuals' utility. The utility index of $i = 1, ..., n_X$ can be written as follows:

$$u_i = u(\mathbf{x}_i). \tag{3}$$

The utility index of individuals in society Y can be represented analogously.

We make the following assumption on the utility index.

Assumption 1. The utility index has the following properties:

(i) u is nondecreasing in each attribute. That is, for $\Delta\mathbf{x}_i \geq \mathbf{0}$,

$$u(\mathbf{x}_i + \Delta\mathbf{x}_i) \geq u(\mathbf{x}_i).$$

(ii) u is concave in $\mathbf{x}_i$. That is, for $\lambda \in [0, 1]$, $\mathbf{x}_i'$ and $\mathbf{x}_i''$, the following inequality holds:

$$u[\lambda\mathbf{x}_i' + (1-\lambda)\mathbf{x}_i''] \geq \lambda u(\mathbf{x}_i') + (1-\lambda)u(\mathbf{x}_i'').$$

(iii) u is continuous.

Hereafter, we denote the set of the utility indices satisfying Assumption 1 as Ω_u.

2.3 Social Evaluation Function (SEF)

For $u \in \Omega_u$, consider a SEF represented as follows: for $J = X, Y$,

$$SW_J = SW(\mathbf{u}_J), \tag{4}$$

where $\mathbf{u}_J \equiv [u_1, ..., u_{n_J}]$ denotes a vector of the utility indices. We make the following assumptions on the SEF.

Assumption 2. The SEF is an increasing and Schur concave in $\mathbf{u}_J$.

We denote the set of the SEFs satisfying Assumption 2 as Ω_W. If we concentrate our attention on the differentiable function, Schur concave SWF is characterized by the conditions that SW is symmetric with respect to $\mathbf{u}$ and $(u_i - u_j)(\partial SW/\partial u_i - \partial SW/\partial u_j) \leq 0$ holds for all i, $j = 1, ..., n_J$.[9]

Example 2 (Schur concave function). The following social welfare functions belong to Ω_W.

$$SW_B = \sum_{i=1}^{n_J} (u_i)^{\beta}, \quad \text{for } \beta \in (0, 1],$$

$$SW_S = \frac{\sum_{i=1}^{n_J} u_i}{n_J} (1 - GINI_u).$$

where $GINI_u$ denotes the Gini coefficient of the distribution of utility index. □

In the above SWFs, SW_S, can be interpreted as a utility version of Sen's social welfare function.

3 Generalization of Shorrocks Theorem

In this section, we consider a situation in which two societies have equal population. For notational simplicity, the number of person in each society is denoted by $n \equiv n_X = n_Y$.

[9] For detail discussion, see Marshall et al. (2011, ch. 3) and Bhatia (1997, II.3).

3.1 Generalized Lorenz Dominance

Before considering multivariate dominance criterion, we confirm properties of the univariate GL. If we concentrate on the case of equal population, the GL criterion can be described as follows:

Definition 1 (Generalized Lorenz Dominance). Let $\mathbf{x}$ and $\mathbf{y}$ be n-dimensional real row vectors with nonnegative entries. The vector $\mathbf{x}$ dominates $\mathbf{y}$ in the sense of generalized Lorenz dominance if and only if $\sum_{i=1}^{k} x_i^{\uparrow} \geq \sum_{i=1}^{k} y_i^{\uparrow}$ holds for $k = 1, ..., n$, where $x_i^{\uparrow}$ and $y_i^{\uparrow}$ are entries of $\mathbf{x}$ and $\mathbf{y}$ in increasing order, respectively.

In the literature of the theory of majorization, GL is said to be a weak supermajorization.[10] For two distributions with equal population, if $\mathbf{x}$ generalized Lorenz dominates $\mathbf{y}$, we can say that $\mathbf{x}$ is weakly supermajorized by $\mathbf{y}$ (in symbols, $\mathbf{x} \prec^{w} \mathbf{y}$). Weak supermajorization can be characterized by using doubly stochastic matrix, which is a nonnegative square matrix whose row and column sums are all one. We denote the set of all doubly stochastic matrices with order n as Ψ_n.

Lemma 1. Let $\mathbf{x}$ and $\mathbf{y}$ be n-dimensional row vectors. The following two conditions are equivalent.

(i) $\mathbf{x} \prec^{w} \mathbf{y}$.

(ii) There exist some $P \in \Psi_n$ and nonnegative vector $\mathbf{s} \in \mathbb{R}_{+}^{n}$ such that

$$\mathbf{y}P + \mathbf{s} = \mathbf{x}. \tag{5}$$

Proof. Suppose that (i) holds. Then, we can find an nonnegative scalar α such as $\sum_{i=1}^{n} x_i^{\uparrow} = \sum_{i=1}^{n} y_i^{\uparrow} + \alpha$. Now, we define $\hat{y}_n^{\uparrow} \equiv y_n^{\uparrow} + \alpha$. Since $[x_1^{\uparrow}, ..., x_n^{\uparrow}]$ is majorized by $[y_1^{\uparrow}, ..., y_{n-1}^{\uparrow}, \hat{y}_n^{\uparrow}]$, there exists $P \in \Psi_n$ such that $\mathbf{y}P + \hat{\mathbf{s}} = \mathbf{x}$ holds, where $\hat{\mathbf{s}} = [0, ..., 0, \alpha]P \geq \mathbf{0}$. This implies (ii). Conversely, suppose that (ii) holds. We have $\mathbf{y}P = \mathbf{x} - \mathbf{s}$ for some $P \in \Psi_n$ which implies $\sum_{i=1}^{k} (x_i - s_i)^{\uparrow} \geq \sum_{i=1}^{k} y_i^{\uparrow}$ for $k = 1, ..., n-1$ and $\sum_{i=1}^{n} (x_i - s_i)^{\uparrow} = \sum_{i=1}^{n} y_i^{\uparrow}$. Clearly, $\sum_{i=1}^{k} x_i^{\uparrow} \geq \sum_{i=1}^{k} (x_i - s_i)^{\uparrow}$ holds for $k = 1, ..., n$, which implies (i). ■

It should be noted that in the univariate case, the characterization of majorization by doubly stochastic matrix coincides with that by T-transform. However, these two characterizations lead to different consequences when we consider the multidimensional majorization.[11] In the next subsection, we consider an extension of GL criterion in line with (5).

[10] For example, see Saposnik (1993).

[11] See Weymark (2006). In Marshall et al. (2011, ch. 15), the uniform majorization is referred to as simply majorization. On the other hand, they use the term of chain majorization if a distribution is obtained from products of T-transformation of the other distribution.

3.2 Multivariate Majorization

The following definition of multivariate majorization can be regarded as a straightforward extension of (5).

Definition 2 (Uniform Supermajorization). Let X and Y be m-by-n matrices. X is said to be uniformly supermajorized by Y (in symbols, $X \prec^w Y$) if

$$YP \leq X, \tag{6}$$

holds for a doubly stochastic matrix P.

If (6) holds with equality X is said to be uniformly majorized by Y (in symbols, $X \prec Y$). It is clear that $X \prec Y$ implies $X \prec^w Y$. Thus, UM can be regarded as a special case of uniform supermajorization. It is also clear that GL is a special case of uniform supermajorization.

Example 3 (Uniform supermajorization). Consider a society consisting of three persons and two attributes. Let X and Y be two matrices defined as follows:

$$X = \begin{bmatrix} 45 & 65 & 35 \\ 30 & 50 & 45 \end{bmatrix}, \quad Y = \begin{bmatrix} 10 & 50 & 80 \\ 40 & 20 & 60 \end{bmatrix}.$$

We can confirm $X \prec^w Y$, since there exists a doubly matrix,

$$P = \begin{bmatrix} 0.3 & 0.1 & 0.6 \\ 0.6 & 0.3 & 0.1 \\ 0.1 & 0.6 & 0.3 \end{bmatrix},$$

such that $YP \leq X$ holds. Indeed,

$$YP = \begin{bmatrix} 41 & 64 & 35 \\ 30 & 46 & 44 \end{bmatrix} \leq \begin{bmatrix} 45 & 65 & 35 \\ 30 & 50 & 45 \end{bmatrix} = X.$$

holds.

We can confirm some basic properties of uniform supermajorization as follows:

Remark 1. Let X, Y, Z be m-by-n matrices.

(i) $X \prec^w X$.

(ii) If $X \prec^w Y$ and $Y \prec^w Z$ hold then $X \prec^w Z$ holds.

(iii) If $X \prec^w Y$ and $Z \prec^w Y$ hold then $\lambda X + (1-\lambda) Z \prec^w Y$ holds for $\lambda \in [0,1]$.

(iv) If $X \prec^w Y$ holds then $X[J] \prec^w Y[J]$ holds for each $J \subset \{1, ..., m\}$, where $X[J]$ is a submatrix of X whose rows are the rows of X indexed by the elements in J.

In Remark 1, (i) and (ii) represent reflexivity and transitivity, respectively. (iii) implies pre-order induced by $\prec^w$ is a convex order. (iv) implies that a necessary condition for $X \prec^w Y$ is that $\mathbf{x}^j$ generalized Lorenz dominates $\mathbf{y}^j$ for $j = 1, ..., m$.

3.3 Relation to Other Majorization Criteria

In the theory of stochastic dominance, uniform supermajorization is known as increasing-concave (second-order) stochastic ordering (ICV).[12] In this sense, the concept of uniform supermajorization is already-known rather than novel. However, by considering ICV based on empirical distributions, we can relate ICV to other majorization criteria.

First, it should be noted that the doubly stochastic matrix can be regarded as a special case of row-stochastic matrix. A s-by-k dimensional nonnegative matrix whose row sum are all one is called row stochastic. Let $\Psi^R_{s,k}$ be the set of all s-by-k row stochastic matrices. That is, for $R \in \Psi^R_{s,k}$, $R\mathbf{e}_k^T = \mathbf{e}_k^T$ and $R > O$ hold, where $\mathbf{e}_k$ denotes a k-dimensional row vector whose entries are one.[13] The set of all row stochastic square matrices with n-dimension is denoted as Ψ^R_n.

Dahl (1999) proposed (right) matrix majorization. For X and $Y \in \mathbb{R}^{m,n}$, X is (right) matrix majorized by Y if there exists $R \in \Psi^R_n$ such that $YR = X$ holds. In Dahl's matrix majorization, R is not necessarily square matrix: variable size in population is allowed. On the other hand, since $YR = X$ implies $Y\mathbf{e}_n^T = X\mathbf{e}_n^T$, total amount of each attribute have to be the same between X and Y.

Kolm (1977) and List (1999) consider non-negative price majorization. For X and $Y \in \mathbb{R}^{m,n}$, X is said to be non-negative price majorized by Y if $\mathbf{v}X \prec \mathbf{v}Y$ for all $\mathbf{v} \in \mathbb{R}^m_+$. That is, $\mathbf{v}YP = \mathbf{v}X$ holds for some $P \in \Psi_n$. Price majorization is referred to as non-negative linear-combinations majorization in Marshall et al. (2011, ch. 15). Since $\mathbf{v}YP\mathbf{e}_n^T = \mathbf{v}Y\mathbf{e}_n^T = \mathbf{v}X\mathbf{e}_n^T$, non-negative price majorization considers a situation in which both population size and attribute mean are the same between the two distributions.

Savaglio (2011) considered rs-majorization. That is, for X and $Y \in \mathbb{R}^{m,n}$, X is said to be rs-majorized by Y if there exists a row stochastic matrix R such that $RY^T = X^T$ holds. In Pería et al. (2005), rs-majorization is referred to as weak matrix majorization. Marshall et al. (2011, ch. 15) use the term of "column-stochastic majorization" in place of rs-majorization, since $RY^T = X^T$ can be rewritten as $YR^T = X$. In rs-majorization, total amount of each attribute may differ while the size in population is fixed.

Two types of matrix majorizations proposed by Dahl (1999) and Savaglio (2011) are closely related to UM as understood from that the relation between two matrices are related by equality. Clearly, $YP = X$ for some $P \in \Psi_n$ implies $YR = X$ and $YR^T = X$ for some $R \in \Psi^R_n$.

We can consider the inequality versions of the majorizations described above as follows: Let X and Y be m-by-n matrices.

[12] For example, see Shaked and Shanthikumar (1994, ch. 5), and Marshall et al. (2011. ch.17).

[13] In what follows, superscript T denotes the transpose operation.

(RM) There exists some row stochastic matrix R such that $YR \leq X$ holds.

(PM) For all non-negative m-dimensional vector $\mathbf{v}$, $\mathbf{v}X \prec^w \mathbf{v}Y$ holds.

(CM) There exists some row stochastic matrix R such that $YR^T \leq X$ holds.

It is clear (**RM**), (**PM**) and (**CM**) correspond to the inequality versions of (right) matrix majorization, non-negative price majorization and rs-majorization, respectively. The relations between uniform supermajorization, (**RM**), and (**PM**) can be summarized as follows:

Remark 2. Let $X, Y \in \mathbb{R}^{m,n}$. If $X \prec^w Y$ then,

(i) $YR \leq X$ holds for some $R \in \Psi_n^R$;

(ii) $\mathbf{v}X \prec^w \mathbf{v}Y$ holds for all $\mathbf{v} \in \mathbb{R}_+^m$.

Proof. (i) follows from $\Psi_n \subset \Psi_n^R$. $YP \leq X$ implies $\mathbf{v}YP \leq \mathbf{v}X$ due to $\mathbf{v} \in \mathbb{R}_+^m$. ■

Furthermore, we obtain the following property relating (**PM**) to (**CM**), which is a straightforward application of Pería et al. (2003, Proposition 3.3).

Remark 3. Let $X, Y \in \mathbb{R}^{m,n}$. If $\mathbf{v}X \prec^w \mathbf{v}Y$ holds for all $\mathbf{v} \in \mathbb{R}_+^m$ then $YR^T \leq X$ holds for some $R \in \Psi_n^R$.

Proof. See Appendix. ■

Thus, from Remark 2, it is verified that $X \prec^w Y \Rightarrow$ (**RM**). In addition, together Remark 3 with Remark 2, we obtain $X \prec^w Y \Rightarrow$ (**PM**) $\Rightarrow$ (**CM**). However, the reciprocal implications are not always true as shown in the following examples.

Example 4 ((**RM**) does not imply $X \prec^w Y$). Consider the following two matrices.

$$X = \begin{bmatrix} 5 & 10 & 40 \\ 5 & 1 & 15 \end{bmatrix}, \quad Y = \begin{bmatrix} 10 & 10 & 30 \\ 5 & 5 & 5 \end{bmatrix}.$$

Considering a matrix

$$R_0 = \begin{bmatrix} 0.2 & 0 & 0.8 \\ 0.2 & 0 & 0.8 \\ 0 & 0.2 & 0.8 \end{bmatrix},$$

we can confirm

$$YR_0 = \begin{bmatrix} 4 & 6 & 40 \\ 2 & 1 & 12 \end{bmatrix} \leq \begin{bmatrix} 5 & 10 & 40 \\ 5 & 1 & 15 \end{bmatrix} = X.$$

On the other hand, it is easily verified that $X[1] \not\prec^w Y[1]$ and $X[2] \not\prec^w Y[2]$. This implies $X \not\prec^w Y$ from Remark 1(iv).

Example 5 ((**PM**) does not imply $X \prec^w Y$). Let X and Y be two matrices.

$$X = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 4 & 6 \end{bmatrix}, \quad Y = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 6 \end{bmatrix}.$$

We can easily confirm (**PM**) holds. On the other hand, consider a matrix $A = [a_{ij}]$ such that $YA \leq X$, $A\mathbf{e}_3^T = \mathbf{e}_3^T$ and $\mathbf{e}_3 A = \mathbf{e}_3$ hold. Since $\mathbf{y}^1\mathbf{a}_1 \leq 1$ and $\mathbf{a}_1\mathbf{e}_3^T = 1$ must hold, $\mathbf{a}_1$ becomes $\mathbf{a}_1 = (1,0,0)^T$ as long as $\mathbf{a}_1 \geq \mathbf{0}$. This means $\mathbf{a}_2 = (0, \hat{a}, 1-\hat{a})^T$ and $\mathbf{a}_3 = (0, 1-\hat{a}, \hat{a})^T$. Furthermore, $\mathbf{y}^2\mathbf{a}_2 \leq 4$ must hold for $YA \leq X$. Noting that $a_{22} = 1 - a_{32}$ for $\mathbf{e}_3 A = \mathbf{e}_3$, we have to say $a_{32} \leq -1$. Thus, there does not exist doubly stochastic matrix such that $YA \leq X$.

Example 6 ((**CM**) does not imply (**PM**)). Let X and Y be two matrices.

$$X = \begin{bmatrix} 10 & 10 & 16 \\ 10 & 12 & 16 \end{bmatrix}, \quad Y = \begin{bmatrix} 10 & 10 & 20 \\ 10 & 12 & 20 \end{bmatrix}.$$

Consider a matrix $R_C \in \Psi_n^R$ such that

$$R_C^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0.5 \\ 0 & 0 & 0.5 \end{bmatrix}.$$

By using R_C^T, we obtain

$$YR_C^T = \begin{bmatrix} 10 & 10 & 15 \\ 10 & 12 & 16 \end{bmatrix} \leq \begin{bmatrix} 10 & 10 & 16 \\ 10 & 12 & 16 \end{bmatrix} = X.$$

However, taking $\mathbf{v}$ as $\mathbf{v}_0 = (1,1)$, we have $\mathbf{v}_0 X\mathbf{e}^T \leq \mathbf{v}_0 Y\mathbf{e}^T$ which implies $\mathbf{v}_0 X \not\prec^w \mathbf{v}_0 Y$.

3.4 Welfare Ordering

We turn to welfare implications carried by the uniform supermajorization. In the case of uniform majorization, it is well known that if and only if $X \prec Y$ holds, then $\sum_{i=1}^n u(\mathbf{x}_i) \geq \sum_{i=1}^n u(\mathbf{y}_i)$ holds for all concave functions (e.g. Kolm, 1977; Karlin and Rinott, 1983; Mosler, 1994; Dahl, 1999). In what follows, we will show that a similar welfare implication holds for uniform supermajorization. To this end, we present the following lemmas.

Lemma 2 (Gale 1960, Theorem 2.8). Let A and $\mathbf{b}$ be an m-by-n matrix and m-dimensional vector, respectively. Exactly one of the following alternatives holds. Either the inequality

$$A\mathbf{x} \leq \mathbf{b},$$

has a nonnegative solution $\mathbf{x} \in \mathbb{R}^n_+$, or the inequalities

$$\mathbf{v}A \geq 0,$$

$$\mathbf{v}\mathbf{b} < 0,$$

have a nonnegative solution $\mathbf{v} \in \mathbb{R}^m_+$.

Proof. See Gale (1960). ∎

This Lemma can be interpreted as a separation theorem of convex sets. The next lemma characterizes a property of piecewise linear function with concavity.

Lemma 3. Let

$$V \equiv \{[c_1, \mathbf{v}_1], ..., [c_s, \mathbf{v}_s] : c_i \in \mathbb{R}, \mathbf{v}_i \in \mathbb{R}_+\},$$

be a set of s-tuples of $m+1$- dimensional row vectors where $\mathbf{v}_i$ is nonnegative. For $\mathbf{x} \in \mathbb{R}^m_+$,

$$u(\mathbf{x}) = \min_{[c_i, \mathbf{v}_i] \in V} \left\{ [c_i, \mathbf{v}_i] \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix} \right\}, \tag{7}$$

is a nondecreasing concave function in $\mathbf{x}$.

Proof. Let $\mathbf{x}'$, $\mathbf{x}''$, and $\lambda \in [0, 1]$. Concavity follows from a direct calculation as

$$\begin{aligned} u[\lambda \mathbf{x}' + (1-\lambda)\mathbf{x}''] =& \min_{[c_i, \mathbf{v}_i] \in V} \left\{ [c_i, \mathbf{v}_i] \left(\lambda \begin{bmatrix} 1 \\ \mathbf{x}' \end{bmatrix} + (1-\lambda) \begin{bmatrix} 1 \\ \mathbf{x}'' \end{bmatrix} \right) \right\} \\ \geq& \lambda \min_{[c_i, \mathbf{v}_i] \in V} \left\{ [c_i, \mathbf{v}_i] \begin{bmatrix} 1 \\ \mathbf{x}' \end{bmatrix} \right\} + (1-\lambda) \min_{[c_i, \mathbf{v}_i] \in V} \left\{ [c_i, \mathbf{v}_i] \begin{bmatrix} 1 \\ \mathbf{x}'' \end{bmatrix} \right\} \\ =& \lambda u(\mathbf{x}') + (1-\lambda) u(\mathbf{x}''). \end{aligned}$$

Since $\mathbf{v}_i \geq 0$, u is nondecreasing in $\mathbf{x}$. ∎

The utility index addressed in (7) can be interpreted as a piecewise linear approximation of general nondecreasing concave utility index. Now, we state the main theorem. The proof of proposition is motivated by Dahl (1999, Theorem 3.3).

Proposition 1. Let $X, Y \in \mathbb{R}^{m,n}$ be m-by-n matrices consisting of n-dimensional column vectors $\mathbf{x}_i$ and $\mathbf{y}_i$, respectively. The following two conditions are equivalent:

(i) $X \prec^w Y$.

(ii) $\sum_{i=1}^{n} u(\mathbf{x}_i) \geq \sum_{i=1}^{n} u(\mathbf{y}_i)$ holds for all $u \in \Omega_u$.

Proof. $(i) \Rightarrow (ii)$ follows from the nondecreasingness and concavity of u. Now we turn to $(ii) \Rightarrow (i)$. Let us consider a situation where

$$YP \nleq X, \tag{8}$$

for all $P \in \Psi_n$. Vectorizing (8), we obtain an equivalent condition as follows[14]:

$$\tilde{A}\hat{\mathbf{p}} \leq \tilde{\mathbf{b}}, \tag{9}$$

does not have nonnegative solution $\hat{\mathbf{p}}$, where

$$\tilde{A} \equiv \begin{bmatrix} I_n \otimes Y \\ \mathbf{e}_n \otimes I_n \\ -\mathbf{e}_n \otimes I_n \\ I_n \otimes \mathbf{e}_n \\ -I_n \otimes \mathbf{e}_n \end{bmatrix}, \quad \tilde{\mathbf{b}} \equiv \begin{bmatrix} \operatorname{vec} X \\ \mathbf{e}_n^T \\ -\mathbf{e}_n^T \\ \mathbf{e}_n^T \\ -\mathbf{e}_n^T \end{bmatrix}.$$

In the above expressions, $\otimes$ and vec stand for the Kronecker product and column stacking operators, respectively. In addition, $\mathbf{e}_n$ denotes an n-dimensional row vector whose entries are one, and I_n is an n-dimensional identity matrix. From Lemma 2, it can be verified that

$$\mathbf{v}\tilde{A} \geq \mathbf{0}, \tag{10}$$

and

$$\mathbf{vb} < \mathbf{0}, \tag{11}$$

have a nonnegative solution. Let $\mathbf{v}^* \equiv [\mathbf{v}_1^*, ..., \mathbf{v}_n^*, \check{\mathbf{z}}^*, \hat{\mathbf{z}}^*, \check{\mathbf{c}}^*, \hat{\mathbf{c}}^*]$ be a solution of (10) and (11), where $\mathbf{v}_i^* \in \mathbb{R}_+^m$ for $i = 1, ..., n$ and $\check{\mathbf{z}}^*, \hat{\mathbf{z}}^*, \check{\mathbf{c}}^*, \hat{\mathbf{c}}^* \in \mathbb{R}^n$. Thus, from (10), we obtain $\mathbf{v}_i^*\mathbf{y}_j + z_j^* + c_i^* \geq 0$ for $i, j = 1, ..., n$, where $z_j^* \equiv \check{z}_j^* - \hat{z}_j^*$ and $c_i^* \equiv \check{c}_i^* - \hat{c}_i^*$. Furthermore, from (11), we have $\sum_{i=1}^{n} (\mathbf{v}_i^*\mathbf{x}_i + z_i^* + c_i^*) < 0$. Defining $V \equiv \{[c_1^*, \mathbf{v}_1^*], ..., [c_n^*, \mathbf{v}_n^*]\}$, we can easily verify that

$$\sum_{i=1}^{n} (\mathbf{v}_i^*\mathbf{x}_i + z_i^* + c_i^*) \geq \sum_{i=1}^{n} \min_{[c_j, \mathbf{v}_j] \in V} \left\{ [c_j^*, \mathbf{v}_j^*] \begin{bmatrix} 1 \\ \mathbf{x}_i \end{bmatrix} \right\} + \sum_{i=1}^{n} z_i^*,$$

$$\sum_{i=1}^{n} \min_{[c_j, \mathbf{v}_j] \in V} \left\{ [c_j^*, \mathbf{v}_j^*] \begin{bmatrix} 1 \\ \mathbf{y}_i \end{bmatrix} \right\} + \sum_{i=1}^{n} z_j^* \geq 0.$$

Thus, we have

[14] In this vectorizing procedure, P is converted to $\hat{\mathbf{p}}(\equiv \operatorname{vec} P)$.

$$\sum_{i=1}^{n} \min_{[c_j,\mathbf{v}_j]\in V} \left\{ [c_j^*, \mathbf{v}_j^*] \begin{bmatrix} 1 \\ \mathbf{y}_i \end{bmatrix} \right\} > \sum_{i=1}^{n} \min_{[c_j,\mathbf{v}_j]\in V} \left\{ [c_j^*, \mathbf{v}_j^*] \begin{bmatrix} 1 \\ \mathbf{x}_i \end{bmatrix} \right\}.$$

From Lemma 3, $u(\mathbf{x}) = \min_{[c_i,\mathbf{v}_i]\in V} \left\{ [c_i, \mathbf{v}_i] \begin{bmatrix} 1 \\ \mathbf{x} \end{bmatrix} \right\}$ is nondecreasing and concave. By considering a utility index that takes the form of (7), we can conclude that

$$\sum_{i=1}^{n} u(\mathbf{x}_i) < \sum_{i=1}^{n} u(\mathbf{y}_i), \tag{12}$$

holds. Therefore, if $\sum\limits_{i=1}^{n} u(\mathbf{x}_i) \geq \sum\limits_{i=1}^{n} u(\mathbf{y}_i)$ holds for all $u \in \Omega_u$ then $YP \leq X$ holds for some $P \in \Psi_n$. ■

In the literature of inequality, a SEF based on individual utility is sometimes employed. The result obtained by Proposition 1 does not change even if we adopt a two-step approach to obtain a welfare implication from multivariate distribution.[15]

Corollary 1. Let $X, Y \in \mathbb{R}^{m,n}$. The following conditions are equivalent.

(i) $X \prec^w Y$.

(ii) $SW(\mathbf{u}_X) \geq SW(\mathbf{u}_Y)$ holds for all $u \in \Omega_u$ and $SW \in \Omega_W$.

Proof. $(ii) \Rightarrow (i)$ is obvious because $\sum_{i=1}^{n} u(\mathbf{x}_i) \in \Omega_W$ and from Proposition 1. Conversely, suppose that $X \prec^w Y$ holds. From Proposition 1, we obtain that

$$\sum_{i=1}^{n} u(\mathbf{x}_i) \geq \sum_{i=1}^{n} u(\mathbf{y}_i),$$

holds for all $u \in \Omega_u$. Let V^{**} be a set such that $V^{**}(z_k) = \{[z_k, 0], [0, 1]\}$ for $z_k \in (0, \infty)$. From Lemma 3,

$$g\left(\mathbf{x}_i, z_k\right) = \min_{\mathbf{v}^{**}\in V^{**}(z_k)} \left\{ \mathbf{v}^{**} \begin{bmatrix} 1 \\ u(\mathbf{x}_i) \end{bmatrix} \right\}, \tag{13}$$

is nondecreasing concave. Thus, applying Proposition 1 to (13), we obtain $\sum_{i=1}^{n} g\left(\mathbf{x}_i, z_k\right) \geq \sum_{i=1}^{n} g\left(\mathbf{y}_i, z_k\right)$ for all $z_k \in (0, \infty)$. Let us define $\mathbf{u}_X^{\uparrow} \equiv [u(\mathbf{x}_1)^{\uparrow}, ..., u(\mathbf{x}_n)^{\uparrow}]$, whose elements are arranged in increasing order of $u(\mathbf{x}_i)$. We define $\mathbf{u}_Y^{\uparrow}$ in a similar manner. We choose z_k as $u(\mathbf{y}_k)^{\uparrow}$ for $k \in \{1, ..., n\}$. Substituting $z_k = u(\mathbf{y}_k)^{\uparrow}$ into (13) and summing, we can verified that

[15] For example, Trannoy and Weymark (2007) considered the properties of the generalized Lorenz dominance for a utility distribution.

$$\sum_{i=1}^{n} g[\mathbf{y}_i, u(\mathbf{y}_k)^{\uparrow}] = \sum_{i=1}^{k} u(\mathbf{y}_i)^{\uparrow} + (n-k)u(\mathbf{y}_k)^{\uparrow},$$

and

$$\begin{aligned}
\sum_{i=1}^{n} g[\mathbf{x}_i, u(\mathbf{y}_k)^{\uparrow}] &= \sum_{i=1}^{n} u(\mathbf{x}_i)^{\uparrow} - \sum_{i=1}^{n}[u(\mathbf{x}_i)^{\uparrow} - u(\mathbf{y}_k)^{\uparrow}]^{+} \\
&= \sum_{i=1}^{k} u(\mathbf{x}_i)^{\uparrow} + \sum_{i=k+1}^{n} \left\{ u(\mathbf{x}_i)^{\uparrow} - [u(\mathbf{x}_i)^{\uparrow} - u(\mathbf{y}_k)^{\uparrow}]^{+} \right\} \\
&\quad - \sum_{i=1}^{k}[u(\mathbf{x}_i)^{\uparrow} - u(\mathbf{y}_k)^{\uparrow}]^{+} \\
&\leq \sum_{i=1}^{k} u(\mathbf{x}_i)^{\uparrow} + \sum_{i=k+1}^{n} \min\{u(\mathbf{x}_i)^{\uparrow}, u(\mathbf{y}_k)^{\uparrow}\},
\end{aligned}$$

hold for $k = 1, ..., n$, where $a^{+} \equiv \max\{0, a\}$. Noting that $(n-k)z_k \geq \sum_{i=k+1}^{\tilde{n}} \min\left\{u(\mathbf{x}_i)^{\uparrow}, z_k\right\}$ and (12), we obtain $\sum_{i=1}^{k} u(\mathbf{x}_i)^{\uparrow} \geq \sum_{i=1}^{k} u(\mathbf{y}_i)^{\uparrow}$, for $k = 1, \ldots, n$. This inequality implies that $\mathbf{u}_Y$ weakly supermajorizes $\mathbf{u}_X$. Since SW is assumed to be increasing and Schur concave in $\mathbf{u}$,

$$SW(\mathbf{u}_X) \geq SW(\mathbf{u}_Y),$$

holds. ■

Proposition 1 and its corollary state that the Shorrocks Theorem can be extended to case of multivariate attributes. However, in practice, we have to compare situations involving populations of different size. In the next section, we will turn to this problem.

4 Different Population Sizes

4.1 SEF with Replication Invariance

In this section, we characterize uniform supermajorization under different population sizes. In the case of univariate attributes, the generalized Lorenz curve provides complete information to a SEF that satisfies replication invariance.[16] For the multivariate case, we can obtain the welfare implications based on a doubly stochastic matrix.

First, we rewrite the SEF (4) so as to make the replication explicit. Consider the following representation using the natural number μ:

[16] For further discussion, see Aboudi et al. (2010).

$$sw\left(u_J, \mu n_J\right) = sw(\underbrace{u_1, ..., u_1}_{\mu}, ..., \underbrace{u_{n_J}, ..., u_{n_J}}_{\mu}). \tag{14}$$

In (14), $\mu = 1$ implies a SEF with no scale adjustment. In order to maintain comparability, we focus on SEFs satisfying the following property.

Assumption 3. The SEF (14) is homogenous degree of zero in n_J: for $\mu > 0$,

$$sw\left(u_J, n_J\right) = sw\left(u_J, \mu n_J\right), \tag{15}$$

holds.

Assumption 3 is referred to as the replication invariance principle.[17] We denote the set of SEFs satisfying Assumption 2 and 3 as Ω_w. Clearly, $\Omega_w \subset \Omega_W$.

4.2 Uniform Supermajorization under Different Population Sizes

Let us replicate X and Y so as to have a population of $\tilde{n} \equiv n_X n_Y$ explained as follows. In order to obtain a replicated distribution of attributes, we define replication matrices G_Y and G_X as follows:

$$G_Y \equiv I_{n_Y} \otimes \mathbf{e}_{n_X} \in \mathbb{R}_+^{n_Y, \tilde{n}},$$

$$G_X \equiv I_{n_X} \otimes \mathbf{e}_{n_Y} \in \mathbb{R}_+^{n_X, \tilde{n}},$$

By using the replication matrix G_J, $J = X, Y$, we obtain the distribution of a $\tilde{n} \equiv n_X n_Y$ population based on X and Y. The replicated distributions X^* and Y^* are respectively defined as follows:

$$Y^* \equiv Y G_Y \in \mathbb{R}_+^{m, \tilde{n}}, \tag{16}$$

$$X^* \equiv Y G_X \in \mathbb{R}_+^{m, \tilde{n}}. \tag{17}$$

Thus, we can apply Proposition 1 to the distributions Y^* and X^*. However, it is difficult to solve the matrix inequality (6) due to the extreme size of the matrices. Fortunately, we can reduce the size of vector and matrices to investigate the welfare orderings. To this end, we define a compressed form of doubly stochastic matrix. This is essentially same as the matrix defined in Aboudi et al. (2010, Definition 3.1).

Definition 3 (Compressed Doubly Stochastic Matrix). A n_Y-by-n_X matrix Q is said to be a compressed doubly stochastic matrix for n_Y and n_X if the following properties are satisfied:

[17] For example, see Tsui (1999).

$$\mathbf{e}_{n_Y} Q = n_Y \mathbf{e}_{n_X}, \tag{18}$$

$$Q \mathbf{e}_{n_X}^T = n_X \mathbf{e}_{n_Y}^T, \tag{19}$$

and

$$Q \geq O. \tag{20}$$

The compressed doubly stochastic matrix is a nonnegative matrix whose column- and row-sum are respectively n_Y and n_X. Hereafter, we denote the set of all compressed doubly stochastic matrices for n_Y and n_X as $\Psi(n_Y, n_X)$. Clearly, $Q \in \Psi(n_Y, n_X)$ implies $Q^T \in \Psi(n_X, n_Y)$. In addition, we define a matrix $\tilde{X}$ as follows:

$$\tilde{X} \equiv n_Y X. \tag{21}$$

The following lemma gives an equivalent condition for uniform supermajorization when the populations differ in size.

Lemma 4. Let $X \in \mathbb{R}^{m,n_X}$, $Y \in \mathbb{R}^{m,n_Y}$, and $X^*, Y^* \in \mathbb{R}^{m,\tilde{n}}$. The following two conditions are equivalent:

(i) $X^* \prec^w Y^*$.

(ii) $YQ \leq \tilde{X}$ holds for some $Q \in \Psi(n_Y, n_X)$.

Proof. It is sufficient to prove that $Y^*P = X^* \Leftrightarrow YQ = \tilde{X}$ for $P \in \Psi_{\tilde{n}}$ and $Q \in \Psi(n_Y, n_X)$. Suppose that there exists $P \in \Psi_{\tilde{n}}$ such that $Y^*P = X^*$. Noting that $Y^* = YG_Y$, we obtain $Y^*PG_X^T = YG_YPG_X^T = X^*G_X^T$, where $G_YPG_X^T$ is an n_Y-by-n_X matrix with nonnegative elements. In addition, $G_YPG_X^T\mathbf{e}_{n_X}^T = n_X\mathbf{e}_{n_Y}^T$ and $\mathbf{e}_{n_Y}G_YPG_X^T = n_Y\mathbf{e}_{n_X}$ hold. Therefore, we obtain $G_YPG_X^T \in \Psi(n_Y, n_X)$. Furthermore, it is clear that $X^*G_X^T = n_YX$. Conversely, suppose that $YQ = \tilde{X}$ holds for some $Q \in \Psi(n_Y, n_X)$. We can easily confirm that $\tilde{G}_Y \equiv (1/n_X)G_Y^T$ and $\tilde{G}_X \equiv (1/n_Y)G_X^T$ are the generalized inverse matrices of G_Y and G_X, respectively.[18] Noting that $Y = Y^*\tilde{G}_Y$, we obtain $YQ = Y^*\tilde{G}_YQ$. Furthermore, we have $\tilde{X}\tilde{G}_X^T = X^*$. Therefore, $YQ = \tilde{X}$ implies that $Y^*\tilde{G}_YQ\tilde{G}_X^T = X^*$. In this equation, we can easily confirm that

$$\tilde{G}_YQ\tilde{G}_X^T\mathbf{e}_{\tilde{n}}^T = \tilde{G}_YQ\mathbf{e}_{n_X}^T = n_X\tilde{G}_Y\mathbf{e}_{n_Y}^T = \mathbf{e}_{\tilde{n}}^T,$$

and

$$\mathbf{e}_{\tilde{n}}\tilde{G}_YQ\tilde{G}_X^T = \mathbf{e}_{n_Y}Q\tilde{G}_X^T = n_Y\mathbf{e}_{n_X}\tilde{G}_X^T = \mathbf{e}_{\tilde{n}},$$

[18] A matrix $\tilde{G}$ is said to be a generalized inverse of G if and only if $\tilde{G}G\tilde{G} = G$ holds. See Rao and Mitra (1971, ch. 2).

hold. Moreover, it is obvious that $\tilde{G}_Y Q \tilde{G}_X^T \geq O$ holds. Hence, we obtain $\tilde{G}_Y Q \tilde{G}_X^T \in \Psi_{\tilde{n}}$. Thus, $Y^* P = X^* \Leftrightarrow YQ = \tilde{X}$ is proved. ■

Putting together Proposition 1 with Lemma 4, we have the following result.

Proposition 2. Let $X \in \mathbb{R}^{m,n_X}$ and $Y \in \mathbb{R}^{m,n_Y}$ be two distributions of attributes. The following two conditions are equivalent:

(i) There exists $Q \in \Psi(n_Y, n_X)$ satisfying

$$YQ \leq \tilde{X}. \tag{22}$$

(ii) $sw(\mathbf{u}_X, n_X) \geq sw(\mathbf{u}_Y, n_Y)$ for all $u \in \Omega_u$ and $sw \in \Omega_w$.

Proof. This follows from Proposition 1, Corollary 1, and Lemma 4. ■

If we consider a distribution consisting of one attribute e.g., income, we can infer the welfare implications for two distributions without relying on Proposition 2. That is, the generalized Lorenz dominance criterion can be easily checked by using the familiar graphical representation. However, in the present situation where individual's well-being depends on more than two attributes, it is difficult to obtain the welfare implications by geometrical procedures. In the next section, we show an algebraic procedure to confirm Proposition 2.

5 Empirical Procedures

5.1 Equivalent Linear Programming Model

In this section, we describe the empirical procedures to confirm Proposition 2. In the theory of majorization, Dahl (1999, Theorem 3.6) suggests a characterization of majorization using a linear programming model.[19] In line with his perspective, we propose a procedure for the empirical analysis.

First, vectorizing the matrix inequality (22), we obtain

$$[I_{n_Y} \otimes Y] \operatorname{vec} Q \leq \operatorname{vec} \tilde{X}. \tag{23}$$

Next, $Q \in \Psi(n_Y, n_X)$ implies that

$$\begin{bmatrix} \mathbf{e}_{n_X} \otimes I_{n_Y} \\ I_{n_X} \otimes \mathbf{e}_{n_Y} \end{bmatrix} \operatorname{vec} Q = \begin{bmatrix} n_X \mathbf{e}_{n_Y}^T \\ n_Y \mathbf{e}_{n_X}^T \end{bmatrix}, \tag{24}$$

and $\operatorname{vec} Q \geq 0$ must hold.

[19] See also Mosler and Scarsini (1991, Example 3.4).

From (23) and (24), we can consider the following linear programming problem.

Problem 1 (P1).

$$\min_{\mathbf{q}} \mathbf{a}\mathbf{q}, \tag{25}$$

subject to

$$A_{eq}\mathbf{q} = \mathbf{b}_{eq}, \tag{26}$$

$$\mathbf{q} \geq 0, \tag{27}$$

where

$$\mathbf{a} \equiv [\mathbf{0}_{\tilde{n}+mn_X}, \mathbf{w}_L] \in \mathbb{R}_+^{\tilde{n}+2mn_X},$$

$$A_{eq} \equiv \begin{bmatrix} I_{n_X} \otimes Y & I_{mn_X} & -I_{mn_X} \\ \mathbf{e}_{n_X} \otimes I_{n_Y} & O & O \\ I_{n_X} \otimes \mathbf{e}_{n_Y} & O & O \end{bmatrix} \in \mathbb{R}_+^{(m+1)n_X+n_Y, \tilde{n}+2mn_X},$$

$$b_{eq} \equiv \begin{bmatrix} \operatorname{vec} \tilde{X} \\ n_X \mathbf{e}_{n_Y}^T \\ n_Y \mathbf{e}_{n_X}^T \end{bmatrix} \in \mathbb{R}_+^{(m+1)n_X+n_Y},$$

and $\mathbf{w}_L \in \mathbb{R}_{++}^{mn_X}$.

For (P1), the dual problem can be written as $\max_{\tilde{\mathbf{q}}} \tilde{\mathbf{q}}\mathbf{b}_{eq}$ subject to $\tilde{\mathbf{q}}A_{eq} \leq \mathbf{a}$. It can be easily shown that the solution to the dual problem exists.[20] By the duality theorem, therefore, (P1) also has a bounded optimal solution. Thus, we have the following result.

Proposition 3. The following two conditions are equivalent:

(i) The optimal value of (P1) is zero.

(ii) $sw(\mathbf{u}_X, n_X) \geq sw(\mathbf{u}_Y, n_Y)$ holds for all $u \in \Omega_u$ and $sw \in \Omega_w$.

Proof. This is obvious from (23) and (24). ■

According to Proposition 3, we can easily verify the relationship of welfare dominance for two societies. Given two distributions X and Y, we first solve (P1) and obtain the optimal value. If the optimal value is equal to zero, then the social welfare in society X is higher than that in society Y in the sense of uniform supermajorization. In contrast, if

[20] It is obvious that $\tilde{\mathbf{q}} = \mathbf{0}$ is a feasible solution. We decompose the vector $\tilde{\mathbf{q}}$ as $\tilde{\mathbf{q}} = [\tilde{\mathbf{q}}_n, \tilde{\mathbf{q}}_v]$, where $\tilde{\mathbf{q}}_{\tilde{n}} \in \mathbb{R}^{\tilde{n}}$ and $\tilde{\mathbf{q}}_v \in \mathbb{R}^{n_X+n_Y}$. Since $\tilde{\mathbf{q}}_{\tilde{n}}$ is restricted by $\mathbf{0} \geq \tilde{\mathbf{q}}_{\tilde{n}} \geq -\mathbf{w}_L$, the optimal solution is bounded.

the optimal value is strictly positive, then we have two possibilities. One is that the social welfare in Y is higher than in X. The other is that neither society is superior in terms of social welfare. Interchanging the roles of X and Y in (P1), one can identify which is the case.

Of course, uniform supermajorization gives a preorder to distribution of attributes. Thus, we may face a situation where the welfare orderings between two distributions cannot be determined. Even in such a situation, solving (P1), we obtain useful information. Consider a situation in which $X \not\prec^w Y$ and $Y \not\prec^w X$. First, it should be noted that in (P1), the vector, $\mathbf{w}_L$, can be interpreted as shadow prices of the respective attributes. That is, the optimal value represents a minimum amount of resources to be added in situation X for achieving more preferable situation than that of Y. Thus, if the policy maker has information on the unit cost required for improving each attribute of individuals, he can easily find a strategy for improving the welfare based on the optimal solution of (P1). Second, the optimal vector of the dual problem provides the parameter of the utility index which violates the majorization criterion.

One difficulty of applying this procedure in practice lies in the size of data. In order to solve (P1), the number of variables is $\tilde{n} + 2mn_X$ and the number of constraints including nonnegativity constraints is $(3m+1)n_X + n_Y + \tilde{n}$. In this sense, this procedure will be appropriate for grouped data or sampled data with small size. However, recent progress in computer capabilities partially mitigates this difficulty.

5.2 Numerical Example

Consider two societies denoted as X and Y. In society X, there are five persons, whose attributes are denoted as follows:

$$X = \begin{bmatrix} 230 & 500 & 430 & 305 & 150 \\ 140 & 200 & 250 & 230 & 245 \\ 175 & 150 & 325 & 265 & 150 \end{bmatrix}.$$

Similarly, society Y has four persons, whose attributes are represented as follows:

$$Y = \begin{bmatrix} 100 & 50 & 800 & 300 \\ 300 & 400 & 100 & 50 \\ 50 & 400 & 250 & 150 \end{bmatrix}.$$

We set $\mathbf{w}_L = 500 \times \mathbf{e}_5$. Solving (P1), we obtain the optimal value as $\mathbf{aq}^* \approx 5000 > 0$. This means that the distribution X does not dominates Y.[21] Solving the dual problem of (P1), we have the following vectors[22]

[21] Clearly, Y does not uniform supermajorize X, since $(1/n_X)X\mathbf{e}_5^T \geq (1/n_Y)Y\mathbf{e}_4^T$ holds.

[22] The vectors $\mathbf{v}_i$ and the scalar c_i are generated by partitioning the optimal solution of the dual problem as $\tilde{\mathbf{q}} = [-\mathbf{v}_1, ..., -\mathbf{v}_5, \mathbf{z}, -c_1, ..., -c_5]$, where $\mathbf{z} \in \mathbb{R}^4$.

$$\begin{bmatrix} c_1 & c_2 & c_3 & c_4 & c_5 \\ \mathbf{v}_1^T & \mathbf{v}_2^T & \mathbf{v}_3^T & \mathbf{v}_4^T & \mathbf{v}_5^T \end{bmatrix} = \begin{bmatrix} -165568 & -3803 & 3727 & -3063 & -7368 \\ 500 & 0 & 0 & 2.928 & 17.623 \\ 382.363 & 1.485 & 1.973 & 0 & 11.304 \\ 0 & 49.969 & 19.650 & 38.236 & 37.106 \end{bmatrix}.$$

Let V be a set defined as $V \equiv \{[c_1, \mathbf{v}_1], ..., [c_5, \mathbf{v}_5]\}$. Take the SEF as

$$sw(\mathbf{u}_X, n_X) = \frac{1}{n_X} \sum_{j=1}^{n_X} \min_{[c_i, \mathbf{v}_i] \in V} \left\{ [c_i, \mathbf{v}_i] \begin{bmatrix} 1 \\ \mathbf{x}_j \end{bmatrix} \right\}.$$

Thus, we obtain $sw(\mathbf{u}_X, n_X) = 5726.37 < 5976.15 = sw(\mathbf{u}_Y, n_Y)$. It should be noted that for $j = 1, 2, 3$, $\mathbf{x}^j$ dominates $\mathbf{y}^j$ in the sense of GL. This example stresses the importance of simultaneous consideration of income and other attributes.

Detail expression of the optimal solution can be represented as follows:

$$\mathbf{aq}^* = \mathbf{w}_L \times \begin{bmatrix} 10 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Thus, if the attribute 1 for individual 1 in X is increased by $\Delta x_{11} = 10/n_Y = 10/4 = 2.5$, then the social welfare will be improved. Indeed, if we consider a modified distribution, X', in which the $\mathbf{x}_1$ term, $[230, 140, 175]^T$, is replaced by $[232.5, 140, 175]^T$, the optimal value of (P1) becomes zero. Furthermore, from the optimal solution, we can find a doubly stochastic matrix such that $YQ \leq \tilde{X}'$ holds as follows:

$$Q^* = \begin{bmatrix} 0.6 & 2.0 & 0.0 & 0.4 & 2.0 \\ 0.6 & 0.0 & 2.0 & 1.6 & 0.8 \\ 0.0 & 2.0 & 2.0 & 1.0 & 0.0 \\ 2.8 & 0.0 & 0.0 & 1.0 & 1.2 \end{bmatrix}.$$

It can be easily verified that $Q^* \in \Psi(n_Y, n_X)$.

6 Empirical Illustration Using Chinese Provincial Data

In this section, we illustrate the empirical procedure by applying to Chinese provincial data.[23] We investigate well-being of the residents during the period of 2005 to 2009. China, which exhibits the most rapid growth in GDP in this period, faces a serious income disparity between the coastal and inland areas. Since the implementation of the open door

[23] In what follows, we focus on the province-level divisions including province, autonomous area, and direct-controlled municipalities. Hereinafter, we refer these divisions as province.

policy, the Chinese government has promoted economic growth in urban areas under the policy that regions with most potential are preferentially developed. As a result, the economic disparity between the urban and rural areas has expanded.[24] After 2000, in order to reduce this regional inequality, the government of China has emphasized the development of inland areas.

An increase in regional inequality resulting from the rapid growth affects some aspects of well-being of individuals. The first is environmental degradation in industrialized areas. For example, urban air pollution causes a serious environmental problem, especially in the industrialized areas of northern and central China (Vennemo et al., 2009). In 11th five-year plan, China's government has set compulsory targets for environmental protection (Cao et al., 2009). Second, during the transition period from a command economy, accessibility to health care services has been impaired. The number of health care workers in rural areas has decreased by 35.9% (Liu et al., 2007) and senior level health professionals tend to concentrate in urban facilities (Liu et al., 2007; Bloom, 2001).

Therefore, focusing on the interregional disparities of income, environmental quality, and accessibility of health care services, we consider the welfare implications of the recent development of the Chinese economy.[25]

6.1 Modified Analytical Framework

The theoretical framework presented in the previous sections should be slightly modified since we employ the grouped data by province. Let $\mathbf{x}_i$ and $\mathbf{y}_i$ be the attributes of individual who lives province i. Individuals who live in a same province are regarded as having the same attributes. That is, the distribution of attributes can be written as follows:

$$X = [\underbrace{\mathbf{x}_1, ..., \mathbf{x}_1}_{n_{1(X)}}, ..., \underbrace{\mathbf{x}_r, ..., \mathbf{x}_r}_{n_{r(X)}}],$$

$$Y = [\underbrace{\mathbf{y}_1, ..., \mathbf{y}_1}_{n_{1(Y)}}, ..., \underbrace{\mathbf{y}_r, ..., \mathbf{y}_r}_{n_{r(Y)}}],$$

where $n_{i(J)}$ denotes the number of population in province i for the distribution J. Thus,

$$\sum_{i=1}^{r} n_{i(J)} = n_J,$$

where r is the total number of province. The distribution of attributes can be written compactly as follows:

[24] Tsui (1996) analyzed the interprovincial inequality of income in the post-1978 period. Lee (2009) argued that the source of regional inequality in output has shifted from intraprovincial to interprovincial inequality since the early 1980s. Fan and Sun (2008) also considered the interprovincial inequality in output by using recent statistical data.

[25] Brajer et al. (2010) and Groom et al. (2010) considered the relationship between income inequality and environmental quality. Zhang and Kanbur (2005) considered spatial inequality in education and health care in China.

$$\hat{X} \equiv [\mathbf{x}_1, ..., \mathbf{x}_r], \tag{28}$$

$$\hat{Y} \equiv [\mathbf{y}_1, ..., \mathbf{y}_r], \tag{29}$$

where $\mathbf{x}_i$ denotes the attributes of individual who lives province i.

The SEF can be written as follows:

$$sw(\mathbf{u}_J, \mathbf{n}_J) = sw(\underbrace{u_{1(J)}, ..., u_{1(J)}}_{n_{1(J)}}, ..., \underbrace{u_{r(J)}, ..., u_{r(J)}}_{n_{r(J)}}),$$

for $J = X, Y$, where $\mathbf{n}_J \equiv (n_{1(J)}, ..., n_{r(J)})$ denotes a vector consisting of the population number for each province for J. We assume that for $\mu > 0$, $sw(\mathbf{u}_J, \mathbf{n}_J) = sw(\mathbf{u}_J, \mu\mathbf{n}_J)$ holds, which implies Assumption 3 holds.

Second, the compressed doubly matrix is also slightly modified. Instead of the compressed doubly stochastic matrix defined by Definition 3, we consider a r-by-r matrix as follows.

Definition 4. An r-by-r matrix $\hat{Q}$ is said to be modified compressed doubly stochastic if the following conditions are satisfied:

$$\mathbf{e}_r\hat{Q} = \frac{1}{n_X}\mathbf{n}_X,$$

$$\hat{Q}\mathbf{e}_r^T = \frac{1}{n_Y}\mathbf{n}_Y^T,$$

and

$$\hat{Q} \geq O.$$

Let $\Psi(\mathbf{n}_Y, \mathbf{n}_X)$ be the set of modified compressed doubly stochastic matrices. That is, a matrix $\hat{Q} \in \Psi(\mathbf{n}_Y, \mathbf{n}_X)$ is a nonnegative square matrix whose row-sum is equal to the population share for Y and whose column-sum is equal to that for X.

Furthermore, we define a vector corresponding to (21) in the basic model.

$$\hat{X}^* \equiv \frac{1}{n_X}\left[n_{1(X)}\mathbf{x}_1, ..., n_{r(X)}\mathbf{x}_r\right] \in \mathbb{R}^r. \tag{30}$$

The following proposition is a modified version of uniform supermajorization.

Proposition 4. For the distributions $\hat{Y}$ and $\hat{X}$, the following two conditions are equivalent:

(i) $\hat{Y}\hat{Q} \leq \hat{X}^*$ holds for some $\hat{Q} \in \Psi(\mathbf{n}_Y, \mathbf{n}_X)$.

(ii) $sw(\mathbf{u}_X, \mathbf{n}_X) \geq sw(\mathbf{u}_Y, \mathbf{n}_Y)$ holds for all $u \in \Omega_u$ and $sw \in \Omega_w$.

Proof. It follows from Proposition 1 and Lemma 1. ■

Thus, based on Proposition 4, we can set up the linear programming model corresponding to (P1).

6.2 Data and Results

The data includes 22 provinces, 4 autonomous areas, and 4 direct-controlled municipalities consisting of Beijing, Tianjin, Shanghai and Chongqing.[26] We compile the data from the China Statistical Yearbook, 2006-2010. As a proxy variable of individual income, we employ per capita household consumption in each region, which is deflated by the GDP deflator.[27] Industrial and household SO_2 emissions per land area are used as a proxy variable representing environmental quality.[28] Accessibility to health care service, where we concentrate our attention on the physical accessibility, is measured by the number of doctors per thousand population for each province.

Summary statistics are represented in Table 1. From Table 1, we can confirm rapid growth in consumption. Except for the period from 2005 to 2006, the mean value of SO_2 decreased. In addition, the number of doctors increased during the period. Since the univariate GL is a necessary condition for uniform supermajorization, Table 1 shows that pair-wise comparison for 2005 and 2006 is nonrankable.

Before investigating uniform supermajorization, we briefly show the result of the item-by-item approach that implies the GL comparison of each attribute. Table 2 summarizes the results. As easily expected, the pair-wise comparison of household consumption indicates that recent situations are more preferable than those of earlier years except for 2007 and 2008. The distribution of environmental quality is improved after 2007. However, we cannot say that the allocation of medical resources represented by the number of doctors improved during the period. As a result, only three pairs, {2005, 2009}, {2007, 2009} and {2008, 2009}, are candidates of uniform supermajorization.

Table 3 shows the optimal value of (P1) for each pair of years. From Table 3, we can conclude that the distribution in 2009 uniformly supermajorizes the distributions of 2007 and 2008. Although the item-by-item approach implies that the distribution in 2009 is preferable to that in 2005, the optimal value of (P1) is different from zero, which means the 2009 distribution is not necessarily preferred to that of 2005.

As noted in the previous section, further investigation on the nonrankable distributions provides useful information. For example, let us compare the two distributions 2006 and 2009. From Table 3, the optimal value of (P1) is 0.0499. If we look at optimal vector $\mathbf{q}$ in detail, we can find $q_{1062} = 9.975 * 10^{-5}$, where q_{1062} corresponds to the number of doctors

[26] Tibet autonomous area is excluded from the sample due to lack of data.

[27] We also considered per capita GRP as the income variable, and the basic results are not affected.

[28] Since SO_2 emission negatively affects the well-being of people, we define environmental quality in province i as $E - SO_2 emission$, where E denotes the reference level of environmental quality, which is assumed to be the same among both provinces and periods. Our analysis does not depend on the value of E.

in Guizhou province, located in the southwestern part of China.[29] Based on this result, we can conclude that if in 2009 the number of doctors in Guizhou province was larger by 132 than the actual data, the distribution in 2009 would dominate that in 2006 in the sense of uniform supermajorization.[30]

7 Concluding Remarks

We have considered an extension of the GL criterion to the case of multivariate attributes. Based on the uniform majorization by Kolm (1977) and Marshall and Olkin (1979), we proposed a uniform supermajorization that allows for attributes with different means. It can be verified that uniform supermajorization has a similar implication to uniform majorization for welfare ordering. Furthermore, we show that this criterion will work under comparison of distributions with unequal population sizes. We also provide a procedure for empirical investigations which is in the standard form of linear programming problems.

The methodology proposed here has a clear welfare implication as well as GL. In addition, the result obtained from the linear programming problem may provide useful information to policy makers and researchers: the optimal value indicates the minimum value for achieving welfare superior distribution compared with the base period. If the appropriate shadow prices are available, the procedure reveals the cost for redistribution.

In the present paper, we do not provide procedures for statistical inference on uniform supermajorization. In order to connect the theoretical analysis with the practical investigation, it will be important to check the statistical significance of the results. We leave this for future research.

8 Appendix

8.1 Proof of Remark 3

First, vectorizing the inequality $YR^T \leq X$, we obtain

$$\breve{A}\mathbf{r} \leq \breve{\mathbf{b}}, \tag{A.1}$$

where

$$\breve{A} \equiv \begin{bmatrix} I_n \otimes Y \\ I_n \otimes \mathbf{e}_n \\ -I_n \otimes \mathbf{e}_n \end{bmatrix}, \quad \breve{\mathbf{b}} \equiv \begin{bmatrix} \operatorname{vec} X \\ \mathbf{e}_n^T \\ -\mathbf{e}_n^T \end{bmatrix},$$

[29] In the optimal vector of (P1), $\mathbf{q}_i$ for $i = 1, ..., r^2$, represents the elements of the compressed doubly stochastic matrix, and $\mathbf{q}_j$ for $j = r^2 + 1, ..., r^2 + mr$ represnts the slack variables in (6).

[30] The population of Guizhou province in 2009 is 39780 (thousands) which is 2.89% of the whole population in the country. If the number of doctors per thousand population were increased by $q_{1062}/2.89$, the distribution of 2009 would be preferable to that of 2006 in the sense of unform supermajorization. Thus, we obtain the number of doctors to be increased in Guizhou province as $(q_{1062}/2.89)\% \times 39780 \approx 132$.

and $\mathbf{r} = \operatorname{vec} R$. Thus, if there exists $R \in \Psi_n^R$ such as $YR^T \leq X$ then $\breve{A}\mathbf{r} \leq \breve{\mathbf{b}}$ has nonnegative solution $\mathbf{r}^*$. Now, suppose that (A.1) does not have nonnegative solution, which is equivalent to $YR^T \nleq X$ for all $R \in \Psi_n^R$. From Lemma 2 stated in Section 3.4, it can be verified that

$$\mathbf{v}\breve{A} \geq \mathbf{0}, \tag{A.2}$$

and

$$\mathbf{v}\breve{\mathbf{b}} < \mathbf{0}, \tag{A.3}$$

have a nonnegative solution $\mathbf{v}^* \in \mathbb{R}_+^{n(m+2)}$. We decompose $\mathbf{v}^*$ into the sub-vectors as $\mathbf{v}^* = [\mathbf{v}_1^*, ..., \mathbf{v}_n^*, \check{\mathbf{c}}^*, \hat{\mathbf{c}}^*]$ where $\mathbf{v}_i^* \in \mathbb{R}_+^m$ for $i = 1, ..., n$ and $\check{\mathbf{c}}^*, \hat{\mathbf{c}}^* \in \mathbb{R}_+^n$. From (A.2), we obtain $\mathbf{v}_i^*\mathbf{y}_j + c_i^* \geq 0$ for $i, j = 1, ..., n$, where $c_i^* \equiv \check{c}_i^* - \hat{c}_i^*$. It can be easily verified that

$$\sum_{i=1}^{n} \min_{1\leq k\leq n} (\mathbf{v}_i^*\mathbf{y}_k) + \sum_{i=1}^{n} c_i^* \geq 0. \tag{A.4}$$

On the other hand, (A.3) can be written as $\sum_{i=1}^n \mathbf{v}_i^*\mathbf{x}_i + \sum_{i=1}^n c_i^* < 0$. This implies

$$\sum_{i=1}^{n} \min_{1\leq k\leq n} (\mathbf{v}_i^*\mathbf{x}_k) + \sum_{i=1}^{n} c_i^* < 0. \tag{A.5}$$

Together (A.4) with (A.5), we have $\sum_{i=1}^n \min_{1\leq k\leq n} (\mathbf{v}_i^*\mathbf{y}_k) > \sum_{i=1}^n \min_{1\leq k\leq n} (\mathbf{v}_i^*\mathbf{x}_k)$. Thus, there exists at least one index i^* such that

$$\min_{1\leq k\leq n} (\mathbf{v}_{i^*}^*\mathbf{y}_k) > \min_{1\leq k\leq n} (\mathbf{v}_{i^*}^*\mathbf{x}_k),$$

holds. This implies that $\mathbf{v}_{i^*}^* X \not\prec^w \mathbf{v}_{i^*}^* Y$ because $(\mathbf{v}_{i^*}^*\mathbf{y}_k)_1^\uparrow > (\mathbf{v}_{i^*}^*\mathbf{x}_k)_1^\uparrow$. Therefore, if $\mathbf{v}X \prec^w \mathbf{v}Y$ holds for all $\mathbf{v} \in \mathbb{R}_+^m$ then $YR^T \leq X$ holds for some $R \in \Psi_n^R$. ■

References

[1] Aboudi, R., Thon, D., and Wallace, S.W. (2010). Inequality comparisons when the populations differ in size. *Journal of Economic Inequality*, 8, 47-70.

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

[3] Atkinson, A. B., and Bourguignon, F. (1987). Income distribution and differences in needs. In: G. R. Feiwel (ed.), *Arrow and the Foundations of the Theory of Economic Policy* (pp. 350-370). London: Macmillan.

[4] Bhatia, R. (1997). *Matrix analysis*. New York: Springer.

[5] Bishop, J.A., Formby, J.P., and Smith, W.J. (1993). International comparisons of welfare and poverty: dominance orderings for ten countries. *Canadian Journal of Economics*, 26, 707-726.

[6] Blackorby, C., Bossert, W., and Donaldson, D. (2009). Population ethics. In P. Arnold, P. Pattanaik, and C. Puppe (eds.), *Handbook of Rational and Social Choice* (pp. 483-500). New York: Oxford University Press.

[7] Bloom, G. (2001). Equity in health in unequal societies: meeting health needs in contexts of social change. *Health Policy*, 57, 205-224.

[8] Bourguignon, F. (1999). Comment to 'Multidimensional Approaches to Welfare Analysis' by Maasoumi, E. In J. Silber (ed.), *Handbook of Income Inequality Measurement* (pp.477-484). Boston: Kluwer.

[9] Bourguignon, F., and Chakravarty, S. R. (2003). The measurement of multidimensional poverty. *Journal of Economic Inequality*, 1, 25-49.

[10] Brajer, V., Mead, R. W., and Xiao, F. (2010). Adjusting Chinese income inequality for environmental equity. *Environment and Development Economics*, 15, 341-362.

[11] Brandolini, A. (2008). *On applying synthetic indices of multidimensional well-being: health and income inequalities in selected EU countries*. Bank of Italy Temi di Discussione No. 668.

[12] Cao, J., Garbaccio, R., and Ho, M.S. (2009). China's 11th five-year plan and the environment: reducing SO_2 emissions. *Review of Environmental Economics and Policy*, 3, 231-250.

[13] Chiou, J-R. (1996). A dominance evaluation of Taiwan's official income distribution statistics, 1976-1992. *China Economic Review*, 7, 57-75.

[14] Dahl, G. (1999). Matrix majorization. *Linear Algebra and its Applications,* 288, 53-73.

[15] Fan, C. C., and Sun, M. (2008). Regional inequality in China, 1978-2006. *Eurasian Geography and Economics*, 49, 1-20.

[16] Gale, D. (1960). *The theory of linear economic model.* Chicago: University of Chicago Press.

[17] Gajdos, T., and Weymark, J. A. (2005). Multidimensional generalized Gini indices. *Economic Theory*, 26, 471-496.

[18] Groom, B., Grosjean, P., Kontoleon, A., Swanson, T., and Zhang, S. (2010). Relaxing rural constraints: a 'win-win' policy for poverty and environment in China. *Oxford Economic Papers*, 62, 132-156.

[19] Kakwani, N. C. (1984). Welfare ranking of income distributions. In R. L. Basman, and G. F. Rhodes, G.F. (eds.), *Advances in Econometrics* (pp. 191-213). Greenwich, CT: JAI Press.

[20] Karlin, S., and Rinott, Y. (1983). Comparison of measures, multivariate majorization and application to statistics. In S. Karlin, T. Amemiya, and L. A. Goodman (eds.), *Studies in Econometrics, Time Series, and Multivariate Statistics* (pp. 465-489). New York: Academic Press.

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

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

[23] Koshevoy G., and Mosler, K. (2007). Multivariate Lorenz dominance based on zonoids. *AStA Advances in Statistical Analysis*, 91, 57-76.

[24] Lambert, P. J. (1993). *The distribution and redistribution of income: a mathematical analysis*, 2nd ed. Manchester: Manchester University Press.

[25] Lee, J. (2009). Changes in the source of China's regional inequality. *China Economic Review*, 11, 232-245.

[26] List, C. (1999). *Multidimensional inequality measurement: a proposal.* Nuffield College Working Paper in Economics, No. 1999-W27.

[27] Liu, M., Zhang, Q., Lu, M., Kwon, C-S., and Quan, H. (2007). Rural and urban disparity in health services utilization in China. *Medical Care*, 45, 767-774.

[28] Lugo, M. A. (2007). Comparing multidimensional indices of inequality: methods and application. *Research on Economic Inequality*, 14, 213-236.

[29] Martínes Pería, F. D., Massey, P. G., and L. E. Silvestre. (2005). Weak matrix majorization. *Linear Algebra and its Applications*, 403, 343-368.

[30] Maasoumi, E. (1986). The measurement and decomposition of multi-dimensional inequality. *Econometrica*, 54, 991-997.

[31] Marshall, A.W., Olkin, I., Arnold, B. C. (2011). *Inequalities: Theory of Majorization and its Applications.* 2nd. ed. New York: Springer.

[32] Mosler, K. (1994). Majorization in economic disparity measures. *Linear Algebra and its Applications*, 199, 91-114.

[33] Mosler, K. and Scarsinin, M. (1991). Some theory of stochastic dominance. In K. Mosler and M. Scarsini (eds.), *Stochastic orders and decision under risk* (pp.261-284). Institute of Mathematical Statistics Lecture Notes-Monograph Series 19. Hayward: Institute of Mathematical Statistics.

[34] Mukhopadhaya, P. (2003). The ordinal and cardinal judgment of social welfare changes in Singapore, 1982-99. *Developing Economies*, 41, 65-87.

[35] Muller, C., and Trannoy, A. (2011). A dominance approach to the appraisal of the distribution of well-being across countries. *Journal of Public Economics*, 95, 239-246.

[36] Nilsson, T. (2010). Health, wealth and wisdom: exploring multidimensional inequality in a developing country. *Social Indicators Research* 95, 299-323.

[37] Rao, C.K., and Mitra, S. K. (1971). *Generalized inverse of matrices and its applications.* New York: John Wiley and Sons.

[38] Savaglio, E. (2011). On multidimensional inequality with variable distribution mean. *Journal of Mathematical Economics*, 47, 453-461.

[39] Saposnik, R. (1993). A note on majorization theory and the evaluation of income distributions. *Economics Letters*, 42, 179-183.

[40] Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic orders and their applications.* San Diego: Academic Press.

[41] Shorrocks, A. F. (1983). Ranking income distributions. *Economica*, 50, 1-17.

[42] Trannoy, A., and Weymark, J.A. (2007). Dominance criteria for critical-level generalized utilitarianism. Working Paper No. 07-W07, Department of Economics, Vanderbilt University.

[43] Tsui, K. (1995). Multidimensional generalization of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach. *Journal of Econnomic Theory* 67, 251-265.

[44] Tsui, K. (1996). Economic reform and interprovincial inequalities in China. *Journal of Development Economics*, 50, 353-368.

[45] Tsui, K. (1999). Multidimensional inequality and multidimensional generalized entropy measures: an axiomatic derivation. *Social Choice Welfare*, 16, 145-157.

[46] Vennemo, H., Aunan, K., Lindhjem, H., and Seip, H. M. (2009). Environmental pollution in China: status and trends. *Review of Environmental Economics and Policy*, 3, 209-230.

[47] Weymark, J. A. (2006). The normative approach to the measurement of multidimensional inequality. In F. Farina and E. Savaglio (eds.), *Inequality and economic integration* (pp.303-328). New York: Routledge.

[48] Zhang, X., and Kanbur, R. (2005). Spatial inequality in education and health care in China. *China Economic Review*. 16, 189-204.

Table 1 Descriptive statistics

	2005	2006	2007	2008	2009
Per capita household consumption (in ten thousands 2005 RMB)					
Mean	0.560	0.610	0.657	0.705	0.778
Std. Dev.	0.263	0.289	0.310	0.324	0.351
Min	0.330	0.354	0.382	0.366	0.421
Max	1.840	1.997	2.146	2.253	2.450
SO_2 emissions per provincial area (t / km^2) [a)]					
Mean	7.277	7.287	6.958	6.473	6.082
Std. Dev.	7.654	7.589	7.450	6.763	5.906
Min	0.173	0.181	0.187	0.188	0.189
Max	62.265	61.658	60.422	54.145	45.989
Number of doctors per thousand population [b)]					
Mean	1.510	1.545	1.550	1.592	1.770
Std. Dev.	0.376	0.380	0.382	0.388	0.413
Min	1.080	1.095	1.007	1.024	1.092
Max	3.293	3.339	3.385	3.484	3.581
Total population (in ten thousands)					
	128046	128850	129635	130540	131371
GDP deflator (2005=100)					
	100.000	103.807	111.735	120.410	119.669 [c)]
Number of provinces and municipalities					
	30	30	30	30	30

Source: National Bureau of Statistics of China, *China Statistical Yearbook*.
Notes: a). Based on the total amount of sulphur dioxide emissions by industry and consumption.
b). Based on the total number of licensed doctor including assistant doctor.
c). Estimated value by IMF, *The World Economic Outlook*.

Table 2 Univariate GL dominance

		Given Year: *x*				
		2005	2006	2007	2008	2009
Base year: *y*	2005	-	C, D	C, E	C, E	C, E, D
	2006	none	-	C, E	C, E	C, E
	2007	none	none	-	E, D	C, E, D
	2008	none	none	none	-	C, E, D
	2009	none	none	none	none	-

Notes: C=household consumption, E= SO_2 emission, D=number of doctors. In each cell, J (=C, E, D) indicates variable J in the base year is GL dominated by that in the given year.

Table 3 Optimal value of the linear programming model (P1)[a)]

		Given Year: X				
		2005	2006	2007	2008	2009
Base year: Y	2005	-	4.905E+00	1.349E+00	1.020E+00	8.854E-03
	2006	7.858E+01	-	3.471E+00	1.208E+00	4.987E-02
	2007	2.304E+02	1.944E+02	-	6.595E-01	**4.930E-11**
	2008	5.161E+02	4.792E+02	2.875E+02	-	**3.622E-13**
	2009	8.364E+02	7.984E+02	6.081E+02	3.209E+02	-

Notes: a). $\boldsymbol{w}_L$=500. Calculations have been implemented by the 'linprog' function of MATLAB. TolFun, which is a lower bound on the change in the value of the objective function during a step, is set at 10E-8.