Measurement of Inequality
with a Finite Number of Pay States:
The Majorization Set and its Applications

Ramses H. Abul Naga [*]

October 10, 2016

Abstract

In this paper we examine the Lorenz ordering when the number of pay states is finite, as is most often the case in public sector employment. We characterize the majorization set: the set of pay scales such that some distribution u is more egalitarian than another distribution v, with u and v being two distributions of a given sum total. We show that while this set is infinite, it is generated as the convex hull of a finite number of points. We then discuss several applications of the result, including the problem of reducing inequality between groups, conditions under which different pay scales may reverse the ordering of two Lorenz curves, and the use of the majorization set in relation to optimal income taxation.

Keywords: Measurement of inequality, discrete data, majorization set, majorization matrix, convex sets and their separation.

JEL codes: I3, I1.

[*]Business School and Health Economics Research Unit, University of Aberdeen, Aberdeen AB24 3QY, UK. *email*: r.abulnaga@abdn.ac.uk.

1. Introduction

In recent years there has been a growing interest in measuring inequality in relation to qualitative data such as self assessed health, happiness and life satisfaction more broadly. The research community has provided new methods of analyzing qualitative data [1] alongside a well established literature on the analysis of continuous income data, founded on generalizations of the Lorenz ordering and related income inequality measures (Atkinson, 1970; Ok, 1996; Foster and Shneyerov, 1999; Gadjos and Weymark, 2005; Savaglio, 2006; and del Río and Alonso-Villar, 2010). Yet the problem of measuring inequality in presence of a finite number of discrete pay states is yet to be examined in fuller detail. The analysis of this problem, involving discrete data–the way qualitative data are discrete–yet cardinal–the way income data are cardinal–is the purpose of the current paper.

An economic theory of competitive earnings differentials predicts that earnings rise with the disutility of work associated with various occupations, or various roles within the same enterprise. A tournament theory of internal labour markets also produces a discrete set of pay states, where "...the hierarchical structure of the firm rests on a given grid of remunerations in which wages change discontinuously, and only with promotion ." (Cahuc et al. 2014, p. 365.) Accordingly, that income or earnings present themselves in a discrete and finite number of states is certainly often the norm in public sector employment. For instance, Britain's Senior Civil Service has four pay states, namely Assistant Secretary, Deputy Secretary, Permanent Secretary and Head of Civil Service (Department of Finance and Personnel, 2014). Additionally, there are annual pay increments between the first and second, second and third and third and fourth of these pay classes. A similar pay structure may be found in the government sector of other EU countries.

The particular perspective we shall follow on the measurement of inequality with a finite number of pay states will be that of characterizing the *majorization set*: that is, the entire set of pay scales that allow for a Lorenz curve ranking of two distributions (say public sector workers in the London area versus those of other parts of Britain.) In the mathematical sciences, Brualdi (1984) characterizes the set of doubly stochastic matrices given two fixed vectors u and v, that enable one to express u as the product of a doubly stochastic matrix and the vector v, and calls the resulting set the majorization polytope. Following this lead, the present

---

[1]See for instance Allison and Foster (2004), Apouey (2007), Abul Naga and Yalcin (2008), Zheng (2011), Gravel and Moyes, (2012), Chakravarty and Zoli (2012), Cowell and Flachaire (2012), Kobus and Miłoś (2012), Andreoli and Zoli (2013), and Gravel et al. (2014).

paper explores the concept of the majorization set in economics. Consider the general problem of selecting parameter values for a policy vector (this could be a pay scale, a combination of an income allowance and a uniform tax rate, etc.) in a way as to reduce inequality over an initial situation. The majorization set is the set of policy vectors that achieve an ordering of Lorenz curves of a benchmark distribution and the distribution of interest.

The characterization of the majorization set and its related properties can be informative for examining several existing problems in the income distribution literature. The first result of the paper is that the majorization set is a closed and bounded type of convex set known as a polytope [2]. Such a set is the convex hull of a finite number of points (pay scales), known as the *vertices* of the majorization set. Central to this result, is a *majorization matrix* that we introduce in the paper, that expresses the partial sums (inequalities) defining Lorenz dominance. The vertices of the majorization set may be thought of as policy benchmarks, providing the limits of feasible pay policies. The vertices are then jointly sufficient statistics for generating the majorization set. The second result of the paper is that the majorization set and its dual set (the set of pay scales that allow for the reverse ordering of the same two distributions) are separable convex sets. The coordinates of the separating hyperplanes are provided by the non-zero rows of the majorization matrix.

Equipped with these results, we discuss five applications of the majorization set. Firstly we revisit the problem of characterizing the set of income transformations that preserve the Lorenz ordering of two distributions, a question studied by Moyes (1994) in the context of continuous income data. We exploit the convexity of the majorization set to revisit this problem in the context of a finite number of pay scales, where we establish a simple relation between the set of transformations of the pay scale that preserve the Lorenz ordering of a pair of distributions and the vertices of the majorization set. Following Bourguignon (1979) and Shorrocks (1980), it is now a well established practice in empirical work to express inequality in a population as a function of inequality within groups and inequality between groups. One important question to address therefore in presence of a finite number of pay states is how different pay scales impact on within groups and between groups inequality separately. We show how the answer to this question relates to the rank of the majorization matrix.

Thirdly, the results of this paper enable us to shed light on an important prob-

---

[2] A polytope may be thought of as a generalization of the concept of a polygon in higher dimensional spaces. An exact definition of a polytope will be given below.

lem discussed in relation to the measurement of inequality with qualitative data. It was first recognized by Allison and Foster (2004) and Zheng (2011) that when the numerical scale defining the economic status of individuals is undetermined, two different numerical scales may reverse the ordering of a pair of Lorenz curves. Using the framework of the majorization set, we are able to characterize necessary and sufficient conditions for Lorenz curve reversals *not* to occur. These conditions are expressed as inequalities involving products of the majorization matrix and vectors defining the vertices of a set that contains the majorization set.

It is evident that in practice incentive, participation, and other institutional constraints provide theoretical limits to income redistribution (Roberts, 1984; Hindriks and Myles, 2013 ch. 18). Incentive contracts typically place restrictions on the size of increments between successive pay states. Our fourth application therefore examines how the majorization set may shrink in the light of incentive considerations in the context of tournament theories of pay and promotion. Our fifth and final application is to discuss the structure of the majorization set in the context of the design of a linear income tax scheme. This final application is motivated by the need to illustrate that applications of the majorization set are not confined to the design of equitable pay schemes, and that this framework is equally relevant in all aspects of income redistribution policies such as in the design of social insurance or differential commodity taxation.

To set the scene for the problem that will occupy us in much of Section 3, consider the following hypothetical example. Suppose you are manager of two big-band orchestras, comprising 30 musicians each. The two orchestras differ in that their brass, keyboard, and rhythm sections have different numbers of individuals. You cannot change the composition of these two orchestras; but you are free to choose pay scales (that will apply to both orchestras) subject to two types of restrictions. Firstly, pay is ordered in a given fashion: a brass musician must not earn less than a keyboard player, who must not earn less than a rhythm section musician. Secondly, you must choose the pay structure in a way that the lowest pay cannot be negative, and the highest pay must be finite. You recognize that the pay structure you will adopt will impact directly on the welfare of your orchestras, and you are interested in studying how a particular pay scale will impact on the ordering of the two pay distributions in terms of inequality. This paper aims at characterizing the set of pay structures such that the distribution of earnings in one orchestra is more egalitarian than in the other orchestra.

The next section of the paper contains definitions of some key concepts. Section 3 presents the two central results pertaining to the majorization set. Section

4 discusses areas of application of the majorization set, while Section 5 concludes the paper.

2. Notation and definitions

In this paper we study the problem of measuring inequality in the context of an economy where the economic status of a person $j = 1, ..., l + 1$ is defined over $i = 1, ..., n$ possible states. We associate with these n states a column vector $x = (x_1, ..., x_n)'$ and we refer to x as a pay scale. There are two assumptions carried over throughout the results we shall obtain in this paper:

Assumption 1 There are n economic states ordered from 1 (the lowest) to n (the highest.)

Assumption 2 Any given pay scale $x = (x_1, ..., x_n)'$ is bounded from below by 0, that is $x_1 \geq 0$, and is bounded from above by a finite number $x_{\max}$, such that $x_n \leq x_{\max}$.

That is, any pay scale x is an element of the set of increasing scales:

$$\mathcal{C} = \left\{ x \in \mathbb{R}^n : 0 \leq x_1 \leq x_2 \leq ... \leq x_n \leq x_{\max} \right\}. \tag{2.1}$$

Throughout, the letters h, i, j, k, l, m, n are all used to refer to natural numbers. We let $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$ respectively denote the sets of natural numbers, integers and real numbers. With the exception of $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$, we will let $\mathcal{S}$ denote a set, S denote a matrix, and s denote a vector. Unless otherwise specified, all vectors are implicitly defined as being column vectors. If a and b are two column vectors, then a' being the transpose of a, $a'b$ is the inner product of a and b. If M denotes a matrix, the null space of M is written as $\mathcal{N}ull(M)$. We define $\iota_m = (1, ..., 1)'$ as an $m-$dimensional vector of ones, so that for any vector $v \in \mathbb{R}^m$, $\iota'_m v = v' \iota_m$ is equal to the sum of the elements of v. We also define $\mathbf{0}_m$ as an $m-$dimensional vector of zeroes, and $\mathbf{0}_{m,n}$ as a $m \times n$ matrix of zeroes. The notation $v \leq \mathbf{0}_m$ signifies that every element of v is non-positive.

A distribution is some vector $u = (u_1(x), ..., u_{l+1}(x))'$. As our judgement on the level of inequality is anonymous, we adopt the convention that the vector u is ordered from richest to poorest, and we write $u \in \mathcal{D}$, where

$$\begin{aligned}
\mathcal{D} = \{ (u_1, ..., u_{l+1})' \in (x_1, ..., x_n)^{l+1} : \\
u_1 \geq u_2 \geq ... \geq u_{l+1} \}
\end{aligned} \tag{2.2}$$

is the set of distributions that is the subject of our study in this paper. It is also convenient to summarize our data with counts. We define the function count$(.)$:

$\mathcal{D} \to \mathbb{Z}^{1 \times n}$, where the *ith* element of count(u) equals m_i, m_i being the number of realizations of state i in the distribution u. Stated more simply, count(u) is an n-dimensional row vector and count$(u)/(l+1)$ is the frequency distribution associated with u.

Each individual j is thus associated with a particular economic state i. Thus, while count(u) is fixed for each given distribution $u \in \mathcal{D}$, it is clear that u is a function of x, so that for two distinct scales $x^1, x^2 \in \mathcal{C}$, it is generally the case that $u(x^1) \neq u(x^2)$. For this reason, if v denotes another distribution, $u = v$ will signify that $u(x) = v(x)$ for all $x \in \mathcal{C}$, or equivalently, count$(u) = $ count(v).

Our central interest throughout is in the Lorenz ordering, the Hardy-Littlewood-Polya relation defined over $(l+1)-$dimensional vectors $u, v \in \mathcal{D}$. Indeed, this is the building block of the income inequality literature, as well as being at the heart of the majorization literature in the mathematical sciences [3]. For a given $x \in \mathcal{C}$, u has less inequality than v, written $u(x) \prec v(x)$ if and only if the following $l+1$ conditions are satisfied:

$$\sum_{j=1}^{k} u_j(x) \leq \sum_{j=1}^{k} v_j(x) \qquad k = 1, ..., l \tag{2.3}$$

$$\sum_{j=1}^{l+1} u_j(x) = \sum_{j=1}^{l+1} v_j(x) \tag{2.4}$$

We call the above set of equations *the $l+1$ partial sums.* Note that given the above definition, either $u(x) \prec v(x)$, $v(x) \prec u(x)$, or $u(x)$ and $v(x)$ are not comparable, in which case we write $u(x) \parallel v(x)$.

Because the above partial sums define a system of $l+1$ linear inequalities in the vector x, it is possible to express them in compact matrix notation. Specifically, letting A denote an $l \times n$ matrix and letting δ denote an n-dimensional row vector, the inequalities (2.3–2.4) may be written as follows:

$$Ax \leq \mathbf{0}_l \tag{2.5}$$

$$\delta x = 0 \tag{2.6}$$

---

[3]Beyond the measurement of inequality in the social sciences, the majorization ordering finds many applications in the mathematical sciences. These include combinatorial analysis, linear algebra and eigen-value analysis, probability and statistics as well as geometric inequalities. See Ando (1994) for a survey, as well as Marshall et al. (2010).

To clarify this notation, it is useful to consider a simple example.

Example 1 Assume there are four individuals ($l + 1 = 4$) and $n = 6$ pay states and consider two distributions $u = (x_5, x_5, x_3, x_2)'$ and $v = (x_6, x_4, x_4, x_1)'$. Then in this context, the pay scale $x \in \mathcal{C}$ renders $u(x)$ more egalitarian than $v(x)$, provided the following four inequalities are satisfied:

$$x_5 \ \leq \ x_6 \qquad\qquad (2.7)$$
$$2x_5 \ \leq \ x_6 + x_4 \qquad\qquad (2.8)$$
$$2x_5 + x_3 \ \leq \ x_6 + 2x_4 \qquad\qquad (2.9)$$
$$2x_5 + x_3 + x_2 \ = \ x_6 + 2x_4 + x_1 \qquad\qquad (2.10)$$

These inequalities are easily expressed in the matrix form (2.5-2.6) where in this context the matrix A has three rows, six columns, and is easily obtained:

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 1 & -2 & 2 & -1 \end{pmatrix}, \qquad\qquad (2.11)$$

while δ is the row vector

$$\delta = \begin{pmatrix} -1 & 1 & 1 & -2 & 2 & -1 \end{pmatrix}. \qquad\qquad (2.12)$$

$\bigcirc$

Consider then summing up the highest k incomes in each distribution, and taking their difference. This difference is clearly a linear function of x, and can be written as $\sum_{j=1}^{k} [u_j(x) - v_j(x)] = a_{k1}x_1 + \cdots + a_{kn}x_n$. The row vector $(a_{k1}, ..., a_{kn})$ then defines the kth row of the majorization matrix A. Likewise by writing $\sum_{j=1}^{l+1} [u_j(x) - v_j(x)] = \delta_1 x_1 + \cdots + \delta_n x_n$, the scalars $\delta_1, \cdots, \delta_n$ are the elements of the row vector δ. We formally define in the appendix of the paper a function $\Phi : \mathcal{D} \times \mathcal{D} \longrightarrow \mathbb{Z}^{l \times n}$ that maps u and v to the matrix A, and we write $A = \Phi(u, v)$. Likewise, we define a function Δ that maps u and v to the row vector $\delta = \Delta(u, v)$. We shall refer to the image A of the map Φ as the *majorization matrix*. As this matrix plays a central role in our analysis below, we first gather some properties of this matrix in the following lemma.

Lemma 2.1 (properties of the majorization matrix) *Let the majorization matrix $A = \{a_{ji}\}$, with jth row $a_j = (a_{j1}, ..., a_{jn})$ denote the image of (u, v) via the map Φ. Then, there holds the following:*

7

$[PA1]$ $A\iota_n = \mathbf{0}_l$ *where ι_n is an n-dimensional vector of ones.*

$[PA2]$ $A = \mathbf{0}$ *if and only if $u = v$.*

$[PA3]$ $1 \leq \dim \mathcal{N}ull(A) \leq n - 1$ *for all $u \neq v$.*

$[PA4]$ *Let $B = \Phi(v, w)$ and $C = \Phi(u, w)$. Then $C = A + B$.*

$[PA5]$ $a_{ji} \in \{-j, ..., 0, ..., j\}$ *for all $j = 1, ..., l$.*

$[PA6]$ *For all $j = 1, ..., l - 1$, if $a_j = (0, ..., 0)$, then $a_{j+1} \in \{-1, 0, 1\}^n$ and at most two elements of a_{j+1} are different from zero.*

$[PA7]$ $(a_{j1}, ..., a_{jn}) = (a_{j+1,1}, ..., a_{j+1,n})$ *if and only if $u_{j+1} = v_{j+1}$.*

Property $[PA2]$ states that A is generally a non-zero integer matrix, and is only identically zero when u equals v. $[PA4]$ is a property of linearity of the map $\Phi(.,.)$. $[PA5]$ indicates that any element of the jth row of A is no smaller than $-j$ and no larger than j. On the other hand, $[PA6]$ reveals that if the $(j-1)th$ row of A is equal to zero, then each element of the jth row takes one of three values: -1, 0 and 1. $[PA7]$ indicates that the jth and $(j+1)th$ rows of A are identical when the elements u_{j+1} and v_{j+1} of the vectors u and v are identical. Finally, $[PA1]$ and $[PA3]$ together indicate that the rank of A cannot exceed $n - 1$, and that the vector of ones ι_n lies in the null space of A.

The set of solutions to a system of linear inequalities in $\mathbb{R}^n$; $\{x \in \mathbb{R}^n : Bx \leq \beta\}$ where B is an $k \times n$ matrix, is defined as a *polyhedral set*. Note that, by definition, every polyhedral set is the intersection of a finite number of closed half-spaces, and accordingly is a closed convex set. The convex hull of a finite set of points is called a *polytope*, or convex polytope, in this paper.

Let $\mathcal{S}$ be a set in $\mathbb{R}^n$. A point ξ in $\mathcal{S}$ is called an extreme point, if ξ cannot be expressed as a convex combination of any two other elements of $\mathcal{S}$. That is, $\xi = \lambda v + (1 - \lambda)\omega$ with $v, \omega \in S$ and $\lambda \in [0, 1]$ entails that $\xi = v = \omega$. We shall make repeated use of a theorem pertaining to convex sets, that we shall refer to as *Minkowski's theorem*. The theorem states that every closed, bounded and convex set $\mathcal{S} \subseteq \mathbb{R}^n$ is equal to the convex hull of its extreme points. The extreme points of convex polytopes are commonly referred to as vertices [4]. Further notation and definitions, specific to particular sections of the paper, will be introduced where needed.

---

[4] A triangle and its interior for instance is a convex polytope in $\mathbb{R}^2$ and its three vertices are the points of intersection of its three edges.

3. The majorization set

Our first task in this paper is to characterize the majorization set $\mathcal{C}_{uv} \subseteq \mathcal{C}$, such that $u(x) \prec v(x)$ for all $x \in \mathcal{C}_{uv}$. We shall approach this task by examining first the geometry of the set $\mathcal{C}$. We shall then study an intermediate set $\mathcal{S}_{uv} \subseteq \mathcal{C}$, defining the set of scales that render $u(x)$ and $v(x)$ two distributions of the same sum total. Clearly the majorization set $\mathcal{C}_{uv}$ is a subset of $\mathcal{S}_{uv}$. It turns out however that $\mathcal{S}_{uv}$ has all the relevant geometric properties of the majorization set that are of interest to us. Nonetheless, $\mathcal{S}_{uv}$ is considerably simpler to characterize than $\mathcal{C}_{uv}$.

3.1. The set of increasing pay scales and related subsets

Consider first the set $\mathcal{C}$. This set is a bounded and closed set in $\mathbb{R}^n$. It is also a convex set. Specifically, $\mathcal{C}$ is an $n-$dimensional simplex with $n+1$ vertices. The set $\mathcal{V}er(\mathcal{C})$ of vertices of $\mathcal{C}$ has the following elements:

$$\mathcal{V}er(\mathcal{C}) = x_{\max}. \left\{ \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} \vdots \\ 0 \\ 1 \\ 1 \end{pmatrix}, ..., \begin{pmatrix} 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \right\}, \tag{3.1}$$

with any $x \in \mathcal{C}$ being the convex combination of the $n+1$ elements of $\mathcal{V}er(\mathcal{C})$. When $n = 2$, $\mathcal{C}$ is a triangle with vertices $x_{\max}.(0,0)'$, $x_{\max}.(0,1)'$ and $x_{\max}.(1,1)'$. Set the value of $x_{\max}$ at unity. Figure 1 provides a sketch of the set of increasing scales $\mathcal{C}$ when $n = 3$, and $x_{\max} = 1$. $\mathcal{C}$ has four vertices: $e^1 = (0,0,0)'$, $e^2 = (0,0,1)'$, $e^3 = (0,1,1)'$ and $e^4 = (1,1,1)'$. The set $\mathcal{C}$ has a pyramidal shape. The convex hull of a subset of vertices defines a *face* of the simplex $\mathcal{C}$, and the vertices (taken individually) are commonly referred to as the *zero-dimensional faces* (Ziegler, 1995; ch. 1). The convex hull of any three vertices spans a triangle, while the convex hull of the four vertices spans the entire set $\mathcal{C}$ in three-dimensional space. Note furthermore that every subset of $\mathcal{C}$ (such as the majorization set $\mathcal{C}_{uv}$) is by definition also a bounded set in $\mathbb{R}^n$.

Let u and v be two elements of $\mathcal{D}$. Clearly, a necessary condition for $u(x)$ to be more egalitarian than $v(x)$ is that $u(x)$ and $v(x)$ are two distributions of a given sum total. Accordingly, define the set

$$\mathcal{S}_{uv} := \left\{ x \in \mathcal{C} : \sum_{j=1}^{l+1} u_j(x) = \sum_{j=1}^{l+1} v_j(x) \right\}. \tag{3.2}$$

9

This set is a subset of the null space

$$\mathcal{N}ull(u - v) := \{x \in \mathbb{R}^n : \delta x = 0\}, \tag{3.3}$$

that is, the set of vectors x that equate the sum totals of $u(x)$ and $v(x)$. Geometrically, $\mathcal{N}ull(u - v)$ is a hyperplane that is generated by $n - 1$ linearly independent vectors in $\mathbb{R}^n$. The set $\mathcal{S}_{uv}$ however is the intersection of $\mathcal{N}ull(u - v)$, and the set of increasing pay scales $\mathcal{C}$. As such, $\mathcal{S}_{uv}$ is *at most* generated by $n - 1$ linearly independent vectors, and we write the inequality:

$$\dim \mathcal{S}_{uv} \leq \dim \mathcal{N}ull(u - v) = n - 1 \tag{3.4}$$

Because the majorization set is a subset of $\mathcal{S}_{uv}$, we may define $\mathcal{C}_{uv}$ as follows:

$$\mathcal{C}_{uv} = \{x \in \mathcal{S}_{uv} : u(x) \prec v(x); \qquad u, v \in \mathcal{D}\} \tag{3.5}$$

Also, because we do not fix the scale x defining the distributions u and v, there remains the possibility that some increasing scales order u to be more egalitarian than v, while other scales reverse the majorization ordering. Thus, in what follows we shall also define the set of scales

$$\mathcal{C}_{vu} := \{x \in \mathcal{S}_{uv} : v(x) \prec u(x); \qquad u, v \in \mathcal{D}\}. \tag{3.6}$$

We shall call $\mathcal{C}_{vu}$ the dual majorization set.

Finally, there remains the possibility that some scale $x \in \mathcal{S}_{uv}$ does not order u and v. This is the set $\mathcal{C}_{u\|v} := \mathcal{S}_{uv} \backslash (\mathcal{C}_{uv} \cup \mathcal{C}_{vu})$, and we have that $\mathcal{S}_{uv} = \mathcal{C}_{uv} \cup \mathcal{C}_{vu} \cup \mathcal{C}_{u\|v}$.

3.2. Characterization of the majorization set

The approach we shall follow here will be to characterize the majorization set in relation to its vertices. This is a well-treaded path in the mathematical sciences. For instance, the set of doubly stochastic matrices is commonly referred to as the Birkhoff polytope, and the Birkhoff-von Newman Theorem states that the vertices of the Birkhoff polytope are the set of permutation matrices (Barvinok, 2002; pp. 56-57). Consider the context of majorization with continuous data, where u and v are vectors and P is a doubly stochastic matrix. The set of doubly stochastic matrices P such that $u = Pv$ is commonly referred to as the majorization polytope. Brualdi (1984) provides a characterization of this set in

relation to its vertices. Dahl (1999) similarly extends the results of Brualdi (1984) in the context of matrix majorization.

We begin by stating a lemma pertaining to the properties of $\mathcal{S}_{uv}$. The proof of this result is simple and helps to clarify the geometric properties of the majorization set.

Lemma 3.1 *Consider two distributions $u, v \in \mathcal{D}$. Then the set of pay scales $\mathcal{S}_{uv}$, such that u and v are two distributions of a given sum total, is a non-empty convex polytope.*

Proof Let $\delta = \Delta(u, v)$ and construct the following matrix:

$$M_1 := \begin{pmatrix} -\delta \\ \delta \end{pmatrix} \tag{3.7}$$

Then, some pay scale x satisfies the condition $\delta x = 0$ if and only if $M_1 x \leq 0$. The restriction on the vector x to belong to the set of increasing pay scales may be conveniently expressed as $M_2 x \leq (\mathbf{0}_n \ x_{\max})'$, where the matrix $M_2 \in \mathbb{Z}^{(n+1) \times n}$ is defined as follows:

$$M_2 := \begin{bmatrix} -1 & 0 & 0 & \cdots & 0 \\ 1 & -1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -1 \\ 0 & & \cdots & 0 & 1 \end{bmatrix} \tag{3.8}$$

Next, construct the following $(n + 3) \times n$ matrix:

$$M := \begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \tag{3.9}$$

and define the vector $\alpha := (\mathbf{0}_{n+2} \ x_{\max})'$. Then it is clear that $\mathcal{S}_{uv} = \{x \in \mathbb{R}^n : Mx \leq \alpha\}$. Thus, $\mathcal{S}_{uv}$ is a polyhedral set. This set is a bounded, closed and convex subset of $\mathbb{R}^n$. It is defined as the solution to a finite system of linear inequalities. Accordingly, $\mathcal{S}_{uv}$ is a convex polytope.

Note finally that the set $\mathcal{S}_{uv}$ is generally non-empty. Consider a constant r^a, such that $x_{\max} \geq r^a \geq 0$. Then, the constant vector $x^a := (r^a, ..., r^a)'$ will always satisfy the system of inequalities $Mx \leq \alpha$, and hence x^a is always an element of $\mathcal{S}_{uv}$. □

In simple terms, the lemma reveals that the subset $\mathcal{S}_{uv}$ of the set of increasing pay scales is a solution to a system of $n + 3$ linear inequalities in the vector x. As

11

a result, $\mathcal{S}_{uv}$ is a convex set, taking the form of a polytope [5]. Let P denote the following $(l + n + 3) \times n$ matrix:

$$P := \begin{pmatrix} A \\ M \end{pmatrix} \tag{3.10}$$

The result below summarizes the key geometric properties of the sets $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$, as well as provides the founding block of all subsequent results obtained in the paper. In particular, it reveals that the majorization set is also a non-empty convex polytope:

Theorem 3.2 *Consider two distributions $u, v \in \mathcal{D}$, and let $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$ respectively denote the sets of pay scales that equate the sum totals of the two distributions, and the majorization set. Then:*

(i) The set $\mathcal{C}_{uv}$ is a non-empty convex polytope defined as the solution to a system of linear inequalities

$$Px \leq \begin{pmatrix} \mathbf{0}_{l+n+2} \\ x_{\max} \end{pmatrix}, \tag{3.11}$$

(ii) the origin $\mathbf{0}_n$ and the vector $\varepsilon := x_{\max} \iota_n$, where ι_n is an $n-$dimensional column vectors of ones, are both vertices of $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$,

(iii) every vertex of $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$ other than $\mathbf{0}_n$ is of the form $e^i = x_{\max} F^i$ where F^i is the cumulative distribution of some random variable defined over n states.

Before we take up each of these three statements in turn, we first mention some related results from the mathematical sciences. In the context of continuous data, Dahl (1998) provides a similar characterization of extreme points (and rays) when the set $\mathcal{C}_{uv}$ takes the form of a convex polyhedral cone. In the context of matrix majorization with continuous data (i.e. where u and v are matrices), Dahl (1999) similarly characterizes the vertices of the majorization polytope $\mathcal{C}_{uv}$ consisting of the row stochastic matrices P such that $u = Pv$.

Consider now statement (i) of Theorem 3.2: because $\mathcal{C}_{uv}$ is a non-empty closed and bounded convex set, Minkowski's theorem entails that this set may be characterized as the convex hull of its vertices (see Corollary 3.3 below). Next, consider (ii) and any scalar $r^a \in [0, x_{\max}]$. Because the constant pay scale $x^a = (r^a, ..., r^a)'$

---

[5] The result also arises naturally from a geometric perspective as the outcome of constructing $\mathcal{S}_{uv}$ as the intersection of a hyperplane with a simplex (see Ziegler 1995, ch. 1 for a discussion).

is always an element of both $\mathcal{S}_{uv}$ and the majorization set, and can be expressed as an appropriate convex combination of $\mathbf{0}_n$ and the vector ε, it is clear that the latter two vectors are both vertices of the sets $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$. Finally, (iii) reveals the structure of every vertex of these two sets (other than the origin), as being the product of $x_{\max}$ and the cumulative distribution of some random variable defined over n states. This property is inherited from the ordering $x_1 \leq \cdots \leq x_n$ of the elements of the vector x. In other contexts where the elements of the policy vector are not naturally ordered, this last result (iii) need no longer hold [6].

Geometrically, vertices are boundary points of a given set; hence vertices may be also viewed as the boundaries of redistribution policies. That is, we may view the vertices of $\mathcal{C}_{uv}$ as the limits of pay policies that make the Lorenz curve of $u(x)$ more egalitarian than that of $v(x)$. Because there are $l + n + 3$ inequalities defining the majorization set (i.e. the number of rows of the matrix P of (3.10)), $\mathcal{C}_{uv}$ will possess a finite number of such vertices. Specifically:

Corollary 3.3 *The set of vertices $\{\eta^1, ..., \eta^k\} = \mathcal{V}er(\mathcal{C}_{uv})$ provides a minimal representation of the majorization set $\mathcal{C}_{uv}$, in the sense that:*

(i) $\mathcal{C}_{uv}$ is the convex hull of $\{\eta^1, ..., \eta^k\}$, and

(ii) each $x \in \mathcal{C}_{uv}$ is a convex combination of a family of at most n affinely independent points from $\{\eta^1, ..., \eta^k\}$.

The description of $\mathcal{V}er(\mathcal{C}_{uv})$ as providing a *minimal representation* of the set $\mathcal{C}_{uv}$ is taken from Lay (1982, pp. 116-117); in words it signifies that knowledge of the set of k points defining $\mathcal{V}er(\mathcal{C}_{uv})$ is all that we require to define all the elements of the majorization set $\mathcal{C}_{uv}$.

We next consider a simple example that illustrates Theorem 3.2 and its corollary 3.3.

Example 2 Let there be $n = 3$ states, $u = (x_2, x_2, x_1)$ and $v = (x_3, x_1, x_1)$. Here $\mathcal{S}_{uv}$ is the set of increasing scales such that for each $x \in \mathcal{S}_{uv}$ there holds $\sum_{j=1}^{3} u_l(x) = \sum_{j=1}^{3} v_l(x)$. We first construct the majorization matrix A as well as the

---

[6] We return to this point in Section 4.5 where we examine the majorization set in relation to a policy vector consisting of a universal benefit and a uniform tax rate, that define a linear income tax schedule.

matrices M_1 and M_2 of Theorem 3.2 to obtain here:

$$A = \begin{pmatrix} 0 & 1 & -1 \\ -1 & 2 & -1 \end{pmatrix}$$

$$M_1 = \begin{pmatrix} -1 & 2 & -1 \\ 1 & -2 & 1 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

Stacking vertically A, M_1 and M_2 allows us to construct the matrix P of Theorem 3.2. On such basis, we easily find that $\mathcal{S}_{uv} = \{x \in \mathcal{C} : x_3 = 2x_2 - x_1\}$. We can also obtain the vertices of the set $\mathcal{C}_{uv}$ [7]. These are $\eta^1 = (0\ 0\ 0)'$, $\eta^2 = x_{\max}(0\ 1/2\ 1)'$ and $\eta^3 = x_{\max}(1\ 1\ 1)'$.

In Figure 2, we draw once again the set $\mathcal{C}$ of Figure 1, and we represent the three vertices η^1, η^2 and η^3 where we fix $x_{\max}$ at the value 1. In accordance with Corollary 3.3, these three vertices span the set $\mathcal{C}_{uv}$, here a triangle, in the sense that $\mathcal{C}_{uv}$ is the convex hull of these three points (the shaded area in Figure 2.) With the exception of the origin (the vertex η^1) the vertices η^2 and η^3 are of the form of cumulative distribution functions for a random variable defined over $n = 3$ states.

Any $x \in \mathcal{C}_{uv}$ is expressible in the form of a convex combination $\sum_{k=1}^{3} \lambda_k \eta^k$ of the three vertices. We obtain for instance the following two elements of the majorization set $\mathcal{C}_{uv}$: $x^a = x_{\max}(0.50\ 0.75\ 1.00)' = 0.5(\eta^2 + \eta^3)$ and $x^b = x_{\max}(0.00\ 0.25\ 0.50)' = 0.5(\eta^1 + \eta^2)$. It may be verified that $u(x^a) \prec v(x^a)$ and also that $u(x^b) \prec v(x^b)$. $\bigcirc$

3.3. Separation result

Theorem 3.2 informs us that the set of pay scales x that render $u(x)$ more egalitarian than $v(x)$ is a convex set. Applying the same logic, but defining a majorization matrix $B := -A$, any scale $y \in \mathcal{C}$ that renders v more egalitarian than u satisfies the set of inequalities $By \leq \mathbf{0}_l$. For this reason it follows also that $\mathcal{C}_{vu}$ is a convex polytope, and accordingly a convex set.

---

[7]A systematic method for deriving the vertices of the majorization set is discussed in the appendix proof of Theorem 3.2.

For this reason, it may be useful to conclude this section of the paper with a result that commonly arises in the analysis of convex sets, namely a *separation* result. Let $\mathcal{A}$ and $\mathcal{B}$ denote two subsets of the set of increasing pay scales $\mathcal{C}$. If there exists a linear functional $\gamma : \mathcal{C} \to \mathbb{R}$ (where γ is not identically equal to zero) and a number $r \in \mathbb{R}$ such that $\gamma(x) \leq r$ for all $x \in \mathcal{A}$ and $\gamma(x) \geq r$ for all $x \in \mathcal{B}$, we shall say that the two sets $\mathcal{A}$ and $\mathcal{B}$ are separated by a hyperplane in $\mathcal{C}$.

To state this *separation* result, we need to draw an important distinction between increasing pay scales x that belong to the boundary of $\mathcal{C}$ from those that belong to the interior of this set. Accordingly, we first define the interior of $\mathcal{C}$ as the following subset of $\mathbb{R}^n$:

$$\mathcal{C}^o = \{x \in \mathcal{C} : 0 < x_1 < x_2 < \cdots < x_n < x_{\max}\} \qquad (3.12)$$

The boundary of $\mathcal{C}$, denoted $\partial\mathcal{C}$, is then the set $\mathcal{C}\backslash\mathcal{C}^o$. The result below summarizes the geometric structure of the two sets $\mathcal{C}_{uv}$ and $\mathcal{C}_{vu}$ in relation to one another, and as subsets of the set of increasing pay scales:

Proposition 3.4 *Let u and v denote two distinct distributions, and let $a = (a_1, ..., a_n)$ denote a row of the majorization matrix $A = \Phi(u, v)$, such that $a \neq \mathbf{0}$. Then the hyperplane*

$$\mathcal{H} = \{x \in \mathcal{C} : ax = 0\} \qquad (3.13)$$

separates the majorization sets $\mathcal{C}_{uv}$ and $\mathcal{C}_{vu}$. Furthermore, at most one of the two majorization sets has points in the interior of $\mathcal{C}$, while the other majorization set lies on the boundary of the set of increasing pay scales.

That is, the above result informs us that if there are two pay scales x and y such that $u(x) \prec v(x)$ and $v(y) \prec u(y)$, it must be that at least one of the two pay scales x and y is an element of the boundary of $\mathcal{C}$. That is, at most one of the two majorization sets has points $x \in \mathcal{C}^o$ with the property that the pay x_{i+1} associated with pay class $i + 1$ is strictly greater than x_i, while the other majorization set is contained in the boundary of the set of increasing pay scales. Finally, any non-zero row of the majorization matrix defines the coordinates of a hyperplane $\mathcal{H}$ that separates the two convex sets $\mathcal{C}_{uv}$ and $\mathcal{C}_{vu}$.

4. Applications of the majorization set

Below we shall consider various applications of the framework we have developed in Section 3, together with its central results, Theorem 3.2 and Proposition 3.4.

4.1. Lorenz-preserving transformations of the pay scale

Because $\mathcal{C}_{uv}$ is a convex set, any function on $\mathcal{C}$ that preserves the Lorenz ordering of u and v must, and can only, be a function $f : \mathcal{C}_{uv} \to \mathcal{C}_{uv}$. Specifically, for some $x \in \mathcal{C}_{uv}$, we can view a new pay scale $y \in \mathcal{C}_{uv}$ as a convex combination of x and some other element $z \in \mathcal{C}_{uv}$ such that $f(x; \mu, z) = \{\mu \in [0,1] , \ z \in \mathcal{C}_{uv} \ : \ y \ = \ \mu x \ + \ (1-\mu) z\}$. Note that because z and x are themselves convex combinations of the vertices of $\mathcal{C}_{uv}$, we have at once:

Corollary 4.1 *Let u and v be two distinct distributions, and let x be an element of $\mathcal{C}_{uv}$ such that $u(x) \prec v(x)$. Then a function $f : \mathcal{C} \to \mathcal{C}$ maps x to a new scale y such that $u(y)$ is more egalitarian than $v(y)$ if and only if y is a convex combination of x and the vertices of $\mathcal{C}_{uv}$.*

As such for instance, if $z = x_{\max}(1, ..., 1)'$, then we have that $y = \mu x + (1-\mu) z$ is also an element of $\mathcal{C}_{uv}$. Specifically, for $i = 1, ..., n-1$, we have that $y_i/y_{i+1} \geqslant x_i/x_{i+1}$. This sort of transformation has been discussed by Moyes (1994) in the income inequality literature: y is said to dominate x in relative differences, and in turn here the Lorenz curves of $u(y)$ and $v(y)$ will be ordered in the same fashion as $u(x)$ and $v(x)$.

4.2. Inequality between groups

If a distribution u exhibits a low level of inequality and v has a high level of inequality under an initial scale x°, a new scale x^1 may bring the Lorenz curves of $u(x^1)$ and $v(x^1)$ closer by making $u(x^1)$ and $v(x^1)$ two distributions with a high level of inequality where $u(x^1) \prec v(x^1)$. Conversely, a new scale x^1 may entail two Lorenz curves $u(x^1)$ and $v(x^1)$ close to the equality line with $u(x^1) \prec v(x^1)$. The scope for reducing between groups inequality in the context of discrete data may be examined formally by studying the null space of the majorization matrix A. The null space of the majorization matrix is the set of pay scales that equate the l partial sums (2.3) to zero. Some vectors in the null space of the majorization matrix may belong to the set of increasing pay scales $\mathcal{C}$ while others may not. It is clear that it is the former set, namely $\mathcal{N}ull(A) \cap \mathcal{C}$ that is of interest to us.

The vertices of $\mathcal{C}$ all lie on the boundary of $\mathcal{C}$. It is also the case that those elements of the null space of A that belong to $\mathcal{C}$ must lie on the boundary of the set of increasing scales:

Proposition 4.2 *Consider two distributions $u, v \in \mathcal{D}$, and let the majorization matrix A denote the image of (u, v) via the map $\Phi(., .)$. Consider some pay scale*

$x \in \mathcal{C}^o$. Then $x \in \mathcal{N}ull(A)$ if and only if $A = \mathbf{0}$.

In simple terms, a pay scale that equalizes the Lorenz curves of u and v will typically have at least two identical elements if u and v are distinct distributions. Put differently, if $u \neq v$, the Lorenz curves pertaining to these two distributions may be brought to equality using some increasing pay scale. The scope for such policies is informed by the rank of the matrix A: if $\mathrm{rank}(A) = n - 1$, then essentially the only scale that equates the Lorenz curves of u and v is of the form $x^a = (r, ..., r)'$ where $0 \leq r \leq x_{\max}$. If $\mathrm{rank}(A) = n - k$, then the null space of A has k linearly independent n-dimensional vectors, of which at least one must belong to $\mathcal{C}$ [8]. In other terms, the lower the rank of the majorization matrix the more flexibility there is for the policy maker to reduce between groups inequality:

Corollary 4.3 *Consider two distributions $u, v \in \mathcal{D}$, and let the $A = \Phi(u, v)$ denote the majorization matrix. Furthermore, let $\mathcal{N}ull(A) \cap \mathcal{C}$ denote the subset of pay scales of the null space of A that also belong to the set of increasing pay scales $\mathcal{C}$, and let $r \in [0, x_{\max}]$ be a scalar. Then,*
(i) If $\dim(\mathcal{N}ull(A) \cap \mathcal{C}) = 1$, between groups inequality, as measured by the distance between Lorenz curves, is only zero for uniform pay scales $x^a = (r, ..., r)'$. At $x^a = (r, ..., r)'$ between groups inequality and within groups inequality vanish simultaneously.
(ii) If $\dim(\mathcal{N}ull(A) \cap \mathcal{C}) > 1$, between groups inequality is zero for pay scales other than $x^a = (r, ..., r)'$. For such pay scales other than x^a, between groups inequality may vanish while within groups inequality persists.

Stated differently, when the null space $\mathcal{N}ull(A) \cap \mathcal{C}$ is of dimension one, policies that would equate Lorenz curves do so by confounding both Lorenz curves with the equality line. On the other hand, when the null space is of higher dimension, both Lorenz curves may be confounded while remaining distinct from the equality line.

4.3. Lorenz curve reversals

Having spent some time discussing the scope for equating Lorenz curves, it is natural to take this discussion one step further, in order to inquire about the

---

[8]Returning for instance to Example 2, we may observe that $\mathcal{N}ull(A) \cap \mathcal{C}$ is the set $\{x \in \mathcal{C} : x_3 = x_2 = x_1\}$. This set lies clearly on the boundary of $\mathcal{C}$. Since $\mathrm{rank}(A) = 3 - 1$, the null space is entirely characterized by vectors of the form $x^a = (r, r, r)'$. See Example 1 for a majorization matrix with a null space of higher dimension.

scope for Lorenz curve reversals. It turns out in fact that the result of Theorem 3.2 can be used to shed light on an important problem that was first discussed in the literature on the measurement of health inequality (Allison and Foster, 2004; Zheng 2011). It was recognized by these authors that when the scale defining the economic status of a person is undetermined, two different increasing scales x and y may reverse the Lorenz ordering of two distributions.

The framework developed in this paper is well suited to revisit this problem: specifically, starting from a scale $x^\circ \in \mathcal{C}_{uv}$ such that $u(x^\circ) \prec v(x^\circ)$, any new scale $x^1 \in \mathcal{C}$ will preserve the Lorenz ordering of $u(x^\circ)$ and $v(x^\circ)$ if and only if x^1 is a convex combination of the vertices of $\mathcal{C}_{uv}$. Conversely for a scale $y \in \mathcal{C}$ to reverse the ordering of u and v, such that there results $u(y) \succ v(y)$ or $u(y) \parallel v(y)$, a necessary condition is that $(\mathcal{S}_{uv} \setminus \mathcal{C}_{uv})$ is a non-empty set; in other terms, y must be selected from a set of scales that do not allow u to be more egalitarian than v.

The above reasoning can be taken one step further, to characterize conditions such that the reversal of Lorenz curves never occurs within the set of scales that make u and v two distributions of a given sum total. As $\mathcal{C}_{uv}$ is a subset of $\mathcal{S}_{uv}$, the condition sought after is that $\mathcal{S}_{uv}$ and $\mathcal{C}_{uv}$ are equal sets: when $\mathcal{S}_{uv} = \mathcal{C}_{uv}$ there results that $\mathcal{S}_{uv} \setminus \mathcal{C}_{uv}$ is an empty set. Fortunately, equality of the sets is easily verified with the help of the vertices of $\mathcal{S}_{uv}$:

Proposition 4.4 *Let u and v be two distributions in $\mathcal{D}$, and let $A = \Phi(u,v)$ denote the majorization matrix. Then, $u(x) \prec v(x)$ for all scales $x \in S_{uv}$, equivalently $\mathcal{C}_{uv} = \mathcal{S}_{uv}$, if and only if for every vertex ξ^h of the set $\mathcal{S}_{uv}$ there holds the inequality*

$$A\xi^h \leq \mathbf{0}_l, \qquad h = 1, ..., k \tag{4.1}$$

Stated differently, when for every vertex ξ^h of $\mathcal{S}_{uv}$, the greatest element of the vector $A\xi^h$ is non-positive, the Lorenz curve of $u(x)$ will lie above that of $v(x)$, for any pay scale that renders $u(x)$ and $v(x)$ two distributions of a given sum total.

To illustrate the practical relevance of the above result, we continue with our examination of Example 1 of Section 2.

Example 1 (continued) Return to Example 1 where $n = 6$, and where $u = (x_5, x_5, x_3, x_2)'$ and $v = (x_6, x_4, x_4, x_1)'$. The resulting majorization matrix A and vector δ are given in (2.11) and (2.12). We are interested in characterizing vectors $x \in \mathcal{C}$ that allow us to order $u(x)$ and $v(x)$.

Referring to Proposition 4.4, there are two vertices of $\mathcal{S}_{uv}$, $\xi^1 = x_{\max}(0\ 0\ 0\ 0\ \frac{1}{2}\ 1)'$ and $\xi^2 = x_{\max}(0\ 0\ 0\ \frac{1}{2}\ 1\ 1)'$ such that $A\xi^1 = x_{\max}(-\frac{1}{2}\ 0\ 0)' \leq \mathbf{0}_3$ and $A\xi^2 =$

$x_{\max}(0 \ \frac{1}{2} \ 0)' \geq \mathbf{0}_3$. Proposition 4.4 then alerts us to the possibility that there may be two scales $x, y \in \mathcal{S}_{uv}$ such that $u(x) \prec v(x)$, while $v(y) \prec u(y)$ or $v(y) \parallel u(y)$. This is clearly so here, as $\xi^1 \in \mathcal{C}_{uv}$ and $\xi^2 \in \mathcal{C}_{vu}$. For instance, if we set $x_{\max} = 10$, it is easily verified that $x = (0.5 \ 2 \ 3 \ 4 \ 5 \ 6.5)' \in \mathcal{C}_{uv}$ while $y = (2 \ 2 \ 3 \ 4 \ 5 \ 5)' \in \mathcal{C}_{vu}$.

It is instructive to characterize the sets $\mathcal{C}_{uv}$ and $\mathcal{C}_{vu}$ of this example. Firstly, we have

$$\mathcal{C}_{uv} = \left\{ x \in \mathcal{C} : \delta x = 0 \text{ and } x_5 \leq \left(\frac{x_4 + x_6}{2} \right) \right\} \tag{4.2}$$

For the set of scales such that $v(x) \prec u(x)$ we obtain

$$\mathcal{C}_{vu} = \left\{ x \in \mathcal{C} : x_5 = x_6, \ x_1 = x_2 \text{ and } x_4 = \left(\frac{x_3 + x_5}{2} \right) \right\} \tag{4.3}$$

Note that from the separation result of Proposition 3.4, at most one of the two sets $\mathcal{C}_{uv}$ and $\mathcal{C}_{vu}$ can have points in the interior of $\mathcal{C}$. Since $x = (0.5 \ 2 \ 3 \ 4 \ 5 \ 6.5)'$ is an element of the set (4.2), it is the case that $\mathcal{C}_{uv}$ has a non-empty intersection with $\mathcal{C}^o$. However, because every scale in (4.3) must satisfy the conditions $x_5 = x_6$, and $x_1 = x_2$, it is clearly the case that $\mathcal{C}_{vu}$ belongs to the boundary of the set of increasing scales. $\bigcirc$

4.4. Incentives and the majorization set

Informational assumptions regarding (1) what employers can observe about their employees' characteristics and effort levels, and (2) what employees can observe about the results (profits, surplus etc.) of corporations will typically set limits to what pay scales can be implemented in practice. That incentive considerations set limits on the extent of redistribution that is feasible is a widely discussed and documented phenomenon (Roberts, 1984; Hindriks and Myles, 2013, ch. 18). Our purpose here is to discuss how the majorization set shrinks when it is restricted to certain incentive compatible pay structures.

When incentive constraints are linear (see below for an example), they may be generally expressed in the form $\{x \in \mathcal{C} : Hx \leq \mathbf{0}\}$. The subset of incentive compatible pay scales of the majorization set may then be written in the form

$$\mathcal{I}_{uv} := \{x \in \mathcal{C}_{uv} : Hx \leq \mathbf{0}\} \tag{4.4}$$

Some models of tournament theory for instance entail an increasing pay structure with *pay increments* rising with seniority (see Section 6.4 of Cahuc et al. 2014).

Such incentive compatible pay scales have the property that $x_2 - x_1 \leq x_3 - x_2 \leq \cdots \leq x_n - x_{n-1}$; and find considerable support in empirical investigations (Baker et al. 2014). The resulting property of increasing pay increments can easily be translated into the system of $n-2$ linear inequalities of the form $Hx \leq \mathbf{0}$ discussed above [9].

The insight provided by tournament theory is to enable us to relate the determinants of the pay increments to the contractual environment of the firm and to workers' risk preferences. Specifically, the pay increments may be shown to rise with the standard deviation of the unobserved component of workers' effort distribution in a given pay class, and also with the number of workers m_i who occupy this class i. Furthermore, when workers are risk averse (rather than risk neutral), such theories predict that pay increments should fall, rather than rise, with seniority.

4.5. Optimal taxation

An examination of the optimal tax literature is a useful setting for thinking about extensions to the framework developed here. Specifically, the context discussed below allows one to consider alternative applications of the majorization set, where for instance the set refers to policy parameters rather than ordered pay scales.

Consider then a simple case where there are $n = 2$ ordered states pertaining to productivity levels x_1 and x_2, and where $0 < x_1 < x_2 < x_{\max}$. Each individual has a time endowment of one unit so that x_1 and x_2 also denote the full incomes of the m_1 and m_2 workers of each type, and $\bar{x} = (m_1 x_1 + m_2 x_2)/(m_1 + m_2)$ is the average productivity level. A linear income tax maps the distribution of productivities (full incomes) to a distribution of consumption (i.e. net earnings), with two consumption levels $y_i = b + (1-\tau)\kappa x_i$ for $i = 1, 2$. The parameter $\kappa \in [0, 1]$ captures the distortionary effect of taxation and $1 - \kappa$ is the (constant) marginal propensity to spend on leisure (see for instance Mirrlees, 1971 and Atkinson, 1973). Here b is a lump sum universal benefit and τ is the uniform income tax rate. It is clear that x_1 and x_2 are fixed here, and that only the y vector is to be chosen (via choice of values for the universal benefit and tax rate). The particular choice of policy instruments b and τ will depend among other factors on (i) the level of inequality

---

[9]For instance, if there are say $n = 4$ seniority levels, the system of inequalities takes the form

$$\begin{pmatrix} -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

in the distribution of productivities (the difference between x_2 and x_1), (ii) the population shares of low and high productivity workers (m_1 and m_2) and (iii) the distortionary effect of taxation (how large $1 - \kappa$ is) [10].

Perhaps the most interesting insight this exercise can offer is the need to recognize that in some instances it may be more illuminating to express the inequalities directly in terms of the policy parameters that define the relevant set of pay scales such that the distribution of post-tax earnings (consumption) is more egalitarian than the distribution of productivities (full incomes). In doing so, however, we turn our attention to expressing the $m_1 + m_2$ partial sums $Ay \leq \mathbf{0}$ in a more detailed form, replacing y by functions of κ, x_1, x_2 m_1 and m_2 (the data of the redistribution problem) as well as b and τ (the policy instruments). We come also to accept that the sum total of the distribution of post-tax earnings is lower than that of full incomes when $\kappa < 1$. This leads us to choose b and τ instead to equate the mean post-tax earnings level and $\kappa \overline{x}$, that is a share of the mean, rather than the mean, of the distribution of productivities. In turn, this entails writing the partial sums as a product of the majorization matrix and two other matrices:

$$
\begin{pmatrix}
0 & m_2 & 0 & -m_2 \\
m_1 & m_2 & -m_1 & -m_2 \\
-m_1 & -m_2 & m_1 & m_2
\end{pmatrix}
\begin{pmatrix}
1 & \kappa x_1 & 0 \\
1 & \kappa x_2 & 0 \\
0 & 0 & \kappa x_1 \\
0 & 0 & \kappa x_2
\end{pmatrix}
\begin{pmatrix}
b \\
1 - \tau \\
1
\end{pmatrix}
\leq
\begin{pmatrix}
0 \\
0
\end{pmatrix}. \quad (4.5)
$$

Note here the simple form of the majorization matrix that arises when comparing a fixed distribution u under two different income vectors x (full income) and y (post-tax earnings). The above formulation has the advantage of making the majorization matrix appear clearly as an integer matrix that satisfies the properties of Lemma 2.1. But it is also the case that this formulation now requires the introduction of an additional matrix (the middle matrix) parametrized by a subset of the data of the redistribution problem (the distortionary effect of taxation and the vector of productivities).

The constraints on the policy parameters are that (1) b equals $\tau \kappa \overline{x}$ and (2) that τ lies between 0 and 1. Note that the first of these conditions is implied by the second and third rows of (4.5). Upon substituting $\tau \kappa \overline{x}$ for b, the inequality $\tau \geq 0$ is implied by the first inequality of (4.5). Finally, the inequality $\tau \leq 1$

---

[10]There are additional important considerations needed in order to select the Pareto efficient level of b and τ. These include redistributive objectives, viz the parameterization of the social welfare function embodying society's preferences, and the share of total income that is committed to subsistence expenditure. See Deaton (1983) and Salanié (2003).

is a liability constraint: it precludes tax authorities from levying more tax than workers' earnings. The resulting majorization set in this context becomes the set of ordered pairs (b, τ) that lie within the line segment with vertices $(0, 0)$ and $(\kappa \bar{x}, 1)$. The majorization set expands when the average productivity level increases, and when the distortionary effect of taxation diminishes. Note finally that because (b, τ) is not an ordered vector in the way pay scales are, the first statement of Theorem 3.2 holds, but (ii) and (iii) of the same theorem no longer apply in the context of the linear income tax.

5. Conclusions

This paper has examined the problem of comparing distributions of a variable defined over a finite number of pay states, from the perspective of the Lorenz ordering. The particular problem we have chosen to address was that of characterizing the set of pay scales that allow for a Lorenz ranking of two distributions. This particular perspective has enabled us to address related questions of interest, such as the scope for reducing between groups inequality, as measured by the distance between Lorenz curves, and the extent to which Lorenz curve reversals may occur under two different pay scales.

We have shown that the problem of characterizing the set of pay scales that allow for a comparison of Lorenz curves in presence of a finite number of economic states has for solution a particular type of convex set known as a polytope. That is, the majorization set is a convex set, defined as the convex hull of a finite number of points, known as the vertices of the set. The significance of the above result is the following: even though the set of pay scales that allow for an ordering of Lorenz curves has an infinite number of points, this set has a minimal representation in terms of a finite number of vertices, and every pay scale in the majorization set arises as a convex combination of a subset of these vertices. The vertices are then jointly sufficient statistics for generating the majorization set.

Central to this result and its applications, is a *majorization matrix* that we have introduced in the paper, that allows us to express the partial sums (inequalities) defining Lorenz dominance. Our second result was to apply a well-known result in convex analysis, the separation theorem, to show that in general the majorization set and its dual, the majorization set that reverses the ordering of Lorenz curves, may be separated using a given number of hyperplanes, the coordinates of which are the non-zero rows of the majorization matrix.

Our final application in Section 4 gave a brief overview of the majorization set

in the context of the design of a linear income tax scheme consisting of a lump sum transfer and a unique marginal tax rate. There are clearly many other areas of application of the majorization set that may form the basis of future research, such as for instance its analysis in the context of differential commodity taxation and social insurance schemes more broadly.

6. Appendix

This appendix gathers the proofs of the results stated in Sections 3 and 4, together with the proof of Lemma 2.1. We begin with a formal definition of the function Φ that maps $\mathcal{D} \times \mathcal{D}$ to majorization matrices.

For $1 \leq h \leq l$, define the zero-one matrix $\Gamma_h = [I_h \ \mathbf{0}]$ where I_h is the $h-$dimensional identity matrix and where $\mathbf{0}$ is a $h \times (l + 1 - h)$ matrix of zeroes. Then Γ_h is a $h \times (l+1)$ *selection matrix* such that $\Gamma_h u = (u_1, ..., u_h)'$ for all $u \in \mathcal{D}$. In words, $\Gamma_h u$ is the vector comprising the h richest individuals from the distribution u. Observe also that $\Gamma_1 u = u_1$ and $\Gamma_l u = (u_1, ..., u_l)'$.

We next define the (row) vector function $\phi_j : \mathcal{D} \times \mathcal{D} \longrightarrow \mathbb{Z}^{1 \times n}$ as follows:

$$\phi_j(u, v) := \text{count}(\Gamma_j u) - \text{count}(\Gamma_j v) \qquad j = 1, ..., l \qquad (6.1)$$

We then construct the $l \times n$ matrix $\Phi(.,.)$ with jth row given by the function $\phi_j(u, v)$. Thus, we define for all $u, v \in \mathcal{D}$ the matrix function

$$\Phi(u, v) := \begin{pmatrix} \phi_1(u, v) \\ \vdots \\ \phi_l(u, v) \end{pmatrix} \qquad (6.2)$$

such that $\Phi(u, v)$ maps $\mathcal{D} \times \mathcal{D}$ into the set of $l \times n$ integer matrices $\mathbb{Z}^{l \times n}$.

We also define for $u, v \in \mathcal{D}$,

$$\Delta(u, v) := \text{count}(u) - \text{count}(v), \qquad (6.3)$$

23

where Δ maps $\mathcal{D} \times \mathcal{D}$ into $\mathbb{Z}^{1 \times n}$.

For the proof of Lemma 2.1, we shall require to define $\mathcal{E}$ as the set of $l \times n$ zero-one matrices such that each row $e_j = (e_{j1}, ..., e_{jn})$ of a matrix $E \in \mathcal{E}$ has exactly one unique element equal to 1 and all other elements equal to 0 :

$$\mathcal{E} = \{E \in \mathbb{Z}^{l \times n} : e_{ji} \in \{0, 1\}, \sum_{i=1}^{n} e_{ji} = 1 \qquad \text{for all } j = 1, ..., l\} \qquad (6.4)$$

Recall that $\Gamma_l = [I_l \ \mathbf{0}_l]$ is a $l \times (l+1)$ selection matrix, such that for any $u \in \mathcal{D}$ and any $x \in \mathcal{C}$, $\Gamma_l u(x) = (u_1(x), ..., u_l(x))'$. It follows that $\Gamma_l u(x)$ may be expressed as $\Gamma_l u(x) = E_u x$, where $E_u \in \mathcal{E}$ is a zero-one matrix.

Let $T \in \mathbb{Z}^{l \times l}$ denote a lower triangular matrix with element $t_{ij} = 1$ for all $i \geq j$ and $t_{ij} = 0$ for all $i < j$. Then, for any $a \in \mathbb{R}^l$, Ta is the vector which cumulates the elements of a, and has for jth component $a_1 + ... + a_j$. Accordingly,

$$TE_u = \begin{pmatrix} count\,(\Gamma_1(u)) \\ \vdots \\ count\,(\Gamma_l(u)) \end{pmatrix} \qquad (6.5)$$

Likewise define $E_v \in \mathcal{E}$ such that $(v_1(x), ..., v_l(x))' = E_v x$. Defining the $l \times n$ integer matrix

$$G := E_u - E_v \qquad (6.6)$$

it may readily be observed that $A = \Phi(u, v)$ may equivalently be constructed as

$$A = TG, \qquad (6.7)$$

i.e. the product of two integer matrices.

Observe also that since T is a full rank square matrix, $\mathcal{N}ull(T) = \mathbf{0}_l$. Accordingly, $E_u - E_v$ and $A = T(E_u - E_v)$ are matrices of identical rank. It follows then that for any $x \in \mathcal{C}$, we have $x \in \mathcal{N}ull(A)$ if and only if $x \in \mathcal{N}ull(E_u - E_v)$.

Note furthermore that for the first row of A and G we have the equality $(a_{11}, ..., a_{1n}) = (g_{11}, ..., g_{1n})$ while,

$$\begin{aligned} (a_{21}, ..., a_{2n}) &= (g_{11} + g_{21}, ..., g_{1n} + g_{2n}) \\ &= (a_{11} + ... + a_{1n}) + (g_{21} + ... + g_{2n}) \end{aligned} \qquad (6.8)$$

By repeated application of the above identities, we obtain for all $j = 2, ..., l$:

$$(a_{j1}, ..., a_{jn}) = (a_{j-1,1} + ... + a_{j-1,n}) + (g_{j1} + ... + g_{jn}) \qquad (6.9)$$

24

We are now ready to proceed to the proof of Lemma 5.1.

Proof of Lemma 2.1 $(PA1)$ Consider the jth row of G. If $u_j = v_j$, then the jth row of G is equal to zero. If on the other hand $u_j = x_h$ and $v_j = x_k$ with $u_j \neq v_j$, then $g_{jh} = 1$, $g_{jk} = -1$, and $g_{jm} = 0$ for all $m \neq h, k$. In both cases, we find that $\sum_{i=1}^{n} g_{ji} = 0$; accordingly, $A\iota_n = TG\iota_n = \mathbf{0}_l$.

$(PA2)$ Since $A = TG$ where T is a full rank matrix, $A = \mathbf{0}_{l,n}$ entails $G = \mathbf{0}_{l,n}$. In turn this entails that $u = v$. Conversely, if $u = v$ then $G = \mathbf{0}_{l,n}$ and A is a matrix of zeroes.

$(PA3)$ From $(PA1)$ we have that $\iota_n \in \mathcal{N}ull(A)$, so that $\dim \mathcal{N}ull(A) \geq 1$. Now assume that $u_1 = x_{n-1}$, $v_1 = x_n$ and for all $j = 2, ..., l$ we have $v_j = u_j$. Then $a_{1,n-1} = 1$, $a_{1n} = -1$, and all other elements of A are equal to zero. Since the resulting matrix has rank 1, and n columns, its null space is of dimension $n - 1$. Therefore it is the case that if $u \neq v$ then $\dim \mathcal{N}ull(A) \leq n - 1$.

$(PA4)$ Let $B = \Phi(v, w)$ and $C = \Phi(u, w)$. Then, $C = T(E_u - E_w) = T[(E_u - E_v) + (E_v - E_w)]$. That is, $C = A + B$.

$(PA5)$ Consider two vectors u and v, such that $u_1 = \cdots = u_j = x_h$ and $v_1 = \cdots = v_j = x_k$. Let i be an element of the set of indices $\{1, ..., n\}$. If $i = h$ and $k \neq h$ then $a_{ji} = j$. If $i = k$ and $k \neq h$, then $a_{ji} = -j$. Note also that if $h = k$, then $a_{ji} = 0$. Any a_{ji} is thus bounded by the values $-j$ and j.

$(PA6)$ Since E_u, E_v are elements of $\mathcal{E}$, it follows from (6.9) that if $(a_{j1}, ..., a_{jn}) = (0, ..., 0)$ then $(a_{j+1,1}, ..., a_{j+1,n})$ can only have at most two non-zero elements, and that $a_{j+1,i} \in \{-1, 0, 1\}$.

$(PA7)$ The jth row of G is equal to zero if and only if $u_j = v_j$. Thus, from (6.9) it follows that $(a_{j-1,1}, ..., a_{j-1,n}) = (a_{j1}, ..., a_{jn})$ if and only if $u_j = v_j$. $\qquad \square$

Proof of Theorem 3.2 (i) Let $A = \Phi(u, v)$ denote the image of (u, v) via the map (6.2). Then, a vector $x \in \mathbb{R}^n$ satisfies the l first partial sums (2.3) defining the majorization relation $u(x) \prec v(x)$, if and only if $Ax \leq \mathbf{0}_l$. The $l + 1$ partial sums [2.3-2.4] can now be conveniently be expressed as follows:

$$\begin{pmatrix} A \\ M_1 \end{pmatrix} x \leq \mathbf{0}_{l+2}$$

where M_1 is the matrix (3.7) of Lemma 3.1. The restriction on the vector x to belong to the set of increasing scales is also expressed as $M_2 x \leq (\mathbf{0}_n \ x_{\max})'$, where the matrix M_2 is defined in (3.8). Thus defining $\beta := (\mathbf{0}_{l+n+2} \ x_{\max})'$, we have

25

that $\mathcal{C}_{uv} = \{x \in \mathbb{R}^n : Px \leq \beta\}$, where P is the matrix of (3.10). Thus, $\mathcal{C}_{uv}$ is a polyhedral set. This set is bounded, closed and convex. Also, because P has a finite number of rows, $\mathcal{C}_{uv}$ is a convex polytope.

Note furthermore that the same argument as in the proof of Lemma 3.1 can be used to show that $\mathcal{C}_{uv}$ is generally a non-empty set in $\mathbb{R}^n$, and that its dimension is at least equal to 1.

(ii) Turning to the vertices of $\mathcal{C}_{uv}$, consider first the vector $\varepsilon = x_{\max}\iota_n$. Since ε satisfies the system of inequalities (3.11) of Theorem 3.2, it follows that $\varepsilon \in \mathcal{C}_{uv}$. On the other hand, ε cannot be obtained as a convex combination of any elements from the set $\mathcal{C}_{uv}\backslash\{\varepsilon\}$. It follows that ε must be a vertex of $\mathcal{C}_{uv}$.

Next consider vertices other than $\varepsilon = x_{\max}\iota_n$. Let $\mathcal{J}_n := \{j_1, ..., j_n\}$ denote a set of n indices from the set $\{1, 2, ..., l+n+3\}$, and let $Q := \mathcal{J}_n(P)$ and $a := \mathcal{J}_n(\beta)$ denote respectively the n-dimensional square matrix and vector constructed using rows $j_1, ..., j_n$ of the matrix P and the vector β. From Theorem 2.4.2 of Barvinok (2002), every vertex of $\mathcal{C}_{uv}$ is the solution v to some system of linearly independent equations $Qv = a$, where $v \in \mathcal{C}_{uv}$. Consider first the case where a is a vector of zeroes (that is, $x_{\max} \notin a$; equivalently, the index $l + n + 3$ pertaining to the last row of the matrix P is not an element of the set of indices $\mathcal{J}_n$ used to construct the matrix Q and the vector a). Then, since v is a vertex of $\mathcal{C}_{uv}$, it follows that Q is a full rank matrix. Thus, the only solution to $Qv = 0$ is the trivial solution $v = 0$, and the origin is a vertex of $\mathcal{C}_{uv}$. Thus (ii) holds.

(iii) Consider next the case where $l + n + 3 \in \mathcal{J}_n$, so that $x_{\max} \in a$. Then, the solution to the system of linearly independent equations $Qv = a$ is $v = Q^{-1}a$, a vector of n linear combinations of the elements 0 and $x_{\max}$, where one such elements of v is equal to $x_{\max}$. Since $v \in \mathcal{C}_{uv}$, it follows that $0 \leq v_1 \leq ... \leq v_n = x_{\max}$. Thus, $v = x_{\max}F$, where F is the cumulative distribution of some random variable defined over n states.

Now consider the set $\mathcal{S}_{uv}$. From Lemma 3.1, $\mathcal{S}_{uv}$, as $\mathcal{C}_{uv}$, is a convex polytope contained within $\mathcal{C}$. The proof of ($ii - iii$) in relation to $\mathcal{S}_{uv}$ proceeds exactly as the above arguments in the context of $\mathcal{C}_{uv}$. $\square$

The next lemma states that if there is a pay vector x such that $u(x)$ is more egalitarian than $v(x)$, and this x is in the interior of $\mathcal{C}$, then any y such that $v(y)$ is more egalitarian than $u(y)$ must lie on the boundary of $\mathcal{C}$ (and vice versa).

Lemma A1 *Let $u, v \in \mathcal{D}$, with $u \neq v$. Assume there is a $x \in \mathcal{C}_{uv}\backslash\partial\mathcal{C}$. Then the set $\mathcal{C}_{vu}\backslash\partial\mathcal{C}$ is empty. Likewise if there is a $y \in \mathcal{C}_{vu}\backslash\partial\mathcal{C}$, then the set $\mathcal{C}_{uv}\backslash\partial\mathcal{C}$ is empty.*

Proof of Lemma A1 Let x be an element of the set $\mathcal{C}_{uv} \backslash \partial \mathcal{C}$ and assume, to the contrary, that there is a $x^+ \in \mathcal{C}_{vu} \backslash \partial \mathcal{C}$. Since by assumption $u \neq v$, it follows from $[PA2]$ that $A = \Phi(u, v)$ must be a non-zero matrix. Consider then the first non-zero row $a_j = (a_{j1}, ..., a_{jn})$ of this matrix. Then, from $[PA6]$, there are at most two indices $h, k \in \{1, ..., n\}$ such that $h < k$, $a_{jh} \neq 0$, $a_{jk} = -a_{jh}$, and, by assumption, $a_{jh} x_h + a_{jk} x_k \leq 0$. Since $u(x) \prec v(x)$, it also follows from $[PA6]$ that $a_{jh} = 1$ and $a_{jk} = -1$. From the assumption that $x \in \mathcal{C}^o$, it also follows that $0 < x_h < x_k$.

Now since $x^+ \in \mathcal{C}_{vu} \backslash \partial \mathcal{C}$, it is also the case that $a_{jh} x_h^+ + a_{jk} x_k^+ \geq 0$. But since $x^+ \in \mathcal{C}$, this inequality entails that $x_h^+ = x_k^+$. In turn, $x_h^+ = x_k^+$ entails that $x^+ \in \partial \mathcal{C}$, in contradiction with the assumption that $x^+ \in \mathcal{C}_{vu} \backslash \partial \mathcal{C}$. $\quad \square$

Proof of Proposition 3.4 Define the integer matrix $B := -A$. Then it follows that the dual majorization set may be defined as $\mathcal{C}_{vu} = \{x \in \mathcal{S}_{uv} : Bx \leq \mathbf{0}_l\}$. Accordingly, Theorem 3.2 entails that $\mathcal{C}_{vu}$ is a convex polytope. Furthermore, this set is non-empty as the constant pay scale $(r, ..., r)$, for $r \in [0, x_{\max}]$, belongs to this set.

Next consider a hyperplane $\mathcal{H} = \{x \in \mathcal{C} : ax = 0\}$ where $a \neq (0, ..., 0)$ is a non-zero row of the majorization matrix A. Since any element $x \in \mathcal{C}_{uv}$ satisfies the inequalities $Ax \leq \mathbf{0}_l$, while any $x \in \mathcal{C}_{vu}$ satisfies the inequalities $Ax \geq \mathbf{0}_l$, the hyperplane $\mathcal{H}$ separates the two majorization sets.

Assume that $\mathcal{C}_{uv}$ has a point $x \in \mathcal{C}^o$. Then, from Lemma A1, $\mathcal{C}_{vu} \backslash \partial \mathcal{C}$ is an empty set. It follows that $\mathcal{C}_{vu}$ has an empty intersection with the interior of $\mathcal{C}$. $\square$

Proof of Proposition 4.2 $(\Rightarrow)$ Let $Ax = \mathbf{0}_l$ where $x \in \mathcal{C}^o$. We proceed by induction on the rows $a_1, ..., a_l$ of A to establish that $A = \mathbf{0}$.

Consider first $a_1 = (a_{11}, ..., a_{1n})$. There are (at most) two indices $h, k \in \{1, ..., n\}$ such that $a_{1h} + a_{1k} = 0$, $a_{1h} x_h + a_{1k} x_k = 0$, $x_h \neq x_k$ and $a_{1m} = 0$ for all $m \neq h, k$. If $a_{1h} \neq 0$, then $a_{1k} = -a_{1h}$ and $a_{1h} x_h + a_{1k} x_k = 0$ entails $x_h = x_k$. But this case is ruled out since by assumption x is an interior point of $\mathcal{C}$. Hence we conclude that $a_{1h} = a_{1k} = 0$ and thus that $a_1 = (a_{11}, ..., a_{1n}) = (0, ..., 0)$.

Now assume that for row $j - 1$ of A we have $a_{j-1} = (0, ..., 0)$. Then, from $[PA6]$ of Lemma 2.1, in row j of A there are again at most two indices $h, k \in \{1, ..., n\}$ such that $a_{jh} + a_{jk} = 0$, $a_{jh} x_h + a_{jk} x_k = 0$, $x_h \neq x_k$ and $a_{jm} = 0$ for all $m \neq h, k$. We now invoke the same argument as in the case of a_1 to establish that $a_j = (0, ..., 0)$.

$(\Leftarrow)$ Let $A = \mathbf{0}$. Then $\mathcal{C} \subseteq \mathcal{N}ull(A)$ and hence $\mathcal{C}^o \subseteq \mathcal{N}ull(A)$. $\quad \square$

Proof of Proposition 4.4 It is clear that $\mathcal{C}_{uv} \subseteq \mathcal{S}_{uv}$. Hence we need to prove that $\mathcal{S}_{uv} \subseteq \mathcal{C}_{uv}$ if and only if for every vertex of the set $\mathcal{S}_{uv}$ there holds the

inequality (4.1).

Let $\mathcal{V}er(\mathcal{S}_{uv}) = \{\xi^1, ..., \xi^k\}$. Then, since from Lemma 3.1 $\mathcal{S}_{uv}$ is a convex polytope, Minkowski's theorem entails that every $x \in \mathcal{S}_{uv}$ can be expressed in the form

$$
\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \mu_1 \begin{pmatrix} \xi_1^1 \\ \vdots \\ \xi_n^1 \end{pmatrix} + \cdots + \mu_k \begin{pmatrix} \xi_1^k \\ \vdots \\ \xi_n^k \end{pmatrix}, \tag{6.10}
$$

$$
0 \leq \mu_1, ..., \mu_k \leq 1 \text{ and } \sum_{h=1}^{k} \mu_h = 1. \tag{6.11}
$$

Assume first that $\mathcal{S}_{uv} \subseteq \mathcal{C}_{uv}$. Then this entails from the definition of $\mathcal{C}_{uv}$ that every $x \in \mathcal{S}_{uv}$ satisfies the inequalities $Ax \leq \mathbf{0}_l$. That is, $A(\mu_1\xi^1 + \cdots + \mu_k\xi^k) \leq \mathbf{0}_l$ where $\mu_1, ..., \mu_k$ satisfy the conditions of (6.11). In particular, for $\mu_h = 1$ and $\mu_j = 0$ for all $j \neq h$, we must have $A\xi^h \leq \mathbf{0}_l$.

Conversely, assume that for each $h = 1, ..., k$ there holds $A\xi^h \leq \mathbf{0}_l$. Then, since each μ^h in (6.10) is non-negative, it follows that $A(\mu_1\xi^1 + \cdots + \mu_k\xi^k) \leq \mathbf{0}_l$. Thus, from (6.10) we have that $Ax \leq \mathbf{0}_l$ for each $x \in \mathcal{S}_{uv}$. Accordingly, we conclude that $\mathcal{S}_{uv} \subseteq \mathcal{C}_{uv}$. $\square$

7. References

Abul Naga, R. H.,Yalcin, T.: Inequality measurement for ordered response health data. *Journal of Health Economics* 27, 1614-1625 (2008).

Allison, R., A., Foster, J. : Measuring health inequalities using qualitative data. *Journal of Health Economics* 23, 505-524 (2004).

Ando, T. : Majorization and inequalities in matrix theory. *Linear Algebra and its Applications* 199, 17-67 (1994).

Andreoli, F., Zoli, C. : On the measurement of dissimilarity and related orders. ECINEQ working paper 247 (2013).

Apouey, B. : Measuring health polarization with self-assessed health data. *Health Economics* 16, 875-894 (2007).

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

Atkinson, A.B.: How progressive should income taxation be?" in Parkin M., Robert Nobay A. eds.: *Essays in Modern Economics.* Longman (1973).

Baker, G., Gibbs, M., Holmström, B. : The internal economics of the firm: evidence from personnel data. *Quarterly Journal of Economics* 109, 881-919 (1994).

Barvinok, A. : *A Course in Convexity.* American Mathematical Society (2002).

Bourguignon, F. : Decomposable income inequality measures. *Econometrica* 47, 901-920 (1979).

Brualdi, R.: The doubly stochastic matrices of vector majorization. *Linear Algebra and its Applications* 61, 141-154 (1984).

Cahuc, P., Carcillo, S., Zylberberg, A. : *Labor Economics.* MIT Press (2014).

Chakravarty, S., Zoli, C. : Stochastic dominance relations for integer variables. *Journal of Economic Theory* 147, 1331–1341 (2012).

Cowell, F., Flachaire, E. : Inequality with ordinal data. Manuscript (2012).

Dahl, G. : Majorization, polyhedra and statistical testing problems. *Linear Algebra and its Applications* 272, 205-225 (1998).

Dahl, G. : Majorization polytopes. *Linear Algebra and its Applications* 297, 157-175 (1999).

Deaton, A. : An explicit solution to an optimal tax problem. *Journal of Public Economics* 20, 333-346 (1983).

Department of Finance and Personnel: *Senior civil service pay.* http ://www.dfpni.gov.uk /scs_pay_scales_1_august_2014.pdf (2014).

Foster, J., Shneyerov, A. : A general class of additively decomposable inequality measures. *Economic Theory* 14, 89-111 (1999).

Gajdos, T., Weymark, J. : Multidimensional generalized gini indices. *Economic Theory* 26, 471-496 (2005).

Gravel, N., Moyes, P. : Ethically robust comparisons of bidimensional distributions with an ordinal attribute. *Journal of Economic Theory* 147, 1384-1426 (2012).

Gravel N., Magdalou, B., Moyes, P.: Ranking distributions of an ordinal attribute. Aix-Marseille School of Economics working paper 2014-50 (2014).

Hindriks, J., Myles, G.: *Intermediate Public Economics*, MIT Press (2013).

Kobus, M., Miłoś, P.: Inequality decomposition for population subgroups. *Journal of Health Economics* 31 15-21 (2012).

Lay, S.: *Convex Sets and their Applications.* John Wiley (1982).

Marshall, A., Olkin, I., Arnold, B.:*Inequalities: Theory of Majorization and its Applications.* Springer (2010).

Mirrlees, J. : An exploration in the theory of optimal taxation. *Review of Economic Studies* 38, 175-208 (1971).

Moyes, P. : Inequality reducing and inequality preserving transformations of incomes: symmetric and individualistic transformations. *Journal of Economic Theory* 63, 271-298 (1994).

Ok, E. : Fuzzy measurement of income inequality: some possibility results on the fuzzification of the Lorenz ordering. *Economic Theory* 7, 513-530 (1996).

del Río, C., Alonso-Villar, O.: New unit-consistent intermediate inequality indices. *Economic Theory* 42, 505-521 (2010).

Roberts, K. : The theoretical limits to redistribution. *Review of Economic Studies* 51, 177-195 (1984).

Salanié, B. : *The Economics of Taxation.* MIT Press (2003).

Savaglio, E.: Multidimensional inequality with variable population size. *Economic Theory* 28, 85-94 (2006).
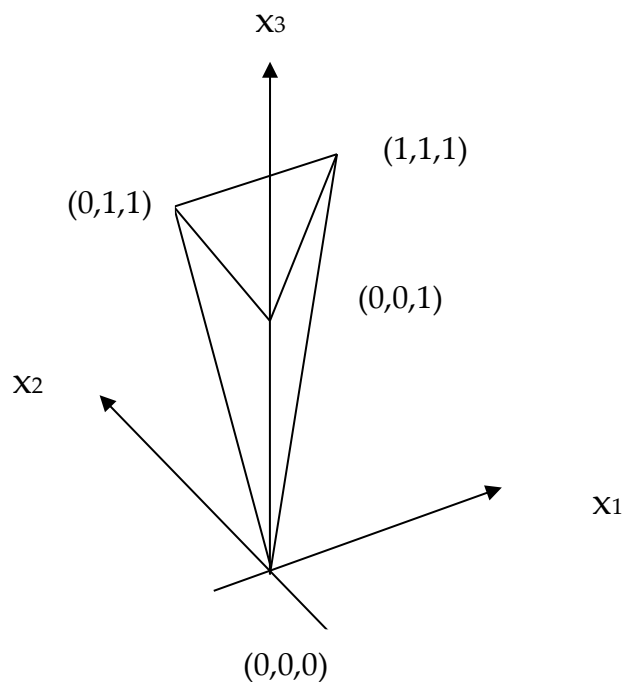
Shorrocks, A. : The class of additively decomposable inequality measures. *Econometrica* 48, 613-625 (1980).

Zheng, B. : A new approach to measure socioeconomic inequality in health", *Journal of Economic Inequality* 9, 555-577 (2011).

Ziegler, G. : *Lectures on Polytopes.* Springer, (1995).

Figure 1

The set C and its four vertices when $n = 3$ and $x_{max} = 1$

Figure 2

The majorization set C_{uv} of Example 2 and its three vertices when $n = 3$ and $x_{max} = 1$