Available online at www.sciencedirect.com
ScienceDirect
Journal of Economic Theory 144 (2009) 1358–1371
JOURNAL OF Economic Theory
www.elsevier.com/locate/jet

Note

A consistent multidimensional Pigou–Dalton transfer principle

Kristof Bosmans [a,b], Luc Lauwers [b], Erwin Ooghe [b,*]

[a] *Department of Economics, Maastricht University, Tongersestraat 53, 6211 LM Maastricht, The Netherlands*
[b] *Center for Economic Studies, Katholieke Universiteit Leuven, Naamsestraat 69, 3000 Leuven, Belgium*

Received 28 November 2006; final version received 26 November 2008; accepted 8 January 2009

Available online 12 February 2009

Abstract

The unidimensional Pigou–Dalton transfer principle demands that a regressive transfer in income—a transfer from worse-off (poor) to better-off (rich)—decreases social welfare. In a multidimensional setting the direct link between income (or any other attribute) and individual well-being is absent. We interpret the social welfare level of a distribution in which each individual has the same bundle as the individual well-being level. We define regressivity on the basis of this individual well-being ranking. In a setting with both transferable and non-transferable attributes, the imposition of the ensuing "consistent" Pigou–Dalton principle forces individual well-being to have a quasi-linear structure in the transferable attributes. Since we allow for transferable and non-transferable attributes, our result provides a normative underpinning for criteria in the distinct literatures of multidimensional inequality measurement (only transferable attributes) and of needs (one transferable and one non-transferable attribute).

JEL classification: D63

Keywords: Pigou–Dalton principle; Multidimensional inequality measurement; Majorization; Budget dominance; Needs; Equivalence scale

* Corresponding author. Fax: +32 (0) 16 32 67 96.
E-mail addresses: k.bosmans@algec.unimaas.nl (K. Bosmans), luc.lauwers@econ.kuleuven.be (L. Lauwers), erwin.ooghe@econ.kuleuven.be (E. Ooghe).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2009.01.003

1. Introduction

In unidimensional inequality measurement, the Pigou–Dalton transfer principle states that a regressive transfer in income—this is a transfer from a poorer to a richer individual—increases income inequality and decreases social welfare. Income, however, is usually considered a poor indicator of individual well-being and should therefore be augmented by other attributes (e.g., wealth, health, and literacy). If the focus is on individual well-being rather than income, then the setting expands to a multidimensional framework and the Pigou–Dalton principle requires modification.

We consider such a multidimensional framework. A social state is a distribution of attribute bundles over the individuals in society, one bundle for each individual. These attribute bundles incorporate all the relevant information, i.e., individuals are treated as identical except for their attribute bundles. Two difficulties arise in developing a multidimensional transfer principle.

First, the idea of a transfer that preserves the total amount of the attribute in society may not be meaningful or desirable for each attribute. We will therefore formulate a multidimensional version of the Pigou–Dalton principle in terms of transferable attributes only.

The *second* problem is less easily tackled: in the multidimensional setting there is in general no direct link between the level of some single attribute and the level of well-being. Consider an example with two attributes, income (transferable) and health (not transferable). Individual i has the attribute bundle (income, health) $= (5, 6)$ while j has the bundle $(9, 2)$. An income transfer from i to j is regressive in income. However, whether or not this income transfer is regressive also in terms of well-being depends upon the ranking of the bundles $(5, 6)$ and $(9, 2)$. In order to be able to label a transfer as regressive in terms of well-being, we need a ranking of the attribute bundles in terms of individual well-being. We appeal to the social ranking—assumed to be complete—to generate such a ranking at the individual level. More precisely, the ranking in terms of individual well-being coincides with the social ranking of distributions in which each individual has the same attribute bundle. In the above example, the bundle $(5, 6)$ yields a lower (higher) level of individual well-being than $(9, 2)$ if the social ranking considers the distribution $((5, 6), (5, 6), \dots, (5, 6))$ to be worse (better) than the distribution $((9, 2), (9, 2), \dots, (9, 2))$. As such, the social ranking induces an individual well-being ranking. Since individuals are considered identical except for their attribute bundles, the ranking in terms of well-being is the same for each individual.

We now formulate our multidimensional extension of the Pigou–Dalton principle: a social ranking satisfies the consistent Pigou–Dalton principle if a transfer considered to be regressive in terms of the induced individual well-being ranking, results in a worse distribution. Our main result states that—in the standard framework of additively separable social welfare functions—the imposition of the consistent Pigou–Dalton principle forces individual well-being to have a quasi-linear structure in the transferable attributes. In the literature on multidimensional inequality measurement—in which all attributes are transferable—our result provides a normative underpinning of Kolm's [10] budget dominance relation. In the needs literature—in which there are two attributes, one transferable (income) and one non-transferable (needs)—our result provides a normative underpinning of Bourguignon's [3] welfare dominance criterion.

The next section introduces notation and presents basic concepts. Section 3 provides a formal definition of the consistent Pigou–Dalton principle. In Section 4 we state and prove the main result (Theorem 1). The remainder of the paper discusses implications of Theorem 1. In Section 5 we consider the framework of multidimensional inequality measurement. We link the consistent Pigou–Dalton principle to the uniform majorization principle and the correlation increasing majorization principle. Furthermore, we show that Kolm's [10] budget dominance relation coincides with the intersection of all social rankings described in Theorem 1. In Section 6 we focus on the needs framework. We interpret our main result as a justification for using absolute (rather than, e.g., relative) equivalence scales to correct for differences in needs. Here, we show that the intersection of all social rankings described in Theorem 1 coincides with an extension of Bourguignon's [3] welfare dominance criterion.

2. Notation

We consider a finite set M of at least two individuals and a non-empty finite set A of attributes. We distinguish transferable from non-transferable attributes. Whether or not an attribute is transferable is ultimately a normative choice: an attribute is transferable if one believes that transferring attribute amounts from better-off to worse-off individuals, while preserving the total amount of this attribute, is desirable. Income would be a typical example of a transferable attribute, whereas health status would be a typical example of a non-transferable attribute. The set T collects the transferable attributes and is non-empty, the set N collects the non-transferable attributes, and $A = T \cup N$. Concerning the sets T and N, two extreme positions have received considerable attention. The multidimensional inequality literature puts $N = \emptyset$ and labels each attribute as transferable. In the needs literature there are two attributes: income (transferable) and needs (non-transferable).

The variable that measures the amount of attribute k runs over some closed interval $A_k \subset \mathbb{R}$. As such, the domains $\mathcal{A}_T$ and $\mathcal{A}_N$ of transferable and non-transferable attribute bundles are cartesian products of closed intervals. The domain of attributes is[1]

$$\mathcal{A} = \underbrace{\prod_{t \in T} A_t}_{\mathcal{A}_T} \times \underbrace{\prod_{n \in N} A_n}_{\mathcal{A}_N} \subset \underbrace{\prod_{t \in T} \mathbb{R}}_{\mathbb{R}^T} \times \underbrace{\prod_{n \in N} \mathbb{R}}_{\mathbb{R}^N}.$$

Each attribute bundle x in $\mathcal{A}$ can be decomposed into (x_T, x_N) with x_T in $\mathcal{A}_T$ the transferable part and x_N in $\mathcal{A}_N$ the non-transferable part. We extend this decomposition to arbitrary vectors in $\mathbb{R}^A$ and we write $\varepsilon = (\varepsilon_T, \varepsilon_N)$ with ε_T in $\mathbb{R}^T$ and ε_N in $\mathbb{R}^N$.

Each individual i in M is endowed with some attribute bundle $x^i = (x^i_k)_{k \in A}$ in $\mathcal{A}$. The number x^i_k in A_k measures the amount of attribute k individual i is endowed with. Superscripts refer to individuals and subscripts to attributes. A distribution of attributes over the set of individuals is an $|A| \times |M|$ matrix X with the attribute bundle x^i at the ith column. The set $\mathcal{D} = \mathcal{A}^M$ is the domain of distributions. We assume that the attribute bundles completely capture the relevant differences between the individuals. In other words, the individuals are identical except for their attribute bundles.

Vector and matrix inequalities are denoted by $\geqslant$, $>$, and $\gg$: we write $X \geqslant Y$ if the inequality $x^i_k \geqslant y^i_k$ holds for each individual i and each attribute k, $X > Y$ if in addition at least one of the inequalities holds strictly, and $X \gg Y$ if all the inequalities hold strictly. We write 0 for zero vectors. For two vectors x and y in $\mathbb{R}^\ell$, we write $x \cdot y$ for the sum $x_1 y_1 + x_2 y_2 + \cdots + x_\ell y_\ell$.

[1] The notation $\mathbb{R}^A$ follows the notation B^A for the collection of maps from a set A to a set B.

A social ranking $\succsim$ is a quasi-ordering in $\mathcal{D}$.[2] The asymmetric and symmetric components of $\succsim$ are denoted by $\succ$ and $\sim$, respectively.[3] The social ranking $\succsim$ induces a quasi-ordering in $\mathcal{A}$. We denote this induced relation by $R_{\succsim}$: for each x and y in $\mathcal{A}$,

$$x R_{\succsim} y \quad \text{if and only if} \quad (x \quad x \quad \cdots \quad x) \succsim (y \quad y \quad \cdots \quad y). \tag{1}$$

It is compelling to interpret the relation $R_{\succsim}$ as the ranking of attribute bundles in terms of individual well-being that underlies the social ranking $\succsim$. Since individuals only differ with respect to attribute bundles, a choice between two distributions in which they all have the same attribute bundle boils down to a choice of the best attribute bundle at the individual level. If the social ranking $\succsim$ in $\mathcal{D}$ is complete, then also the induced relation $R_{\succsim}$ in $\mathcal{A}$ is complete. The asymmetric part of $R_{\succsim}$ is denoted by $P_{\succsim}$.

We now introduce three properties for a social ranking $\succsim$ of distributions. Monotonicity and anonymity are natural requirements.

Monotonicity. For each X and Y in $\mathcal{D}$, the matrix inequality $X > Y$ implies $X \succ Y$.

Anonymity. For each X in $\mathcal{D}$, we have indifference between X and all distributions that are equal to X up to a rearrangement of its columns (individuals).

Monotonicity makes sense in the multidimensional context if each attribute is a *good*—not a bad. A monotonic social ranking registers an increase in an attribute as an improvement. Anonymity imposes that the names of the individuals are not taken into account. This property makes sense since all relevant characteristics of individual i are incorporated in the attribute bundle x^i. In other words, the relevant differences between the individuals are captured by the attribute bundles. The third property incorporates completeness, continuity, and separability.

Additive representability. There exist C^1-maps $u^i : \mathbb{R}^A \to \mathbb{R}$, one for each i in M, such that, for each X and Y in $\mathcal{D}$,

$$X \succsim Y \quad \text{if and only if} \quad \sum_{i \in M} u^i\left(x^i\right) \geqslant \sum_{i \in M} u^i\left(y^i\right). \tag{2}$$

Imposing this property forces the social ranking to be complete, continuous (hence, representable), and separable over individuals (in order to compare two distributions, only those individuals who experience a change in their attribute bundles are taken into account, individuals who experience a status quo have no impact). The combination of monotonicity, anonymity, continuity, and separability implies the representability of the social ranking as in (2) with u^i continuous and $u^i = u$ for each i in M [2]. The map u represents the ordering $R_{\succsim}$ in $\mathcal{A}$ induced by the social ranking $\succsim$ in $\mathcal{D}$. The technical condition that the representation involves maps that are continuously differentiable (or C^1) is used in the proof of Theorem 1.

3. A consistent Pigou–Dalton principle

According to the unidimensional Pigou–Dalton principle, a transfer from poor to rich results in a distribution that is socially worse than the initial distribution. Now consider the multidimen-

[2] A transitive and reflexive binary relation is a quasi-ordering. A complete quasi-ordering is an ordering.

[3] That is, $X \succ Y$ if $X \succsim Y$ and not $Y \succsim X$, and $X \sim Y$ if $X \succsim Y$ and $Y \succsim X$.

sional setting. Let R be a quasi-ordering in $\mathcal{A}$ that ranks attribute bundles in terms of individual well-being. The R-Pigou–Dalton principle requires that whenever individual i is—according to R—not worse off than individual j, then a mean-preserving transfer from j to i in one or more attributes, decreases social welfare.

***R*-Pigou–Dalton principle.** For each X and Y in $\mathcal{D}$, for each i and j in M with $x^i R x^j$, and for each $\varepsilon = (\varepsilon_T, \varepsilon_N)$ in $\mathbb{R}^A$ with $\varepsilon_T > 0$ and $\varepsilon_N = 0$, we have that if

$$Y = \begin{pmatrix} \cdots & x^i + \varepsilon & \cdots & x^j - \varepsilon & \cdots \end{pmatrix},$$

with X and Y coinciding except for individuals i and j, then $X \succ Y$.

In this definition, the transfer from j to i is regressive in terms of well-being. The restriction $\varepsilon_N = 0$ reflects that only transferable attributes are involved. The unidimensional $\geqslant$-Pigou–Dalton principle (with $\geqslant$ the natural ordering in $\mathbb{R}$) coincides with the unidimensional Pigou–Dalton principle. The normative contents of the R-Pigou–Dalton principle crucially depends on the *choice* of the quasi-ordering R in $\mathcal{A}$. It seems natural to choose R equal to the well-being concept underlying $\succsim$, i.e., to choose $R = R_{\succsim}$. We refer to this version of the R-Pigou–Dalton principle as the consistent Pigou–Dalton principle.

Consistent Pigou–Dalton principle. The social ranking $\succsim$ in $\mathcal{D}$ satisfies the consistent Pigou–Dalton principle if it satisfies the $R_{\succsim}$-Pigou–Dalton principle, where the quasi-ordering $R_{\succsim}$ in $\mathcal{A}$ is induced by the social ranking $\succsim$ in $\mathcal{D}$ as in (1).

Note that if a unidimensional social ranking $\succsim$ satisfies monotonicity, then the induced well-being ranking $R_{\succsim}$ coincides with $\geqslant$. In that case, the consistent Pigou–Dalton principle coincides with the unidimensional Pigou–Dalton principle.

4. Main result

The next theorem investigates the effect of imposing monotonicity, anonymity, additive representability, and the consistent Pigou–Dalton principle upon a social ranking of distributions.

Theorem 1. *A social ranking $\succsim$ satisfies monotonicity, anonymity, additive representability, and the consistent Pigou–Dalton principle if and only if there exist*

- *a vector p_T in $\mathbb{R}^T$ with $p_T \gg 0$,*
- *a C^1-map $\varphi : \mathbb{R} \to \mathbb{R}$ which is strictly increasing and strictly concave, and*
- *a C^1-map $\psi : \mathcal{A}_N \to \mathbb{R}$ which is strictly increasing in each variable,*

such that, for each X and Y in $\mathcal{D}$, we have

$$X \succsim Y \quad \text{if and only if} \quad \sum_{i \in M} \varphi\big(p_T \cdot x_T^i + \psi\big(x_N^i\big)\big) \geqslant \sum_{i \in M} \varphi\big(p_T \cdot y_T^i + \psi\big(y_N^i\big)\big).$$

Proof. The particular representation of the social ranking $\succsim$ satisfies the four conditions. We focus on the reverse implication. Therefore, let $\succsim$ be an ordering in $\mathcal{D}$ that satisfies the four conditions. We proceed in five steps.

Step 1. *Starting up the proof*: *organizing partial differential equations.*

Since the social ranking $\succsim$ is anonymous, monotonic, and additively representable, it can be represented by a C^1-map on the domain of distributions:

$$SC : \mathcal{D} \to \mathbb{R} : X \mapsto \sum_{i \in M} u(x^i),$$

with $u : \mathbb{R}^A \to \mathbb{R}$ a strictly increasing function that represents the induced ordering $R_{\succsim}$. The monotonicity of the relation $\succsim$ implies the monotonicity of $R_{\succsim}$ (if $x > y$, then $x P_{\succsim} y$).

Next, we study the indifference surfaces of the map u. The monotonicity of $R_{\succsim}$ implies that the indifference surfaces are thin. Let $t \in T$. Let x and y in $\mathcal{A}$ satisfy $u(x) \geqslant u(y)$. According to consistent Pigou–Dalton, the distribution $(x \quad y \quad z \quad \cdots \quad z)$ is socially preferred to $(x + \varepsilon \quad y - \varepsilon \quad z \quad \cdots \quad z)$ with ε in $\mathbb{R}^A$, $\varepsilon_t > 0$, and $\varepsilon_k = 0$ for each $k \neq t$. Given the additive representability, it follows that

$$u(y) - u(y - \varepsilon) > u(x + \varepsilon) - u(x) > 0.$$

Divide by ε_t, take the limits for ε_t to 0, and obtain $D_t u(y) \geqslant D_t u(x) \geqslant 0$. In sum,

$$\text{for each } t \text{ in } T, \quad \text{if } u(x) \geqslant u(y), \quad \text{then } 0 \leqslant D_t u(x) \leqslant D_t u(y).$$

If $u(x) = u(y)$, then $u(x) \geqslant u(y)$ and $u(y) \geqslant u(x)$, and the partial derivative with respect to a transferable attribute is a constant: $D_t u(x) = D_t u(y)$. The utility level $u(x)$ completely determines the derivative $D_t u(x)$.

As a consequence, for each t in T and for each x in $\mathcal{A}$, it holds that $D_t u(x) = V_t(u(x))$, with $V : \mathbb{R} \to (\mathbb{R}^+)^T$ a vector valued map. This is a system of first-order partial differential equations.

Step 2. *The equation "for each $z = (z_1, z_2, \ldots, z_k)$, we have $D_1 f(z) = g \circ f(z)$" and its solution.*[4]

Let f be C^1 and consider the above partial differential equation. Let G be a primitive of $1/g$, i.e., $G' = 1/g$. We claim that the map f satisfies the equation if and only if f is implicitly defined by

$$G\big(f(z)\big) = z_1 + \psi(z_2, z_3, \ldots, z_k),$$

with ψ an arbitrary map from $\mathbb{R}^{k-1}$ to $\mathbb{R}$. Such an implicitly defined map f clearly satisfies the differential equation. To show necessity, rewrite the differential equation as $\frac{1}{g \circ f(z)} D_1 f(z) = 1$ and integrate with respect to z_1. If g is almost everywhere positive, then G is strictly increasing, the inverse map $\varphi = G^{-1}$ is well defined, and the solution can be rewritten: $f(z) = \varphi(z_1 + \psi(z_2, z_3, \ldots, z_k))$. As f is assumed to be C^1, the map ψ has to be C^1. In this particular case, the above quasi-linear first-order partial differential equation characterizes C^1-transforms of a quasi-linear function.

Step 3. *The case $T = \{1\}$, the equation $D_1 u(x) = V_1(u(x))$.*

According to Step 2—plug in u and V for f and g—we obtain $u(x) = \varphi(x_1 + \psi(x_N))$. The map φ is strictly concave. Otherwise, if $D\varphi(x_1 + \psi(x_N)) = D\varphi(\tilde{x}_1 + \psi(x_N))$ with $x_1 > \tilde{x}_1$, then a small progressive transfer might result in the same social welfare.

[4] Andrei Polyanin (http://eqworld.ipmnet.ru/en/solutions/fpde/fpde2103.pdf) presents a solution of this equation. See also [12].

Step 4. *The case $T = \{1, 2\}$, the equations $D_t u(x) = V_t(u(x))$, $t = 1, 2$.*

Apply Step 3 for both transferable attributes and obtain two representations:

$$u(x_1, x_2; x_N) = \alpha\big(x_1 + \beta(x_2; x_N)\big)$$

and

$$u(x_1, x_2; x_N) = \gamma\big(x_2 + \delta(x_1; x_N)\big).$$

Consider the marginal rates of substitution. For each $(x_1, x_2; x_N)$ and for each n in N, we have

$$\frac{D_2 u(x_1, x_2; x_N)}{D_1 u(x_1, x_2; x_N)} = D_2\beta(x_2; x_N) \overset{\text{(i)}}{=} \frac{1}{D_1\delta(x_1; x_N)}$$

and

$$\frac{D_n u(x_1, x_2; x_N)}{D_1 u(x_1, x_2; x_N)} = D_n\beta(x_2; x_N) \overset{\text{(ii)}}{=} \frac{D_n\delta(x_1; x_N)}{D_1\delta(x_1; x_N)}.$$

Identity (i) implies that $D_2\beta(x_2; x_N)$ only depends on x_N, not on x_2; say, $\beta(x_2; x_N) = x_2 f(x_N) + g(x_N)$. Similarly, identities (ii)—one for each n in N—imply that $D_n\beta(x_2; x_N)$ only depends on x_N, not on x_2; hence $f(x_N)$ must be a constant, say $f(x_N) = a_2$. In conclusion, u obtains the desired representation: $u(x_1, x_2; x_N) = \alpha(x_1 + a_2 x_2 + g(x_N))$. With respect to the strict concavity of α, the argument in Step 3 remains valid.

Step 5. *The case $T = \{1, 2, 3, \ldots\}$, the equations $D_t u(x) = V_t(u(x))$, t in T.*

We tackle this case by induction. Suppose Theorem 1 holds for $|T| < k$. Let $|T| = k$. Apply the induction hypothesis to two different subsets of T of cardinality $k - 1$ and obtain two representations. Combine these two representations—along the lines of Step 4—and conclude that u has the desired representation. □

As Theorem 1 shows, the imposition of the consistent Pigou–Dalton principle in an additively separable framework considerably limits the possibilities to rank distributions in $\mathcal{D}$: the induced well-being ranking in $\mathcal{A}$ obtains a quasi-linear structure. The next two sections apply this result to the framework of multidimensional inequality measurement and to the framework of needs, respectively.

5. Multidimensional inequality measurement

The literature on multidimensional social evaluation assumes that each attribute is transferable, i.e., $A = T$ and $N = \emptyset$. Imposing the four properties results in the following criterion: for each X and Y in $\mathcal{D} = \mathcal{A}_T^M$,

$$X \succsim Y \quad \text{if and only if} \quad \sum_{i \in M} \varphi\big(p \cdot x^i\big) \geqslant \sum_{i \in M} \varphi\big(p \cdot y^i\big), \tag{3}$$

with $\varphi : \mathbb{R} \to \mathbb{R}$ a strictly increasing and strictly concave C^1-map, and $p \gg 0$ in $\mathbb{R}^T$. We discuss the normative approach (5.1) and the dominance approach (5.2). For surveys of these two approaches, we refer to Weymark [17] and Trannoy [15], respectively.

5.1. Here, we link the consistent Pigou–Dalton principle to two other principles in the literature.

First, we consider the uniform majorization principle due to [10]. This principle requires that post-multiplying a distribution by a non-permutation bistochastic matrix increases social welfare.[5]

Uniform majorization principle. The social ranking $\succsim$ satisfies the *weak* uniform majorization principle if, for each distribution X in $\mathcal{D} = \mathcal{A}_T^M$ and each non-permutation bistochastic matrix B, we have $XB \succsim X$. If, in addition, $XB \neq X$ implies $XB \succ X$, then $\succsim$ satisfies the *strong* uniform majorization principle.

A social ranking $\succsim$ of the form (3) satisfies the weak uniform majorization principle. Indeed, the strict concavity of the map φ implies the inequality

$$\sum_{i \in M} \varphi\big(p \cdot (XB)^i\big) \geqslant \sum_{i \in M} \varphi\big(p \cdot x^i\big),$$

with $(XB)^i$ the ith column of XB.[6] Furthermore, the combination of $XB \neq X$ and $p \cdot XB = p \cdot X$ may occur. Hence, a ranking $\succsim$ of the form (3) does not satisfy the strong uniform majorization principle.

Second, we examine the correlation increasing majorization principle due to [16]. Consider two individuals and switch attribute amounts between them until one individual has more of each attribute than the other. The correlation increasing majorization principle demands that such switches—which increase the correlation of attributes—decrease social welfare. For each x and y in $\mathbb{R}^A$, let $x \wedge y = (\min\{x_k, y_k\})_{k \in A}$ and $x \vee y = (\max\{x_k, y_k\})_{k \in A}$.

Correlation increasing majorization principle. For each X and Y in $\mathcal{D} = \mathcal{A}_T^M$, and for each i and j in M, we have that if $x^i \vee x^j$ differs from x^i and x^j, and if

$$Y = \begin{pmatrix} \cdots & x^i \vee x^j & \cdots & x^i \wedge x^j & \cdots \end{pmatrix},$$

with X and Y coinciding except for individuals i and j, then $X \succ Y$.

It is straightforward to see that the consistent Pigou–Dalton principle—in combination with anonymity—implies the correlation increasing majorization principle. Indeed, assume that with respect to distribution X individual i is not worse off than individual j. Then, the move from distribution X to distribution Y involves the transfer $(x^j - x^i) \vee 0$ from j to i. If, with respect to X, j is not worse off than i, then apply anonymity to switch i and j, and repeat the previous argument.

5.2. Expression (3) defines different criteria, one for each choice of the map φ and of the vector p. The intersection of all these criteria defines a partial ranking.

Unanimity criterion $\succsim_T$ in $\mathcal{D} = \mathcal{A}_T^M$. Let X and Y be two distributions in $\mathcal{D}$. Then, $X \succsim_T Y$ if $X \succsim Y$ for each social ranking $\succsim$ that satisfies monotonicity, anonymity, additive representability, and the consistent Pigou–Dalton principle.

[5] A non-negative square matrix is said to be bistochastic if all of its row and column sums are equal to 1. A bistochastic matrix with only zeros and ones is a permutation matrix.

[6] See [11, p. 64].

Let $p \gg 0$ be a vector in $\mathbb{R}^T$. The condition

$$\varphi(p \cdot x^1) + \varphi(p \cdot x^2) + \cdots + \varphi(p \cdot x^n) \geqslant \varphi(p \cdot y^1) + \varphi(p \cdot y^2) + \cdots + \varphi(p \cdot y^n)$$

for each C^1-map $\varphi : \mathbb{R} \to \mathbb{R}$ that is strictly increasing and strictly concave, is one way to express that the n-tuple $(p \cdot x^1, p \cdot x^2, \ldots, p \cdot x^n)$ generalized Lorenz dominates the n-tuple $(p \cdot y^1, p \cdot y^2, \ldots, p \cdot y^n)$. Here, we refer to Kolm [9] and Shorrocks [13].

Kolm [10] interprets p as a price vector, the product $p \cdot x^i$ as the budget of individual i, and introduces the budget dominance criterion. According to this criterion distribution X dominates distribution Y if for each price vector the budgets induced by X generalized Lorenz dominate the budgets induced by Y.

As a consequence, Kolm's criterion coincides with the unanimity criterion $\succsim_T$ and Theorem 1 provides a normative underpinning of Kolm's criterion.

Corollary 1. *Let X and Y be two distributions in $\mathcal{D} = \mathcal{A}_T^M$. Then, $X \succsim_T Y$ if and only if X budget dominates Y.*

6. The needs framework

In the needs literature, one usually considers only two attributes: income and an index of needs.[7] Income is assumed to be transferable and is denoted by x_t in $A_t \subset \mathbb{R}$. The needs index is assumed to be non-transferable and is denoted by x_n in $A_n \subset \mathbb{R}$. Lower values of x_n correspond with higher needs. The domain $\mathcal{D}$ is the cartesian product $(A_t \times A_n)^M \subset (\mathbb{R}^2)^M$. A social ranking that satisfies the four properties ranks the distributions X and Y in $\mathcal{D}$ as follows:

$$X \succsim Y \quad \text{if and only if} \quad \sum_{i \in M} \varphi\big(x_t^i + \psi\big(x_n^i\big)\big) \geqslant \sum_{i \in M} \varphi\big(y_t^i + \psi\big(y_n^i\big)\big), \tag{4}$$

with φ and ψ strictly increasing C^1 and with φ also strictly concave. We discuss the equivalence scale literature (6.1) and Bourguignon's [3] dominance criterion (6.2).

6.1. The standard equivalence scale approach proceeds in two steps. *First*, in order to compare the living standards across individuals with different levels of needs, one determines for each individual an equivalent income. The equivalent income function $E : A_t \times A_n \to \mathbb{R}$ adjusts incomes for needs and is assumed to be strictly increasing and C^1. *Second*, welfare is defined as the sum of utilities of the equivalent incomes $\sum_{i \in M} U(E(x_t^i, x_n^i))$ with U strictly increasing and C^1. The relative and income-independent transform, i.e., $E(x_t, x_n) = x_t/S(x_n)$, is a well known and frequently used example.

Theorem 1—as repeated in expression (4)—indicates that the imposition of the four axioms forces the equivalent income to be a quasi-linear transform of income: $E(x_t, x_n)$ must be defined as $x_t + \psi(x_n)$. In words, the income x_t is adjusted for needs by adding an income-independent equivalence scale $\psi(x_n)$. This particular equivalent income function is known as an absolute and income-independent transform. Relative transforms conflict with the consistent Pigou–Dalton principle and are therefore excluded from expression (4).[8]

[7] See [1,3].

[8] Related results can be found in [4–6,8,14].

6.2. Expression (4) defines different criteria, one for each choice of φ and ψ. As in Section 5.2 we consider the partial ranking generated by the intersection of all these criteria. We define this unanimity criterion and show that it extends the welfare dominance criterion of Bourguignon [3].

Let us first introduce an extension of Bourguignon's [3] welfare dominance criterion. Let X and Y be two distributions in $\mathcal{D} = (A_t \times A_n)^M$. Let

$$L(X, Y) = \{l \mid \text{there exists an } i \text{ in } M \text{ such that } l = x_n^i \text{ or } l = y_n^i\}$$

be the set of all needs values that occur in X or in Y. For each value l in $L(X, Y)$, let $U_l : \mathbb{R} \to \mathbb{R}$ be a C^2 utility function. Then, $X \succsim_B Y$ if

$$\sum_{i \in M} U_{x_n^i}(x_t^i) \geqslant \sum_{i \in M} U_{y_n^i}(y_t^i),$$

for each profile $U = (U_l)_{l \in L(X,Y)}$ of C^2-functions that satisfies

- $U_l'(y) > 0$ and $U_l''(y) < 0$ for each y in $\mathbb{R}$ and for each l in $L(X, Y)$,
- for each l_1 and l_2 in $L(X, Y)$ and for each y in $\mathbb{R}$, we have that if $l_1 < l_2$, then $U_{l_1}'(y) \geqslant U_{l_2}'(y)$,
- $\lim_{y \to \infty} U_l(y) = \bar{u}$ for each l in $L(X, Y)$.

The first two conditions state that marginal utilities are everywhere positive and decreasing, and higher needs imply higher marginal utilities for the same income. The final condition imposes that needs differences do not matter if individuals are infinitely rich. This limit condition is weak (recall that observed incomes are finite) and technical (it simplifies Proposition 1). A profile U that meets these three conditions is said to be an $L(X, Y)$-profile.

In the particular case where the $|M|$-tuples of needs values in X and Y are equal up to a permutation, the relation $\succsim_B$ boils down to Bourguignon's criterion. The next proposition rephrases $\succsim_B$ in terms of distribution functions and is similar to Bourguignon's theorem [3, p. 73].

Proposition 1. *Let X and Y be two distributions in $\mathcal{D} = (A_t \times A_n)^M$. Let $L = L(X, Y)$ be the set of all needs values that occur in X or in Y. For each l in L, let $F_X(\cdot|l)$ be the distribution function of income in X conditional upon the value of needs being l (and similar for Y). Then, $X \succsim_B Y$ if and only if*

$$\sum_{l \in L} \int_{z=-\infty}^{\beta_l} \big(F_X(z|l) - F_Y(z|l)\big) \leqslant 0, \tag{5}$$

for each $|L|$-tuple $(\beta_l)_{l \in L}$ in $\mathbb{R}^L$ that satisfies $\beta_{l_1} \leqslant \beta_{l_2}$ if $l_1 < l_2$.

Proof. We rewrite $\succsim_B$ in terms of dominance. Observe that $X \succsim_B Y$ if and only if, for each L-profile U, we have

$$\Delta_U(X, Y) = \sum_{l \in L} \int_{\alpha=-\infty}^{+\infty} U_l(\alpha)\big(dF_X(\alpha|l) - dF_Y(\alpha|l)\big) \geqslant 0,$$

where dF_X and dF_Y are the densities that correspond to F_X and F_Y. Integration by parts converts $\Delta_U(X, Y)$ into

$$\sum_{l\in L}\lim_{\alpha\to+\infty} U_l(\alpha)\big(F_X(\alpha|l)-F_Y(\alpha|l)\big)-\sum_{l\in L}\int_{\alpha=-\infty}^{+\infty} U_l'(\alpha)\big(F_X(\alpha|l)-F_Y(\alpha|l)\big).$$

As $\lim_{+\infty} U_l(\alpha)=\bar{u}$ for each l, the first sum is equal to 0. Hence, $X \succsim_B Y$ if and only if, for each L-profile U, we have

$$\Delta_U(X,Y)=-\sum_{l\in L}\int_{\alpha=-\infty}^{+\infty} U_l'(\alpha)\big(F_X(\alpha|l)-F_Y(\alpha|l)\big)\geqslant 0.$$

Let l_0 be the smallest value in L. For each l in L, the inequalities $U_l''<0$ and $U_{l_0}'\geqslant U_l'>0$ implicitly define a strictly increasing map f_l by $U_l'(\alpha)=U_{l_0}'(f_l(\alpha))$ for each α. Furthermore, for each α, we have $f_{l_0}(\alpha)=\alpha$ and $f_{l_1}(\alpha)\leqslant f_{l_2}(\alpha)$ if $l_1<l_2$. Also, $\lim_{-\infty} f_l(\alpha)=-\infty$ and $\lim_{+\infty} f_l(\alpha)=+\infty$. Let $\Gamma_l'(\alpha)=F_X(\alpha|l)-F_Y(\alpha|l)$ and write

$$\Delta_U(X,Y)=-\sum_{l\in L}\int_{\alpha=-\infty}^{+\infty} U_{l_0}'\big(f_l(\alpha)\big)\Gamma_l'(\alpha).$$

In each term we change the variable α into $\beta=f_l(\alpha)$ and obtain

$$\Delta_U(X,Y)=-\sum_{l\in L}\int_{\beta=-\infty}^{+\infty} U_{l_0}'(\beta)\Gamma_l'\big(f_l^{-1}(\beta)\big)\big(f_l^{-1}\big)'(\beta).$$

Integrate by parts, put $\lim_{+\infty} U_{l_0}'(\beta)=0$ (a consequence of $\lim_{+\infty} U_{l_0}(\beta)=\bar{u}$), use the linearity of integration, plug in the definition of Γ_l, and conclude that $X \succsim_B Y$ if and only if

$$\Delta_U(X,Y)=\int_{\beta=-\infty}^{+\infty} U_{l_0}''(\beta)\left(\sum_{l\in L}\int_{z=-\infty}^{f_l^{-1}(\beta)} F_X(z|l)-F_Y(z|l)\right)\geqslant 0$$

for each L-profile U. Recall that $U_{l_0}''<0$ and conclude that $X \succsim_B Y$ if

$$\sum_{l\in L}\int_{z=-\infty}^{f_l^{-1}(\beta)}\big(F_X(z|l)-F_Y(z|l)\big)\leqslant 0 \tag{6}$$

for each β in $\mathbb{R}$ and for each profile $f=(f_l)_{l\in L}$ of increasing maps that satisfies, for each α, $f_{l_0}(\alpha)=\alpha$ and $f_{l_1}(\alpha)\leqslant f_{l_2}(\alpha)$ if $l_1<l_2$. Manipulating the values of $U_{l_0}''(\beta)$ one can show that also the reverse holds: "the negation of (6) for a particular profile f and value β" in combination with "$\Delta_U(X,Y)\geqslant 0$ for each profile U" cannot occur. □

Next, we introduce the unanimity social ranking as the intersection of the rankings satisfying the axioms in Theorem 1 for the current domain of distributions.

Unanimity criterion $\succsim_N$ in $\mathcal{D}=(A_t\times A_n)^M$. Let X and Y be two distributions in $\mathcal{D}$. Then, $X \succsim_N Y$ if $X \succsim Y$ for each social ranking $\succsim$ that satisfies monotonicity, anonymity, additive representability, and the consistent Pigou–Dalton principle.

Recall expression (4). The social ranking $\succsim_N$ expresses unanimity among utilitarian planners using a wide set of absolute equivalence scales to correct income for needs. Proposition 2 links the unanimity criterion $\succsim_N$ with the Bourguignon criterion $\succsim_B$.[9]

Proposition 2. *Let X and Y be two distributions in $\mathcal{D} = (A_t \times A_n)^M$. Then, $X \succsim_N Y$ if and only if $X \succsim_B Y$.*

Proof. Let X and Y be two distributions in $\mathcal{D}$. We keep the previous notation: $L = L(X, Y)$ is the set of needs values and for each l in L, $F_X(\cdot|l)$ and $F_Y(\cdot|l)$ are the conditional distributions with $dF_X(\cdot|l)$ and $dF_Y(\cdot|l)$ the corresponding densities.

By the definition of the unanimity relation $\succsim_N$ and Theorem 1, we have that $X \succsim_N Y$ if and only if the welfare difference

$$\Delta_{\varphi,\psi}(X, Y) = \sum_{l \in L} \int_{\alpha=-\infty}^{+\infty} \varphi\big(\alpha + \psi(l)\big)\big(dF_X(\alpha|l) - dF_Y(\alpha|l)\big) \geqslant 0,$$

for each strictly increasing and strictly concave C^1-function φ and for each strictly increasing C^1-function ψ.

We will translate this condition towards Bourguignon's dominance criterion as described in Proposition 1. We start by integrating $\Delta_{\varphi,\psi}(X, Y)$ by parts:

$$\sum_{l \in L} \left(\lim_{+\infty} \varphi\big(\alpha + \psi(l)\big)\big(F_X(\alpha|l) - F_Y(\alpha|l)\big) - \int_{\alpha=-\infty}^{+\infty} \varphi'\big(\alpha + \psi(l)\big)\big(F_X(\alpha|l) - F_Y(\alpha|l)\big) \right).$$

Let us look at the sum of the first terms. For α sufficiently large (i.e., larger than the observed incomes), the differences $\delta_l = F_X(\alpha|l) - F_Y(\alpha|l)$ remain constant. We obtain that

$$\sum_{l \in L} \delta_l \lim_{+\infty} \varphi\big(\alpha + \psi(l)\big) = \sum_{l \in L} \delta_l \lim_{+\infty} \varphi(\alpha) = \lim_{+\infty} \varphi(\alpha) \sum_{l \in L} \delta_l = 0.$$

It follows that $X \succsim_N Y$ if and only if, for each couple (φ, ψ), we have

$$\Delta_{\varphi,\psi}(X, Y) = -\sum_{l \in L} \int_{\alpha=-\infty}^{+\infty} \varphi'\big(\alpha + \psi(l)\big)\big(F_X(\alpha|l) - F_Y(\alpha|l)\big) \geqslant 0.$$

For each l in L, let $\Gamma_l(\alpha) = \int_{z=-\infty}^{\alpha} F_X(z|l) - F_Y(z|l)$ and $\beta = \alpha + \psi(l)$. We obtain

$$\Delta_{\varphi,\psi}(X, Y) = -\sum_{l \in L} \int_{\beta=-\infty}^{+\infty} \varphi'(\beta)\Gamma_l'\big(\beta - \psi(l)\big).$$

Integrate by parts and use the linearity of integration. Conclude that

$$\Delta_{\varphi,\psi}(X, Y) = -\lim_{+\infty} \left(\varphi'(\beta) \sum_{l \in L} \Gamma_l\big(\beta - \psi(l)\big) \right) + \int_{\beta=-\infty}^{+\infty} \varphi''(\beta) \left(\sum_{l \in L} \Gamma_l\big(\beta - \psi(l)\big) \right).$$

[9] Fleurbaey, Hagneré, and Trannoy [7] obtain a similar result using relative equivalence scales.

Recall that φ is strictly increasing and strictly concave.[10] Hence, the conclusion "$\Delta_{\varphi,\psi}(X,Y) \geqslant 0$ for each couple (φ,ψ)" holds if the following two conditions are met:

(A) for each $|L|$-tuple $\beta = (\beta_l)_{l \in L}$ that satisfies $\beta_{l_1} \leqslant \beta_{l_2}$ if $l_1 < l_2$, we have

$$\Delta_\beta(X,Y) = \sum_{l \in L} \int_{z=-\infty}^{\beta_l} \big(F_X(z|l) - F_Y(z|l)\big) \leqslant 0,$$

(B) for $\beta = (+\infty, +\infty, \dots, +\infty)$, we have

$$\Delta_{(+\infty,+\infty,\dots,+\infty)}(X,Y) = \sum_{l \in L} \int_{z=-\infty}^{+\infty} \big(F_X(z|l) - F_Y(z|l)\big) \leqslant 0.$$

Condition (B) can be dropped as it is implied by condition (A). Indeed, for β in $\mathbb{R}$ sufficiently large, the sign of $\Delta_{(\beta,\beta,\dots,\beta)}(X,Y)$ does not change anymore. Hence, if condition (A) is met, then $X \succsim_N Y$.

Manipulating the values of $\varphi'(+\infty)$ and $\varphi''(\beta)$ one can show that also the reverse holds. Indeed, "the violation of condition (A) for a particular $|L|$-tuple" in combination with "$\Delta_{\varphi,\psi}(X,Y) \geqslant 0$ for each profile (φ,ψ)" cannot occur.

In sum, distribution X dominates distribution Y according to the unanimity criterion $\succsim_N$ if and only if condition (A) is met. From Proposition 1 we know that distribution X dominates distribution Y according to the extended Bourguignon criterion $\succsim_B$ if and only if condition (A) is met. Therefore, the relations $\succsim_B$ and $\succsim_N$ coincide. □

Acknowledgments

We thank Koen Decancq, Benjamin Polak, Ernesto Savaglio, Alain Trannoy, and John Weymark for valuable comments. The authors are solely responsible for remaining shortcomings. Kristof Bosmans gratefully acknowledges financial support from the Fund for Scientific Research—Flanders (Grant G.0005.04) and the Interuniversity Attraction Poles network funded by the Federal Public Planning Service, Belgian Science Policy (Grant P5/21-A). Erwin Ooghe gratefully acknowledges the hospitality of GREQAM-IDEP, Marseille, and the accompanying financial support of the Fund for Scientific Research—Flanders.

References

[1] A.B. Atkinson, F. Bourguignon, Income distribution and differences in needs, in: G.R. Feiwel (Ed.), Arrow and the Foundations of the Theory of Economic Policy, Macmillan, London, 1987.

[2] C. Blackorby, D. Donaldson, M. Auersperg, A new procedure for the measurement of inequality within and among population subgroups, Can. J. Econ. 14 (1981) 665–685.

[3] F. Bourguignon, Family size and social utility: Income distribution dominance criteria, J. Econometrics 42 (1989) 67–80.

[4] B. Capéau, E. Ooghe, On comparing heterogeneous populations: Is there really a conflict between welfarism and a concern for greater equality in living standards? Math. Soc. Sci. 53 (2007) 1–28.

[10] According to Aleksandrov's theorem the second order derivative φ'' of the convex function φ exists almost everywhere.

[5] U. Ebert, Social welfare when needs differ: An axiomatic approach, Economica 64 (1997) 233–244.
[6] U. Ebert, P. Moyes, Equivalence scales reconsidered, Econometrica 71 (2003) 319–343.
[7] M. Fleurbaey, C. Hagneré, A. Trannoy, Welfare comparisons with bounded equivalence scales, J. Econ. Theory 110 (2003) 309–336.
[8] P.J. Hammond, Dual interpersonal comparisons of utility and the welfare economics of income distribution, J. Public Econ. 7 (1979) 51–71.
[9] S.-Ch. Kolm, The optimal production of social justice, in: J. Margolis, H. Guitton (Eds.), Public Economics, Macmillan, London, 1969.
[10] S.-Ch. Kolm, Multidimensional egalitarianisms, Quart. J. Econ. 91 (1977) 1–13.
[11] A. Marshall, I. Olkin, Inequalities: Theory of Majorization and Its Applications, Academic Press, New York, 1979.
[12] A.D. Polyanin, V.F. Zaitsev, A. Moussiaux, Handbook of First Order Partial Differential Equations, Taylor & Francis, London, 2002.
[13] A.F. Shorrocks, Ranking income distributions, Economica 50 (1983) 3–17.
[14] A.F. Shorrocks, Inequality and welfare evaluation of heterogeneous income distributions, J. Econom. Ineq. 2 (2004) 193–218.
[15] A. Trannoy, Multidimensional egalitarianism and the dominance approach: A lost paradise? in: F. Farina, E. Savaglio (Eds.), Inequality and Economic Integration, Routledge, London, 2006.
[16] K.-Y. Tsui, Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation, Soc. Choice Welfare 16 (1999) 145–157.
[17] J.A. Weymark, The normative approach to the measurement of multidimensional inequality, in: F. Farina, E. Savaglio (Eds.), Inequality and Economic Integration, Routledge, London, 2006.