

Decomposing bivariate dominance for social welfare comparisons*

Tina Gottschalk Marling[†] Troels Martin Range[‡] Peter Sudhölter[§]
Lars Peter Østerdal[¶]

June 6, 2018

Abstract

The principal dominance concept for inequality-averse multidimensional social welfare comparisons, commonly known as lower orthant dominance, entails less or equal mass on all lower hyperrectangles of outcomes. Recently, it was shown that bivariate lower orthant dominance can be characterized in terms of two elementary mass transfer operations: diminishing mass transfer (reducing welfare) and correlation-increasing switches (increasing inequality). In this paper we provide a constructive algorithm, which decomposes the mass transfers into such welfare reductions and inequality increases.

JEL classification: C63, D63, I31

Keywords: social welfare, lower orthant dominance, first order dominance, algorithm, inequality aversion

1 Introduction

Dominance concepts are increasingly used for multidimensional comparisons of social welfare, inequality, and poverty (see, e.g., Aaberge and Brandolini 2014).¹ Such concepts are appealing, since they provide comparisons of the overall attainment of groups, which are robust for broad classes of individual and social preferences over the (multidimensional) outcomes.

An important and frequently used dominance concept for inequality-averse multidimensional social welfare comparisons is *lower orthant dominance*. The idea of using orthant dominance – and related (less restrictive) concepts – for inequality-averse social welfare comparisons was popularized by Atkinson and

*The authors are grateful to an associate editor and two anonymous referees of this journal for their remarks that helped to improve the paper. Support from Independent Research Fund Denmark (Grant-IDs: DFF-1327-00097 and DFF-6109-000132) is acknowledged, and the third author acknowledges support from the Spanish Ministerio de Economía y Competitividad under Project ECO2015-66803-P (MINECO/FEDER).

[†]SEF Energi A/S, Fåborgvej 44, 5700 Svendborg, Denmark; Email: tgm@sef.dk

[‡]Hospital of South West Jutland, Finsensgade 35, 6700 Esbjerg, Denmark, and Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Alfred Getz veg 3, NO-7491, Trondheim, Norway; Email: Troels.Martin.Range@rsyd.dk

[§]Department of Business and Economics, and COHERE, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark; Email: psu@sam.sdu.dk

[¶]Department of Economics, Copenhagen Business School, Porcelænshaven 16A, 2000 Frederiksberg, Denmark; Email: lpo.eco@cbs.dk

¹Stochastic dominance is not only useful in welfare economics, but also in many other fields. It is, for example, an important tool in decision theory (see, e.g., Levy 1992 or Müller and Stoyan 2002), finance (see, e.g., Sriboonchita, Dhompongsa, Wong, and Nguyen 2009), as well as in probability theory and statistics (see, e.g., Silvapulle and Sen 2011).

Bourguignon (1982), and it has been significantly developed and refined in several articles (see, e.g., Bourguignon and Chakravarty 2003, Duclos, Sahn, and Younger 2006, Duclos, Sahn, and Younger 2007, Gravel, Moyes, and Tarroux 2009, Gravel and Mukhopadhyay 2010, and Muller and Trannoy 2011).²

Suppose that there are multiple dimensions of welfare and that, for each dimension, a wellbeing indicator can take a finite number of possible values.³ We can then describe a population distribution by a probability mass function over the outcomes; i.e., by a function that assigns to each outcome the probability that a randomly selected individual obtains that outcome (or, put differently, it describes the share of all individuals in the population obtaining that outcome). For two probability mass functions (i.e., population distributions) f and g , the function f lower orthant dominates g if

- (1) the cumulative probability mass at f is smaller than or equal to that at g for every lower hyperrectangle.⁴

Until quite recently, a characterization based on elementary operations (i.e., conditions specifying exactly which simple changes in a distribution are allowed to obtain another distribution which dominates it) has been missing.⁵

This and a related gap were recently addressed and partially filled by Meyer and Strulovici (2011, 2015) and Müller (2013). Indeed, for the bivariate case the former authors showed that one probability mass function *supermodular dominates* another if and only if the former probability mass function can be obtained from the latter probability mass function by increasing probability mass transfers and correlation-increasing switches. An increasing probability mass transfer is simply a shift of mass from a worse to a better outcome (i.e., such a transfer is a welfare improvement).⁶ A correlation-increasing switch consists of two simultaneous transfers that move mass from intermediate outcomes to more extreme outcomes without changing the marginal distributions. For example, Tchen (1980), Epstein and Tanny (1980), Atkinson and Bourguignon (1982), Tsui (1999), Decancq (2012), and Sonne-Schmidt, Tarp, and Østerdal (2016) argue that correlation-increasing switches are operations that increase inequality. Lower orthant dominance may be expressed with the help of diminishing transfers (reducing welfare) and correlation-increasing switches (increasing inequality). More precisely, Müller’s (2013) characterization of lower orthant dominance for the general multivariate case directly implies, for the bivariate case, that (1) is equivalent to

- (2) a finite sequence of diminishing bilateral transfers and correlation-increasing switches exists such that g can be obtained by f and where each intermediate transformation leads to a distribution.

In welfare terms, supermodular dominance corresponds to a population that is better off but the inequality is *higher*, whereas lower orthant dominance corresponds to a population that is better off and the

²Note that lower orthant dominance has sometimes been referred to as “first order dominance”, particularly in the welfare economics literature. In order not to risk confusion with the usual stochastic order – the natural dominance concept for multidimensional social welfare comparisons with ordinal data (see, e.g., Arndt et al. 2012, Østerdal 2010, and Range and Østerdal 2017) – we use the term lower orthant dominance as customary in the probability theory literature (e.g. Shaked and Shanthikumar 2007).

³Chakravarty and Zoli (2012) mention a number of applications in which a wellbeing indicator is discrete by nature.

⁴In the continuous bivariate case, it is well-known (Atkinson and Bourguignon 1982) that (1) holds if and only if the average utility of f is at least as high as that of g for any non-decreasing utility function with negative cross derivative.

⁵For example, Moyes (2012) points out in his Footnote 13 that such characterization is missing, even though there are results in the literature that are making steps in this direction.

⁶Indeed, the usual stochastic order is completely characterized by such transfers, as shown by, for example, Strassen (1965) and Kamae, Krengel, and O’Brien (1977).

inequality is *lower*, i.e., the latter concept provides a basis for making inequality-averse social welfare comparisons.

The approach by Meyer and Strulovici (2010, 2015) is constructive, but it is not shown that a distribution can be obtained after each elementary operation. In contrast, Müller (2013) shows the existence of such sequences, where a distribution is obtained after each elementary operation, but an explicit construction is not given.

The main contribution of this paper is to provide a constructive proof of the equivalence between (1) and (2). The proof yields an algorithm that returns a set of diminishing transfers and correlation-increasing switches whenever a lower orthant dominance relationship exists. The algorithm has quadratic time complexity in the number of outcomes. We also mention that the results for upper orthant dominance would be similar.

2 Basics

Let $n, m \in \mathbb{N}$. For $x, y \in \mathbb{R}^m$, $x \leq (\geq) y$ denotes $x_i \leq (\geq) y_i$ for all $i = 1, \dots, m$, and $x < (>) y$ means $x \leq (\geq) y$ and $x \neq y$. Similarly, for two functions $f, g : D \rightarrow \mathbb{R}$ on an arbitrary domain D , we write $f \geq (\leq) g$ if $f(x) \geq (\leq) g(x)$ for all $x \in D$, and $f > (<) g$ if $f \geq (\leq) g$ and $f \neq g$.

Denote $X(n, m) = X = \{x \in \mathbb{N}^2 \mid x \leq (n, m)\}$ the rectangle of size $n \times m$ and $\mathcal{F}(n, m) = \mathcal{F} = \{f : X \rightarrow \mathbb{R}_+\}$ be the set of all real-valued non-negative functions on the domain X . For $\emptyset \neq Y \subseteq X$ let $\max Y = y \in X$ be the componentwise maximum defined by $y_i = \max\{x_i \mid x \in Y\}$ for $i = 1, 2$, and let $\min Y$ be the componentwise minimum defined analogously. Moreover, for $x \in X$, we denote the lower set $\downarrow x = \{y \in X \mid y \leq x\}$ as all elements of X having no component larger than the components of x .

In this paper we will use two fundamental operations. The first operation is a so-called diminishing transfer⁷, while the second is a correlation-increasing switch. For $f, g \in \mathcal{F}$ we say that g results from f

- by a *diminishing (bilateral) transfer* if there exist $x, y \in X$ such that $x < y$, $g(x) - f(x) = f(y) - g(y) > 0$, and $g(z) = f(z)$ for all $z \in X \setminus \{x, y\}$ (the underlying transfer is a transfer *from* y *to* x of size $\varepsilon = g(x) - f(x)$) and we use the notation $g = f_\varepsilon^{x \leftarrow y}$;
- by a *correlation-increasing switch* if there exist $x, y \in X$ such that $f(x) - g(x) = f(y) - g(y) = g(v) - f(v) = g(w) - f(w) > 0$ and $f(z) = g(z)$ for all $z \in X \setminus \{x, y, v, w\}$, where $v = \min(\{x, y\})$ and $w = \max(\{x, y\})$ (note that in this case x and y are incomparable; i.e., $x \not\leq y \not\leq x$, and that the underlying switch transfers $\varepsilon = f(x) - g(x)$ from each x and y to each v and w) and we use the notation $g = f_\varepsilon^{x \rightleftharpoons y}$.

Note that a diminishing transfer may be represented as a composition of even more elementary transfers, where one only transfers mass “horizontally” and the other only “vertically”. However, we use the current “composite” transfer because it is intuitive and simple and its decomposition into the mentioned elementary transfers is straightforward. A diminishing transfer is illustrated in Figure 2.1, where mass is transferred from y to x . It should be noted that a diminishing transfer can be decomposed into a sequence of unit diminishing transfers from $y = (y_1, y_2)$ to either $(y_1 - 1, y_2)$ or $(y_1, y_2 - 1)$. This decomposition is,

⁷Note that the term diminishing transfer has been used with another meaning by Lambert (2001, p. 62). We stick to its current meaning in order to be consistent with Range and Østerdal (2015).

however, not unique. A correlation-increasing switch is illustrated in Figure 2.2. As illustrated, mass is transferred from x to v and a similar mass is transferred from y to w . A symmetric transfer exists where mass is transferred from x to w and the same mass is transferred from y to v .

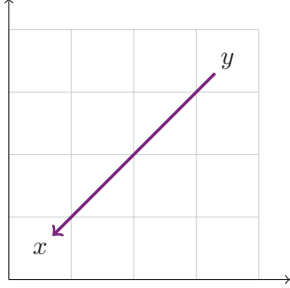


Figure 2.1: Diminishing transfer

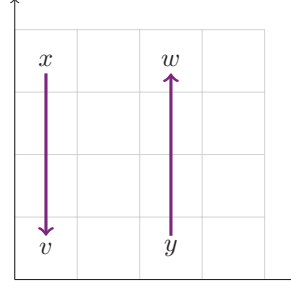


Figure 2.2: Correlation-increasing switch

The following notation is employed. For any $f \in \mathcal{F}(n, m)$ and $x \in X(n, m)$ denote $\tilde{f}(x) = \sum_{y \in \downarrow x} f(y)$, i.e., the accumulated mass below and left of the entry including the row and column in of the entry, and $\tilde{f}(0, i) = \tilde{f}(j, 0) = f(0, 0) = 0$ for all $i \in \{1, \dots, m\}$ and $j \in \{1, \dots, n\}$. Furthermore, we will regard the marginal distribution as the partial distribution of each dimension. The partial distributions are, respectively, the $\tilde{f}(n, i)$ and $\tilde{f}(j, m)$. A diminishing transfer only moves mass in the direction towards the origin (i.e., we only increase \tilde{f} -values). A careful inspection of its definition shows that a correlation-increasing switch preserves the marginal distributions.

Definition 2.1 Let $n, m \in \mathbb{N}$ and $f, g \in \mathcal{F}(n, m)$.

- (1) We say that f lower orthant dominates g , written $f \succeq_L g$, if $\tilde{g}(n, m) = \tilde{f}(n, m)$ and $\tilde{g} \geq \tilde{f}$.
- (2) If there exist $k \in \mathbb{N}$ and $f_1, \dots, f_k \in \mathcal{F}(n, m)$ such that $f = f_1$, $g = f_k$, and, for all $\ell \in \{2, \dots, k\}$, f_ℓ results from $f_{\ell-1}$ by a diminishing transfer or by a correlation-increasing switch, then we write that $f \succeq g$ and say that g arises from f by finitely many (here $k - 1$) diminishing transfers or correlation-increasing switches.

Finally, let \succ_L and \triangleright denote the respective strict relations.

Remark 2.2 Note that, for all $f, g \in \mathcal{F}(n, m)$, if $f \succeq_L g$ or $f \succeq g$, then $\sum_{x \in X(n, m)} f(x) = \sum_{x \in X(n, m)} g(x)$. Hence, all of the following statements on lower orthant domination or diminishing transfers and correlation increasing-switches that refer to $\mathcal{F}(n, m)$ are also valid for $\mathcal{F}_1(n, m) = \{f \in \mathcal{F}(n, m) \mid \sum_{x \in X(n, m)} f(x) = 1\}$, the set of probability distributions on $X(n, m)$. In particular, in Theorem 3.1, according to which $f \succeq g$ if and only if $f \succeq_L g$, the assumption $f, g \in \mathcal{F}(n, m)$ may be replaced by the assumption $f, g \in \mathcal{F}_1(n, m)$, i.e., \tilde{f} and \tilde{g} are cumulative distribution functions.

3 Disentangling lower orthant dominance

The key question is: If one distribution lower orthant dominates another distribution, can we then disentangle the dominance into the two elementary operations of diminishing transfers and correlation-increasing switches? Indeed, the two dominance relations given in Definition 2.1 can be proved to be equivalent, as elegantly done by Müller (2013) using duality theory. This result is given in Theorem 3.1.

Theorem 3.1 *Let $n, m \in \mathbb{N}$ and $f, g \in \mathcal{F}(n, m)$. $f \succeq g$ if and only if $f \succeq_L g$.*

However, the proof given by Müller (2013) is not constructive and cannot be used directly to disentangle the diminishing transfers and correlation-increasing switches. In this paper we will give a constructive proof that can be used to explicitly disentangle the elementary operations.

We will proceed as follows: First we prove that if $f \succeq_L g$, then $f \succeq g$. This proof is split into two parts. The first part shows that we can (always) construct a distribution, h , that (a) still satisfies $h \succeq_L g$, (b) arises from f by exclusively using diminishing transfers, and (c) has the same marginal contributions as g . This part is the main result of the paper. In the second part we give an alternative proof of the result obtained by Tchen (1980) and Decancq (2012), where it is shown that if $f \succeq_L g$ and f and g have the same marginal distributions, then $f \succeq g$. Our two proofs are both constructive, and, as a consequence, they yield a polynomial time complexity algorithm for disentangling lower orthant dominance into elementary operations. For completeness we also provide a short proof for the straightforward direction Proposition 3.5 of Theorem 3.1 (i.e., of the statement “if $f \succeq g$ then $f \succeq_L g$ ”).

3.1 Construction of distribution with identical marginals

In Proposition 3.2 we show that for any f and g with $f \succeq_L g$ we can construct h resulting from f by a sequence of finitely many (possibly zero) diminishing transfers such that h has the same marginal distributions as g and satisfies $h \succeq_L g$.

Proposition 3.2 *Let $n, m \in \mathbb{N}$ and $f, g \in \mathcal{F}(n, m)$ such that $f \succeq_L g$. Then there exists $h \in \mathcal{F}(n, m)$ that arises from f by finitely many diminishing transfers such that $h \succeq_L g$ and for all $i \in \{1, \dots, m\}$ and all $j \in \{1, \dots, n\}$,*

$$\begin{aligned} \tilde{h}(j, m) - \tilde{h}(j, i) &\geq \tilde{g}(j, m) - \tilde{g}(j, i), \\ \tilde{h}(n, i) &= \tilde{g}(n, i) \text{ and} \\ \tilde{h}(j, m) &= \tilde{g}(j, m). \end{aligned}$$

Indeed, the distribution h does not only lower orthant dominate g , but it also satisfies $\tilde{h}(j, m) - \tilde{h}(j, i) \geq \tilde{g}(j, m) - \tilde{g}(j, i)$.

With this intuition we prove the following technical lemma, which is useful for showing Proposition 3.2.

Lemma 3.3 *Let $i_0 \in \{1, \dots, m\}$ and $f \succeq_L g$ such that*

$$\tilde{f}(j, m) - \tilde{f}(j, i) \geq \tilde{g}(j, m) - \tilde{g}(j, i) \quad \forall j \in \{1, \dots, n-1\}, i \in \{i_0, \dots, m\} \text{ and} \quad (3.1)$$

$$\tilde{f}(n, i) = \tilde{g}(n, i) \quad \forall i \in \{i_0, \dots, m\}. \quad (3.2)$$

Then there exists $h \in \mathcal{F}$ that arises from f by finitely many diminishing transfers such that $h \succeq_L g$ and

$$\tilde{h}(j, m) - \tilde{h}(j, i) \geq \tilde{g}(j, m) - \tilde{g}(j, i) \quad \forall j \in \{1, \dots, n-1\}, i \in \{i_0-1, \dots, m\} \text{ and} \quad (3.3)$$

$$\tilde{h}(n, i) = \tilde{g}(n, i) \quad \forall i \in \{i_0-1, \dots, m\}. \quad (3.4)$$

Proof of Lemma 3.3: Due to lower orthant dominance we have $\tilde{g}(n, i_0-1) \geq \tilde{f}(n, i_0-1)$ and $\tilde{g}(n, m) = \tilde{f}(n, m)$,

$$\tilde{f}(n, m) - \tilde{f}(n, i_0-1) \geq \tilde{g}(n, m) - \tilde{g}(n, i_0-1). \quad (3.5)$$

Step 1: We show that there exists h that arises from f by finitely many diminishing transfers such that $h \succeq_L g$ and

$$\tilde{h}(j, m) - \tilde{h}(j, i) \geq \tilde{g}(j, m) - \tilde{g}(j, i) \quad (3.6)$$

for all $j \in \{1, \dots, n\}$ and all $i \in \{i_0 - 1, \dots, m\}$.

Let

$$K(f) = K = \{j \in \{1, \dots, n\} \mid \tilde{f}(j, m) - \tilde{f}(j, i_0 - 1) < \tilde{g}(j, m) - \tilde{g}(j, i_0 - 1)\}$$

and $k = |K|$. We proceed by induction on k . If $k = 0$, then $h = f$ fulfills (3.3). Assume that our statement is correct whenever $k < \ell$ for some $\ell \in \mathbb{N}$. Now, if $k = \ell$, we proceed as follows. Let $j_0 = \max K$. By (3.5), $j_0 < n$. Moreover, because j_0 is maximal, $\alpha := \tilde{f}(j_0 + 1, m) - \tilde{f}(j_0 + 1, i_0 - 1) \geq \tilde{g}(j_0 + 1, m) - \tilde{g}(j_0 + 1, i_0 - 1) := \beta$, we conclude that

$$\begin{aligned} & \left(\alpha - (\tilde{f}(j_0, m) - \tilde{f}(j_0, i_0 - 1)) \right) - \left(\beta - (\tilde{g}(j_0, m) - \tilde{g}(j_0, i_0 - 1)) \right) \\ & \geq (\tilde{g}(j_0, m) - \tilde{g}(j_0, i_0 - 1)) - (\tilde{f}(j_0, m) - \tilde{f}(j_0, i_0 - 1)), \end{aligned}$$

that is,

$$\sum_{i=i_0}^m \left(f(j_0 + 1, i) - g(j_0 + 1, i) \right) \geq (\tilde{g}(j_0, m) - \tilde{g}(j_0, i_0 - 1)) - (\tilde{f}(j_0, m) - \tilde{f}(j_0, i_0 - 1)) =: \varepsilon > 0. \quad (3.7)$$

Let $A = \{i \in \{i_0, \dots, m\} \mid f(j_0 + 1, i) > g(j_0 + 1, i)\}$. There exist $\varepsilon_i \geq 0, i \in A$, such that $\sum_{i \in A} \varepsilon_i = \varepsilon$ and $f(j_0 + 1, i) - \varepsilon_i \geq g(j_0 + 1, i)$ for all $i \in A$. Let h arise from f by transferring ε_i from $(j_0 + 1, i)$ to (j_0, i) for all $i \in A$ (i.e., h arises from f by a sequence of $|A|$ diminishing transfers). As $\tilde{g}(j_0, i_0 - 1) \geq \tilde{f}(j_0, i_0 - 1)$, (3.1) applied to $j = j_0$ yields $\tilde{g}(j_0, i) \geq \tilde{f}(j_0, i) + \varepsilon$ for all $i = i_0, \dots, m$. As $\tilde{h}(j_0, i) - \tilde{f}(j_0, i) \leq \varepsilon$ for $i \in A$ and $\tilde{h}(j, i) = \tilde{f}(j, i)$ for all other pairs (j, i) , we conclude that $h \succeq_L g$. Moreover, $K(h) \subseteq K(f) \setminus \{j_0\}$ by construction. Finally, as only transfers from the right-hand to the left-hand side are employed, (3.6) is still satisfied for all $i \in \{i_0, \dots, m\}$ by (3.1). Hence, by the inductive hypothesis, this step is complete.

Step 2: We now finish the proof by showing (3.4). By Step 1 we may assume that $\tilde{f}(j, m) - \tilde{f}(j, i_0 - 1) \geq \tilde{g}(j, m) - \tilde{g}(j, i_0 - 1)$ for all $j = 1, \dots, n$. Let $\tilde{g}(n, i_0 - 1) - \tilde{f}(n, i_0 - 1) = \rho(f) = \rho \geq 0$. If $\rho = 0$, we may choose $h = f$. Hence, we assume that $\rho > 0$. Let

$$L(f) = L = \{j \in \{1, \dots, n\} \mid f(j, i_0) > g(j, i_0)\},$$

and $\ell = |L|$. Now, $\tilde{f}(n, i_0) = \tilde{g}(n, i_0)$ by (3.2) so that $\rho = \tilde{g}(n, i_0 - 1) - \tilde{g}(n, i_0) - \tilde{f}(n, i_0 - 1) + \tilde{f}(n, i_0) = \sum_{t=1}^n \left(f(t, i_0) - g(t, i_0) \right)$. We conclude that $\ell > 0$ (and, hence, $i_0 > 1$). Assume now that (3.4) is already proven whenever $\ell < r$ for some $r \in \mathbb{N}$. If $\ell = r$, then let $j_1 = \max L$ and $\varepsilon = \min\{f(j_1, i_0) - g(j_1, i_0), \rho\}$. Let h result from f by the diminishing transfer from (j_1, i_0) to $(j_1, i_0 - 1)$ of size ε . Then $h \succ_L g$ and either $\rho(h) = 0$ (if $\varepsilon = \rho(f)$) or $|L(h)| < |L(f)|$. As $\sum_{t=1}^j \left(f(t, i_0) - g(t, i_0) \right) \geq 0$ for all $j \in \{j_1, \dots, n\}$ by construction, (3.6) is still satisfied by (3.1) for all $i \in \{i_0, \dots, m\}$, so that the proof is finished by an inductive argument. **q.e.d.**

Lemma 3.3 reveals the construction of the maximal sequence of diminishing transfers and is used in the following proof.

Proof of Proposition 3.2: The mapping f satisfies the conditions of Lemma 3.3 for $i_0 = m$. Hence, applying the aforementioned lemma successively to $i_0 = m, \dots, 1$ yields a mapping h that arises from f by a finite number of diminishing transfers with $h \succeq_L g$ such that, by (3.3),

$$\tilde{h}(j, m) - \tilde{h}(j, i) \geq \tilde{g}(j, m) - \tilde{g}(j, i) \quad \forall j \in \{1, \dots, n-1\}, i \in \{0, \dots, m\}.$$

Hence, applying this inequality to $i = 0$ yields $\tilde{h}(j, m) \geq \tilde{g}(j, m)$ for all $j \in \{1, \dots, n-1\}$. As $h \succeq_L g$, $\tilde{h}(n, m) = \tilde{g}(n, m)$ and $\tilde{g} \geq \tilde{h}$ so that $\tilde{h}(j, m) = \tilde{g}(j, m)$ for all $j \in \{1, \dots, n\}$. Finally, by (3.4), $\tilde{h}(n, i) = \tilde{g}(n, i)$ for all $i \in \{1, \dots, m\}$. **q.e.d.**

Using Lemma 3.3 allows us to construct a finite sequence of diminishing transfers to obtain the function h of Proposition 3.2. In Algorithm 1 we provide an algorithmic description of how to obtain h . The actual diminishing transfers are given in lines 16 and 27. Observe that the sequence constructed in Lemma 3.3 exists if dominance prevails, and, consequently, if the sequence does not exist, then f does not dominate g . This is exactly what is given in lines 5-8. The complexity of Algorithm 1 to identify the diminishing transfers is $O(n \log n m)$.

Algorithm 1: Construction of distribution with identical marginals

```

1 for  $i_0 = m$  to 1 do
2   Step 1:
3    $K = \{j \in \{1, \dots, n\} \mid \tilde{f}(j, m) - \tilde{f}(j, i_0 - 1) < \tilde{g}(j, m) - \tilde{g}(j, i_0 - 1)\}$ 
4    $j_0 = \max K$ ;
5   if  $j_0 = n$  then
6     Print "There is no dominance";
7     exit;
8   end
9   while  $K \neq \emptyset$  do
10     $A = \{i \in \{i_0, \dots, m\} \mid f(j_0 + 1, i) > g(j_0 + 1, i)\}$ ;
11     $\varepsilon = \tilde{g}(j_0, m) - \tilde{g}(j_0, i_0 - 1) - (\tilde{f}(j_0, m) - \tilde{f}(j_0, i_0 - 1))$ ;
12    for  $i = 1$  to  $m$  do
13       $\varepsilon_i = \min\{f(j_0 + 1, i) - g(j_0 + 1, i), \varepsilon - \sum_{t=i_0}^m \varepsilon_t\}$ ;
14    end
15    foreach  $i \in A$  do
16      Transfer  $\varepsilon_i$  from  $(j_0 + 1, i)$  to  $(j_0, i)$ ;
17    end
18    Remove  $j_0$  from  $K$ ;
19     $j_0 = \max K$ ;
20  end
21  Step 2:
22   $\rho = \tilde{g}(n, i_0 - 1) - \tilde{h}(n, i_0 - 1)$ ;
23   $L = \{j \in \{1, \dots, n\} \mid \sum_{t=1}^j f(t, i_0) > \sum_{t=1}^j g(t, i_0)\}$ ;
24  while  $L \neq \emptyset$  do
25     $j_1 = \max\{j \in L \mid f(j, i_0) > g(j, i_0)\}$ ;
26     $\varepsilon = \min\{f(j_1, i_0) - g(j_1, i_0), \rho\}$ ;
27    Transfer  $\varepsilon$  from  $(j_1, i_0)$  to  $(j_1, i_0 - 1)$ ;
28     $\rho = \rho - \varepsilon$ ;
29  end
30 end

```

3.2 Using correlation-increasing switches to obtain the distribution

So far, we have shown that if $f \succeq_L g$ we can obtain a function h from f by finitely many diminishing transfers that still satisfies $h \succeq_L g$ and, additionally, has the same marginal distributions as g . Then (1980) and Decancq (2012) show that g can be obtained from an arbitrary h by correlation-increasing switches if and only if g and h have coinciding marginal distributions and $h \succeq_L g$. To enable a

description of our algorithm and in order to make the paper self-contained we include the following short proof that proceeds by induction on m and may be illustrated as follows. We construct correlation-increasing switches that successively create a new function $h \succeq_L g$ such that $h(j, 1) = g(j, 1)$ for all $j \in \{1, \dots, n\}$, so that we may delete the first row and apply an inductive argument. Within this row we proceed recursively by first selecting the maximal j with $f(j, 1) < g(j, 1)$, which we denote by j_0 , and secondly selecting the minimal $j > j_0$ (denoted by j_1) such that when we subtract the aggregate weights of $\{j_0, \dots, j\}$ of g from f , this difference is nonnegative. We then move the difference of the aggregate weights of $\{j_0, \dots, j_1 - 1\}$ via correlation-increasing switches from $(j_1, 1)$ to elements $(j, 1)$ with $j_0 \leq j < j_1$, so that either j_0 or j_1 becomes smaller. For the precise definition of these correlation-increasing switches, we refer to the proof.

The following notation is useful. For $f \in \mathcal{F}$ denote

$$\mathcal{F}_f = \{g \in \mathcal{F} \mid f \succeq_L g, \tilde{g}(j, m) = \tilde{f}(j, m), \tilde{g}(n, i) = \tilde{f}(n, i) \forall i \in \{1, \dots, m\}, j \in \{1, \dots, n\}\}.$$

This means that \mathcal{F}_f contains all lower orthant dominated distributions having marginal distributions coinciding with the marginal distributions of f . Let $g \in \mathcal{F}_f$ and $f \succeq h$. Then $h \succeq g$ if and only if h arises from f by a sequence of finitely many correlation-increasing switches and $g \in \mathcal{F}_h$. Moreover, denote

$$j_0(f, g) = j_0 = \max\{j \in \{1, \dots, n\} \mid f(j, 1) < g(j, 1)\},$$

where $\max \emptyset = 0$ is used in this context, and

$$j_1(f, g) = j_1 = \min \left\{ j \in \{j_0 + 1, \dots, n\} \mid \sum_{t=j_0}^j (f(t, 1) - g(t, 1)) \geq 0 \right\}.$$

Note that j_1 exists because $\tilde{g}(j_0 - 1, 1) \geq \tilde{f}(j_0 - 1, 1)$ and $\tilde{g}(n, 1) = \tilde{f}(n, 1)$. Also note that if $j_0 = 0$, then $f(j, 1) = g(j, 1)$ for all $j \in \{1, \dots, n\}$. The following technical lemma is useful.

Lemma 3.4 *Let $g \in \mathcal{F}_f$ such that $j_0(f, g) > 0$. Then there exists $f \triangleright h$ such that $h \succeq_L g$ and $j_0(h, g) < j_0(f, g)$.*

Proof: Let $j_0 = j_0(f, g)$, $j_1 = j_1(f, g)$, and denote

$$X'(f, g) = X' = \{(j, i) \in X \mid j_0 \leq j < j_1, i > 1, f(j, i) > g(j, i)\}.$$

Thus $X'(f, g)$ contains all $(j, i) \in X$ in the interval between j_0 and j_1 where it is feasible to move mass into, since f has more mass here than g .

Let $\varepsilon = \sum_{t=j_0}^{j_1-1} (g(t, 1) - f(t, 1))$. Note that, by the definitions of j_0 and j_1 , $j_0 > 0$ implies $\varepsilon > 0$. Hence, $f(j_1, 1) - g(j_1, 1) \geq \varepsilon$. As $\tilde{f}(j, m) = \tilde{g}(j, m)$ in particular for $j = j_0 - 1$ and $j = j_1 - 1$, the equations

$$\begin{aligned} \tilde{f}(j_1 - 1, m) &= \tilde{f}(j_0 - 1, m) + \sum_{j=j_0}^{j_1-1} f(j, 1) + \sum_{i=2}^m \sum_{j=j_0}^{j_1-1} f(j, i) \text{ and} \\ \tilde{g}(j_1 - 1, m) &= \tilde{g}(j_0 - 1, m) + \sum_{j=j_0}^{j_1-1} g(j, 1) + \sum_{i=2}^m \sum_{j=j_0}^{j_1-1} g(j, i), \end{aligned}$$

imply that $\sum_{i=2}^m \sum_{j=j_0}^{j_1-1} (f(j, i) - g(j, i)) = \varepsilon$, hence $X'(f, g) \neq \emptyset$. By a recursive argument it suffices to construct h such that $f \triangleright h \succeq_L g$ and either

- $j_0(h, g) < j_0$ or
- $j_0(h, g) = j_0$ and $j_1(h, g) < j_1$ or
- $j_0(h, g) = j_0, j_1(h, g) = j_1$ and $|X'(h, g)| < |X'|$.

For this purpose let $i' \in \{2, \dots, m\}$ be minimal such that there exists $(j, i') \in X'$. Moreover let $j' \in \{j_0, \dots, j_1 - 1\}$ be maximal such that $(j', i') \in X'$. Now we define $\varepsilon' = \min\{\varepsilon, f(j', i') - g(j', i')\}$ and consider $h = f_{\varepsilon'}^{(j_1, 1) \leftrightarrow (j', i')}$ and verify that $h \succeq_L g$. If $\varepsilon' = \varepsilon$, then we have moved ε to the element $(j', 1)$ so that $\sum_{j=j_0}^{j_1-1} h(j, 1) = \sum_{j=j_0}^{j_1-1} f(j, 1) + \varepsilon = \sum_{j=j_0}^{j_1-1} g(j, 1)$. If $j' = j_0$ and $\varepsilon = g(j_0, 1) - f(j_0, 1)$, we have $j_0(h, g) < j_0$. If $j' \neq j_0$ or $\varepsilon < g(j_0, 1) - f(j_0, 1)$, we have $j_0(h, g) = j_0$ and $j_1(h, g) < j_1$. Finally, if $\varepsilon' < \varepsilon$, then we have $j_0(h, g) = j_0, j_1(h, g) = j_1$, and $X'(h, g) = X'(f, g) \setminus \{(j', i')\}$ so that the proof is complete. **q.e.d.**

We apply Lemma 3.4 iteratively to obtain Algorithm 2, which either constructs a finite sequence of correlation-increasing switches or shows that the sequence does not exist. Recall that correlation-increasing switches preserve the marginal distribution, and as a consequence, if we cannot find the distribution h from Lemma 3.4, then f does not lower orthant dominate g . The complexity of Algorithm 2 to identify the correlation-increasing switches is $O(n^2m^2)$.

Algorithm 2: Sequence of Correlation-Increasing Switches

```

1 for  $i_0 = 1$  to  $m - 1$  do
2   while  $f(j, i_0) \neq g(j, i_0) \forall j \in \{1, \dots, n\}$  do
3      $j_0 = \max\{j \in \{1, \dots, n\} \mid f(j, i_0) < g(j, i_0)\}$ ;
4      $j_1 = \min\{j \in \{j_0 + 1, \dots, n\} \mid \sum_{t=j_0}^j (f(t, i_0) - g(t, i_0)) \geq 0\}$ ;
5     if  $j_1$  does not exist then
6       Print "There is no dominance";
7       exit;
8     end
9      $\varepsilon = \sum_{t=j_0}^{j_1-1} (g(t, i_0) - f(t, i_0))$ ;
10    while  $\varepsilon > 0$  do
11       $X' = \{(j, i) \in X \mid j_0 \leq j < j_1, i > i_0, f(j, i) > g(j, i)\}$ ;
12       $i' = \min\{i \in \{i_0 + 1, \dots, m\} \mid \exists (j, i') \in X'\}$ ;
13       $j' = \max\{j \in \{j_0, \dots, j_1 - 1\} \mid \exists (j', i') \in X'\}$ ;
14       $\varepsilon' = \min\{\varepsilon, f(j', i') - g(j', i')\}$ ;
15      Transfer  $\varepsilon'$  from  $(j_1, i_0)$  to  $(j', i_0)$  and  $\varepsilon'$  from  $(j', i')$  to  $(j_1, i')$ ;
16       $\varepsilon = \varepsilon - \varepsilon'$ ;
17    end
18  end
19 end

```

3.3 Equivalence of dominance concepts

We now collect the results from the previous sections to prove the two directions of Theorem 3.1 separately. We start with the straightforward direction.

Proposition 3.5 *Let $n, m \in \mathbb{N}$ and $f, g \in \mathcal{F}(n, m)$. If $f \succeq g$, then $f \succeq_L g$.*

Proof: Let $f, g \in X(n, m)$. If g results from f by a diminishing bilateral transfer from y to x or by a correlation-increasing switch from x to $\min\{x, y\}$ and y to $\max\{x, y\}$, then $\tilde{g}(n, m) = \tilde{f}(n, m)$ and it is straightforward to show that $\tilde{g} \geq \tilde{f}$. Therefore, Proposition 3.5 follows by induction on k , the number of functions $f_1, \dots, f_k \in \mathcal{F}$ such that $f_1 = f$, $f_k = g$, and f_ℓ results from $f_{\ell-1}$ by a diminishing transfer or a correlation-increasing switch for each $\ell \in \{2, \dots, k\}$. **q.e.d.**

Proposition 3.6 *Let $n, m \in \mathbb{N}$ and $f, g \in \mathcal{F}(n, m)$. If $f \succeq_L g$, then $f \supseteq g$.*

Proof: By Corollary 3.2 we may assume that $g \in \mathcal{F}_f$. Successively applying Lemma 3.4 if necessary we may also assume that $j_0(f, g) = 0$, i.e., $f(j, 1) = g(j, 1)$ for all $j \in \{1, \dots, n\}$ so that the proof is finished by induction on m . **q.e.d.**

Proof of Theorem 3.1: By combining Proposition 3.5 and Proposition 3.6 the proof of our main result follows. **q.e.d.**

As mentioned earlier, our constructive proof yields an algorithm. We need to follow each step in the proof, and this means that we have to make sets and define variables as described in the proof. There is only one place where the proof does not give us unique transfers, namely in the first step in the diminishing transfer part. Here, we can decide in the algorithm in which order transfers should be made. One possibility is to choose the lexicographical order. When we follow the algorithm it is possible to track every transfer made. The complexity of the entire algorithm will thus be $O(n^2 m^2)$, as Algorithm 2 has the largest leading term.

4 Example

To illustrate the algorithm we give a 4×4 example. Let f and g be given by Figures 4.1 and 4.2.

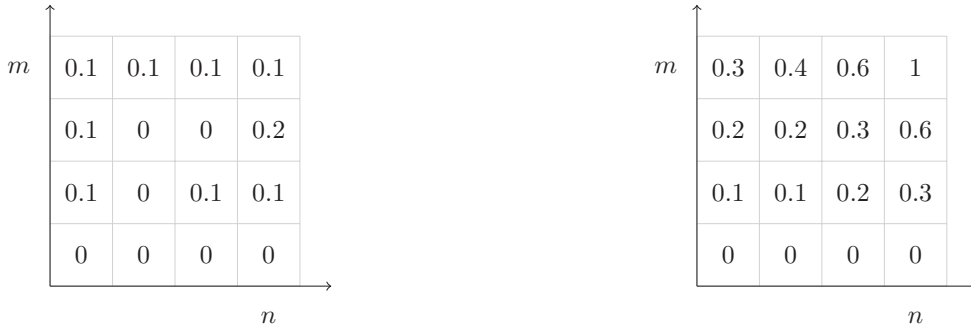


Figure 4.1: The probability mass function f (left) and \tilde{f} (right)

The first part of the algorithm employs diminishing transfers to obtain equal marginal distributions. The computation of sets and values appears from the proof. An overview of the algorithm may also be obtained from inspecting the pseudo codes of Algorithm 1 (see Section 3.1) and Algorithm 2 (see Section 3.2). As an example for $i_0 = 4$ we find the set K of indices j where $\tilde{f}(j, 4) - \tilde{f}(j, 3) < \tilde{g}(j, 4) - \tilde{g}(j, 3)$. This is only true for $j = 1$, where $\tilde{f}(1, 4) - \tilde{f}(1, 3) = 0.3 - 0.2 = 0.1$ and $\tilde{g}(1, 4) - \tilde{g}(1, 3) = 0.4 - 0.2 = 0.2$. As we only have one element in K this will also be the largest element, so $j_0 = 1$. Now we know which column we need to move mass to, so now we need to check what row, $i \in \{i_0, \dots, m\}$, we can move mass from $(j_0 + 1, i)$, i.e., which row index i satisfies $f(j_0 + 1, i) < g(j_0 + 1, i)$. Since $i_0 = 4$ we have $i = 4$,

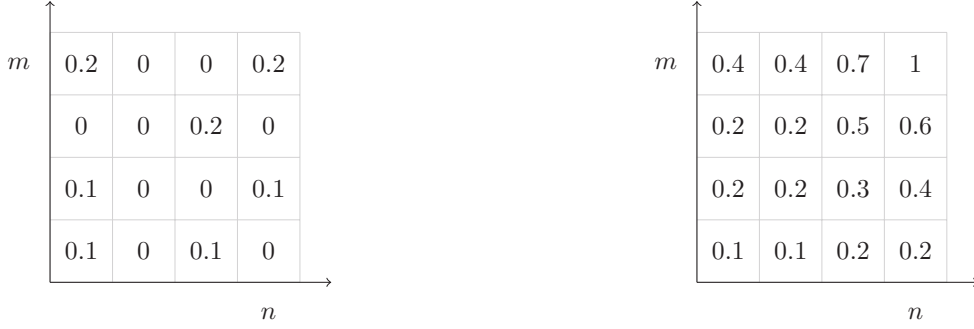


Figure 4.2: The probability mass function g (left) and \tilde{g} (right)

i.e., $A = \{4\}$. Now we find $\varepsilon = \tilde{g}(1, 4) - \tilde{g}(1, 3) - (\tilde{f}(1, 4) - \tilde{f}(1, 3)) = 0.4 - 0.2 - (0.3 - 0.2) = 0.1$ that has to be distributed to all row indices of A . Since A is a singleton this task is simple. In the case that A contains more elements there might be more ways to move the entire mass. In the example, we move as much mass as possible for the lowest indices, and stop when the entire ε is moved. This gives us a diminishing transfer in the horizontal direction. We then check whether transfers in the vertical direction are necessary, i.e., transfers that ensure that $\tilde{f}(n, i_0 - 1) = \tilde{g}(n, i_0 - 1)$. Since $\tilde{f}(4, 3) = 0.6 = \tilde{g}(4, 3)$, no such transfers are needed. For $i_0 = 3$, $K = \emptyset$, but $\tilde{f}(4, 2) = 0.3 \neq 0.4 = \tilde{g}(4, 2)$ so that that $L \neq \emptyset$. We find L as the set of indices j where $\sum_{t=1}^j \tilde{f}(t, i_0) > \sum_{t=1}^j \tilde{g}(t, i_0)$ and conclude that $L = \{1, 2, 4\}$ since $\tilde{f}(1, 3) = 0.1 > 0 = \tilde{g}(1, 3)$, $\sum_{t=1}^3 \tilde{g}(t, 3) = 0.2 > 0.1 = \sum_{t=1}^3 \tilde{f}(t, 3) = \sum_{t=1}^2 \tilde{f}(t, 3) > 0 = \sum_{t=1}^2 \tilde{g}(t, 3)$, and $\sum_{t=1}^4 \tilde{f}(t, 3) = 0.3 > 0.2 = \sum_{t=1}^4 \tilde{g}(t, 3)$. We now find $\rho = \tilde{g}(4, 2) - \tilde{f}(4, 2) = 0.4 - 0.3 = 0.1$ and j_1 which is the largest index j where $\tilde{f}(j, 3) > \tilde{g}(j, 3)$ is true, $\tilde{f}(4, 3) = 0.2 > 0 = \tilde{g}(4, 3)$ so $j_1 = 4$. We are then able to compute $\varepsilon = \min\{\tilde{f}(4, 3) - \tilde{g}(4, 3), \rho\} = \min\{0.2, 0.1\} = 0.1$ and transfer this amount. Now $\rho = \rho - \varepsilon = 0.1 - 0.1 = 0$. If $\rho > 0$ after the transfer, we would continue with a new j_1 and find the new ε . In the table the steps are also shown for $i_0 = 2$ and $i_0 = 1$.

$i_0 = 4$	$K = \{1\}, j_0 = 1, A = \{4\}, \varepsilon = 0.1 = \varepsilon_4$ $L = \emptyset$	Diminishing transfer of 0.1 from (2, 4) to (1, 4)
$i_0 = 3$	$K = \emptyset$ $L = \{1, 2, 4\}, \rho = 0.1, j_1 = 4, \varepsilon = 0.1$	Diminishing transfer of 0.1 from (4, 3) to (4, 2)
$i_0 = 2$	$K = \emptyset$ $L = \{3, 4\}, \rho = 0.2, j_1 = 4, \varepsilon = 0.1$ $L = \{3\}, \rho = 0.1, j_1 = 3, \varepsilon = 0.1$	Diminishing transfer of 0.1 from (4, 2) to (4, 1) Diminishing transfer of 0.1 from (3, 2) to (3, 1)
$i_0 = 1$	$K = \{3\}, j_0 = 3, A = \{1, 3\}, \varepsilon = 0.1 = \varepsilon_1$	Diminishing transfer of 0.1 from (4, 1) to (3, 1)

After having applied the diminishing transfers, the arising new f and \tilde{f} are given by Figure 4.3. When we have secured that the marginal distributions are equal we will continue with the second part where we employ correlation increasing switches in order to obtain identical probability mass functions. Here $i_0 = i$ means that the row i from the full probability mass functions is the bottom row in the present step. This means that in our 4x4-example having $i_0 = 4$ indeed means that we only consider the uppermost row from the full probability mass functions, because the rest are already identical. We start with $i_0 = 1$. Here we find $j_0 = \max\{j \in \{1, \dots, n\} | f(j, 1) < g(j, 1)\}$. We have that $f(1, 1) = 0 < 0.1 = g(1, 1)$, $f(2, 1) = 0 = g(2, 1)$, $f(3, 1) = 0.2 > 0.1 = g(3, 1)$ and $f(4, 1) = 0 = g(4, 1)$, which means that $j_0 = 1$. Then we



Figure 4.3: The new probability mass function f (left) and \tilde{f} (right)

find $j_1 = \min\{j \in \{j_0 + 1, \dots, n\} \mid \sum_{t=j_0}^j (f(t, 1) - g(t, 1)) \geq 0\}$, since $\sum_{t=1}^2 (f(t, 1) - g(t, 1)) = -0.1$ and $\sum_{t=1}^3 (f(t, 1) - g(t, 1)) = 0$ we have that $j_1 = 3$. We have to compute the set $X' = \{(j, i) \in X \mid j_0 \leq j < j_1, i > i_0, f(j, i) > g(j, i)\}$ this gives us $X' = \{(1, 3)\}$. This means that when we have to chose i' minimal such that $(j, i') \in X'$ and j' maximal such that $(j', i') \in X'$ we will have $(j', i') = (3, 1)$. We can compute $\varepsilon = \sum_{t=j_0}^{j_1-1} (g(t, i_0) - f(t, i_0)) = (g(1, 1) - f(1, 1)) + (g(2, 1) - f(2, 1)) = (0.1 - 0) + (0 - 0) = 0.1$ and $\varepsilon' = \min\{\varepsilon, f(j', i') - g(j', i')\} = \min\{0.1, f(1, 3) - g(1, 3)\} = \min\{0.1, 0.1 - 0\} = 0.1$. So now we have identified a correlation increasing switch. Since the row $i = 1$ is identical in f and g we are done with this row. So we can discard it from the algorithm and look on the remaining rows. In the table the rows $i_0 = 2$ and $i_0 = 4$ are blank because each of these rows already coincides with the row corresponding to g when the algorithm for $i_0 = 1$ and $i_0 = 3$, respectively, has been applied already. Hence, it is not necessary to employ any further correlation increasing switches.

$i_0 = 1$	$j_0 = 1, j_1 = 3, i' = 3, j' = 1, \varepsilon = 0.1, \varepsilon' = 0.1$ Correlation increasing switch of 0.1 from (3, 1) to (1, 1) and from (1, 3) to (3, 3)
$i_0 = 2$	
$i_0 = 3$	$j_0 = 3, j_1 = 4, i' = 4, j' = 3, \varepsilon = 0.1, \varepsilon' = 0.1$ Correlation increasing switch of 0.1 from (4, 3) to (3, 3) and from (3, 4) to (4, 4)
$i_0 = 4$	

5 Remarks on the multivariate case

We now consider $d \in \mathbb{N}$, $d > 2$, $n_1, \dots, n_d \in \mathbb{N}$, the discrete hyper-rectangle $X = \{x \in \mathbb{N}^d \mid x \leq (n_d, \dots, n_1)\}$, and $\mathcal{F} = \{f : X \rightarrow \mathbb{R}_+\}$. The expressions lower orthant dominance, diminishing transfer, and correlation increasing switch are easily generalized to this multivariate case by literally copying their formal definitions provided in Section 2. It is straightforward to generalize the proof of Proposition 3.5 to obtain that $f \succeq g$ implies $f \succeq_L g$. Moreover, with the help of an example of Müller and Scarsini (2000) it may be shown that the opposite implication does not hold, hence our main theorem does not hold in the case $d > 2$.

Indeed, if the dimension is higher than 2, further elementary transfers have to be considered. Müller's (2013) general equivalence result employs so-called Δ -antitone transfers and shows that $f \succeq_L g$ if and

only if g can be obtained from f by a finite sequence of Δ -antitone k -transfers, where $k = 1, \dots, d$. Now, Δ -antitone 1-transfers and Δ -antitone 2-transfers are the diminishing transfers and correlation-increasing switches, respectively. Moreover, Decancq’s (2012) result applies to the multivariate case: If f and g have the same marginal distributions, then $f \succeq_L g$ if and only if g arises from f by a finite sequence of Δ -antitone k -transfers where $k = 2, \dots, d$. As our current approach does not aim to present an algorithm that uses the aforementioned equivalence result, we do not mention further details, e.g., the definition of Δ -antitone k -transfers for $k > 2$, in the present paper.

6 Final remarks

As mentioned earlier, Theorem 3.1 was recently proven by Meyer and Strulovici (2010, 2015) and Müller (2013). However, our approach is different. In particular, our proof is constructive and yields an algorithm. The most important feature is that it allows us to decompose domination into diminishing transfers and correlation-increasing switches. In other words, we can disentangle domination into welfare deteriorations and inequality increases.

The algorithm is useful in several ways. First, the decomposition technique enables us to compare lower orthant dominance to more restrictive stochastic dominance concepts, for example to first order dominance, which allows only diminishing transfers, or to increasing interdependence (as in Tchen 1980, Epstein and Tanny 1980, and Decancq 2012), allowing only correlation-increasing switches. Second, if we measure the distance between two distributions by the amount of mass moved times the ℓ_1 distance it is moved, then the algorithm tells us how much the distributions differ, and, in particular, we can decompose the difference to see how much mass is moved by diminishing transfers. Although the sequence of diminishing transfers and correlation-increasing switches leading from one distribution to another is generally not uniquely determined, the mass we have to move with diminishing transfers to obtain the same marginal distributions is always the same.⁸ Thus, in terms of the measure of distance of mass displacement through diminishing transfers, the welfare differences are uniquely determined and hence meaningfully measured for comparison purposes.

References

- AABERGE, R., AND A. BRANDOLINI (2014): “Multidimensional poverty and inequality”, in *Handbook of Income Distribution*, ed. by A. B. Atkinson and F. Bourguignon, Elsevier, vol. 2, 141–216.
- ARNDT, C., R. DISTANTE, M. A. HUSSAIN, L. P. ØSTERDAL, P. L. HUONG, AND M. IBRAIMO (2012): “Ordinal welfare comparisons with multiple discrete indicators: A first order dominance approach and application to child poverty”, *World Development*, 40, 2290–2301.
- ATKINSON, A., AND F. BOURGUIGNON (1982): “The comparison of multi-dimensioned distributions of economic status”, *Review of Economic Studies*, 49, 183–201.

⁸Consider distributions f , g , and h . Suppose that g can be obtained from f by a sequence of diminishing transfers, and similarly, h can be obtained from f by (another) sequence of diminishing transfers. Moreover, assume that g and h have identical partial marginal distributions. We have to prove that the transfer system leading from g to f “moves the same amount of mass” as the system obtaining h from f . First, realize that the result is true for the case that $m = 1$ (i.e., the one-dimensional case) because the partial marginal distributions uniquely determine the distribution. If $m > 1$, then we know exactly how much mass has to be moved from right to left. Similarly, by the one-dimensional case, we also know exactly how much mass has to be moved down. Therefore, we know exactly how much mass to move.

- BOURGUIGNON, F., AND S. R. CHAKRAVARTY (2003): “The measurement of multidimensional poverty”, *The Journal of Economic Inequality*, 1, 25–49.
- CHAKRAVARTY, S. R., AND C. ZOLI (2012): “Stochastic dominance relations for integer variables”, *Journal of Economic Theory*, 147, 1331–1341.
- DECANCO, K. (2012): “Elementary multivariate rearrangements and stochastic dominance on a Fréchet class”, *Journal of Economic Theory*, 147, 1450–1459.
- DUCLOS, J.-Y., D. E. SAHN, AND S. D. YOUNGER (2006): “Robust multidimensional poverty comparisons”, *The Economic Journal*, 116, 943–968.
- (2007): “Robust multidimensional poverty comparisons with discrete indicators of well-being”, in *Inequality and poverty re-examined*, ed. by S. P. Jenkins, and J. Micklewright. Oxford University Press.
- EPSTEIN, L. G., AND TANNY, S. M. (1980): “Increasing generalized correlation: A definition and some economic consequences”, *Canadian Journal of Economics*, 13, 16–34.
- GRAVEL, N., P. MOYES, AND B. TARROUX (2009): “Robust international comparisons of distributions of disposable income and regional public goods”, *Economica*, 76, 432–461.
- GRAVEL, N., AND A. MUKHOPADHYAY (2010): “Is India better off today than 15 years ago? A robust multidimensional answer”, *The Journal of Economic Inequality*, 8, 173–195.
- KAMAE, T., U. KRENGEL, AND G. O’BRIEN (1977): “Stochastic inequalities on partially ordered spaces”, *The Annals of Probability*, 5, 899–912.
- LAMBERT, P. (2001): *The Distribution and Redistribution of Income*. 3rd edition, Manchester University Press, Manchester.
- LEVY, H. (1992): “Stochastic dominance and expected utility: Survey and analysis”, *Management Science*, 38, 555–593.
- MEYER, M., AND B. STRULOVICI (2011): “The supermodular stochastic ordering”, *Unpublished*.
- (2015): “Beyond correlation: Measuring interdependence through complementarities”, *Unpublished*.
- MOYES, P. (2012): “Comparisons of heterogeneous distributions and dominance criteria”, *Journal of Economic Theory*, 147, 1351–1383.
- MÜLLER, A. (2013): “Duality theory and transfers for stochastic order relations”, in *Stochastic Orders in Reliability and Risk in Honor of Professor Moshe Shaked*, ed. by H. Li and X. Li, Lecture Notes in Statistics 208, pp. 41 – 57, New York. Springer.
- MÜLLER, A., AND M. SCARSINI (2000): “Some remarks on the supermodular order”, *Journal of Multivariate Analysis*, 73, 107–119.
- MÜLLER, A., AND D. STOYAN (2002): *Comparison Methods for Stochastic Models and Risks*. John Wiley and Sons.
- MULLER, C., AND A. TRANNOY (2011): “A dominance approach to the appraisal of the distribution of well-being across countries”, *Journal of Public Economics*, 95, 239–246.

- RANGE, T. M., AND L. P. ØSTERDAL (2017): “First order dominance: Stronger characterization and bivariate checking algorithm”, *Mathematical Programming*, published online at <https://doi.org/10.1007/s10107-017-1213-9>.
- SHAKED, M., AND J. SHANTHIKUMAR (2007): *Stochastic Orders*. Springer.
- SILVAPULLE, M. J., AND P. K. SEN (2011): *Constrained Statistical Inference: Order, Inequality, and Shape Constraints*. Vol. 912. Wiley Series in Probability and Statistics.
- SONNE-SCHMIDT, C., F. TARP, AND L. P. ØSTERDAL (2016): “Ordinal bivariate inequality: Concepts and application to child deprivation in Mozambique”, *Review of Income and Wealth*, 62, 559–573.
- SRIBOONCHITA, S., S. DHOMPONGSA, W. K. WONG, AND H. T. NGUYEN (2009): *Stochastic Dominance and Applications to Finance, Risk and Economics*. Chapman & Hall/CRC.
- STRASSEN, V. (1965): “The existence of probability measures with given marginals”, *The Annals of Mathematical Statistics*, 36, 423–439.
- TCHEN, A. H. (1980): “Inequalities for distributions with given marginals”, *The Annals of Probability*, pp. 814–827.
- TSUI, K.-Y. (1999): “Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation”, *Social Choice and Welfare*, 16, 145–157.
- ØSTERDAL, L. P. (2010): “The mass transfer approach to multivariate discrete first order stochastic dominance: Direct proof and implications”, *Journal of Mathematical Economics*, 46, 1222–1228.