Checking Bivariate First Order Dominance

by

Troels Martin Range

and

Lars Peter Østerdal

Discussion Papers on Business and Economics
No. 9/2013

FURTHER INFORMATION
Department of Business and Economics
Faculty of Social Sciences
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark

Tel.: +45 6550 3271
Fax: +45 6550 3237
E-mail: lho@sam.sdu.dk
http://www.sdu.dk/ivoe

ISBN 978-87-91657-87-0

Checking Bivariate First Order Dominance

Troels Martin Range and Lars Peter Østerdal

Department of Business and Economics, and COHERE, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark

May 27, 2013

Abstract

We consider the problem of checking first order dominance for finite bivariate distributions. We observe that this can be formulated as a special bipartite network problem related to the classical transportation problem. We exploit this observation to develop a new characterization of first order dominance and fast dominance-checking algorithms.

Keywords: Multidimensional first order dominance, usual stochastic order, characterization, network problem, checking algorithm.
JEL Code: C61, D63, I31

1 Introduction

Dominance comparisons of distributions are a methodological cornerstone in economics, finance, probability theory, and statistics, among other fields. In welfare economics, for instance, dominance concepts are used to partially order population distributions according to better social welfare or less inequality (e.g. Atkinson and Bourguignon (1982), Gravel and Moyes (2012)), in decision analysis and finance stochastic orderings are used for evaluating risky assets (e.g. Sriboonchita et al. (2009)), and in statistics various order restrictions can form part of a null or alternative hypothesis (e.g. Silvapulle and Sen (2011)). For a general discussion of stochastic dominance theory we refer to Marshall and Olkin (1979), Müller and Stoyan (2002), and Shaked and Shanthikumar (2007).

The canonical stochastic dominance concept is that of *first order dominance*, also known as the *usual (stochastic) order* (Lehmann (1955)). First order dominance captures the intuition that one (dominant) distribution is better, i.e. gives higher outcomes, than the other (dominated) distribution. For two multidimensional finite distributions, f and g, f

first order dominates distribution g if and only if one of the following three (equivalent) conditions hold: (a) it is possible to obtain g from f by moving probability mass from better to worse outcomes, (b) the cumulative probability mass at f is smaller than or equal to that at g for every comprehensive subset of outcomes[1], and (c) the expected utility of f is as least as high as that of g for any non-decreasing utility function.[2] Thus, first order dominance is an ordinal concept that does not rely on assumptions about the relative importance of dimensions or the complementarity/substitutability relationships between dimensions (Arndt et al. (2012)).[3]

In the one-dimensional case, much research into the nature of first order dominance has been conducted (for example, the bibliography by Bawa (1982) contains more than 400 references), and the theory is by now well developed and the applications many. Surprisingly, perhaps, there have to date only been few empirical applications of the first order dominance concept to cases with two or more dimensions. An important reason might be that checking multivariate first order dominance is not so easy with existing methods (unless the total number of outcomes is small), and there is a gap in research on how this can be done efficiently. The most direct way to check first order dominance is to use definition (b). Here one needs to check an inequality for each comprehensive subset of outcomes. However, the number of inequalities to be tested grows dramatically in the total number of outcomes, so it is not an efficient method. Mosler and Scarsini (1991) and Dyckerhoff and Mosler (1997) describe a method based on linear programming for checking first order dominance in the general multivariate finite case, based on definition (a) above. To our knowledge, the first empirical implementations of a method along these lines appear in Arndt et al. (2012) and Arndt et al. (2013). An alternative approach would be to make use of a network flow formulation of the problem, as outlined in Preston (1974) or Hansel and Troallic (1978), and then check for dominance via computation of the maximum flow. We are not aware of any actual implementations of such a method for checking multivariate first order dominance.

In this paper we present two algorithms for identifying first order dominance in the finite bivariate case. The first algorithm is an intuitive and constructive approach for testing first order dominance by identifying a finite sequence of diminishing bilateral transfers or showing that no such sequence exists in $O(n^2)$ time, where n is the number of outcomes. The second algorithm is indirect. It either states that first order dominance exists or shows

[1] A comprehensive subset holds the property that if an outcome is in the subset, then all smaller outcomes are also included in that subset.

[2] Less restrictive dominance criteria for better distributions have been defined by imposing stronger restrictions on the set of admissible utility functions. See, e.g., Levy and Paroush (1974), Harder and Russell (1974), Huang et al. (1978), Atkinson and Bourguignon (1982), Mosler (1984), Russell and Seo (1978), and Scarsini (1988).

[3] In the multidimensional context the first order dominance concept has been used with other meanings than the one given here. In particular, Atkinson and Bourguignon (1982) and others have used the term "first order dominance" to denote a less restrictive stochastic dominance concept (also known as an orthant stochastic order cf. e.g. Dyckerhoff and Mosler (1997)) suitable under a substitutability relationship between the dimensions.

that a comprehensive set violates (b). We use the first (direct) algorithm to prove the correctness of the indirect algorithm in the case where first order dominance exists. The second algorithm has $O(n)$ worst case complexity.

The paper is organized as follows. First, in section 2 the necessary notation and basic definitions are introduced. Next, in section 3 we give some basic insights along with a new characterization of first order dominance for the general multivariate case. In section 4 we turn to algorithms for the bivariate case. Section 5 concludes.

2 Notation and Definitions

An outcome is a vector $\mathbf{x} = (x_1, \ldots, x_m)$, where each attribute x_j is from an attribute set $X_j = \{1, 2, ..., n_j\}$, $j = 1, \ldots, m$, and $m \geq 2$. The outcome set is the product set $X = X_1 \times \ldots \times X_m$ and has cardinality $n = \Pi_{j=1}^{m} n_j$. If $m = 2$, then we have a bivariate case. For any two elements, $\mathbf{x}, \mathbf{y} \in X$, we define $\mathbf{y} \leq \mathbf{x}$ such that $y_j \leq x_j$ for all j, and $\mathbf{y} < \mathbf{x}$ such that $x_j \leq y_j$ for all j and $\mathbf{y} \neq \mathbf{x}$. A set $Y \subseteq X$ is called *comprehensive* if $\mathbf{x} \in Y$, $\mathbf{y} \in X$, and $\mathbf{y} \leq \mathbf{x}$ imply $\mathbf{y} \in Y$.

A *distribution* is a real-valued function f on X, such that $f(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in X$ and $\sum_{\mathbf{x} \in X} f(\mathbf{x}) = 1$. We say that a distribution g can be derived from a distribution f by a *bilateral transfer* (of probability mass) if there are outcomes $\mathbf{x}, \mathbf{y}$ such that $g(\mathbf{z}) = f(\mathbf{z})$ for $\mathbf{z} \neq \mathbf{x}, \mathbf{y} \in X$. Note that if the two distributions are identical except for the values in $\mathbf{x}$ and $\mathbf{y}$, then we can obtain g from f by transferring a suitable amount of probability mass between $\mathbf{x}$ and $\mathbf{y}$.[4] A *diminishing* bilateral transfer is a shift of probability mass from one outcome, $\mathbf{x}$, to another, $\mathbf{y}$, such that $\mathbf{y} < \mathbf{x}$.

Suppose that f and g denote two distribution functions. We say that f *first order dominates* g if one of the following three equivalent properties (A)-(C) hold.[5]

(A) g can be obtained from f by a finite number of diminishing bilateral probability mass transfers.

(B) $\sum_{\mathbf{x} \in Y} g(\mathbf{x}) \geq \sum_{\mathbf{x} \in Y} f(\mathbf{x})$ for any comprehensive set $Y \subseteq X$.

(C) $\sum_{\mathbf{x} \in X} u(\mathbf{x}) f(\mathbf{x}) \geq \sum_{\mathbf{x} \in X} u(\mathbf{x}) g(\mathbf{x})$ for every non-decreasing function u.[6]

[4] As f and g are distribution functions and only differ in $\mathbf{x}$ and $\mathbf{y}$, we must have that if $f(\mathbf{y}) \leq g(\mathbf{y})$ then $f(\mathbf{x}) \geq g(\mathbf{x})$ and $g(\mathbf{y}) - f(\mathbf{y}) = f(\mathbf{x}) - g(\mathbf{x})$. Thus the bilateral transfer we consider is to increase $f(\mathbf{y})$ by $f(\mathbf{x}) - g(\mathbf{x})$ and decrease $f(\mathbf{x})$ by the same amount. The resulting distribution function will then be equivalent to g.

[5] The equivalence between (B) and (C) was proven by Lehmann (1955) and Levhari et al. (1975). The equivalence between (A) and (B) can be obtained as a corollary of a theorem by Strassen (1965) (see e.g. Kamae et al. (1977)) or can be established through an application of the *max-flow min-cut theorem* for flow networks (see e.g. Preston (1974)). Østerdal (2010) gives a direct proof for equivalence between (A) and (B) in the finite case.

[6] A real-valued function u on X is *non-decreasing* if $\mathbf{x}, \mathbf{y} \in X$ and $\mathbf{y} \leq \mathbf{x}$ implies $u(\mathbf{y}) \leq u(\mathbf{x})$.

We add the following notation to ease the description of first order dominance. Given an outcome $\mathbf{x} \in X$, let $L(\mathbf{x}) = \{\mathbf{y} \in X | \mathbf{y} \leq \mathbf{x}\}$ be the lower set and $U(\mathbf{x}) = \{\mathbf{y} \in X | \mathbf{x} \leq \mathbf{y}\}$ be the upper set. For two distributions, f and g, we define the real valued function $s(\mathbf{x}) := f(\mathbf{x}) - g(\mathbf{x})$ for $\mathbf{x} \in X$ along with the sets $P = \{\mathbf{x} \in X | s(\mathbf{x}) > 0\}$ and $R = \{\mathbf{x} \in X | s(\mathbf{x}) < 0\}$. For the elements of R, probability mass has to be added to f for f to become equivalent to g and, likewise, for the elements of P, probability mass has to be subtracted from f for f to become equivalent to g. Thus, for f to first order dominate g, the elements of P have excess probability mass which needs to be transferred to one or more elements of R by diminishing transfers. More precisely, for each element $\mathbf{p} \in P$ we have to transfer $s(\mathbf{p})$ probability mass to a number of elements in $L(\mathbf{p}) \cap R$. In a similar vein, each element $\mathbf{r} \in R$ requires $-s(\mathbf{r})$ probability mass to be transferred from a number of elements in $U(\mathbf{r}) \cap P$ for f to become equivalent to g by diminishing bilateral transfers.

When using diminishing bilateral transfers to obtain g from f it should be noted that we can omit all the elements $\mathbf{x} \in X$ having $s(\mathbf{x}) = 0$. The reason is that if we transfer some probability mass from $\mathbf{x}$ to an element $\mathbf{z}$ of $L(\mathbf{x}) \setminus \{\mathbf{x}\}$, then the same amount of probability mass has to be transferred to $\mathbf{x}$ from some element of $\mathbf{y} \in U(\mathbf{x}) \setminus \{\mathbf{x}\}$. By the definition of $U(\mathbf{x})$ we must have that $L(\mathbf{x}) \subset L(\mathbf{y})$, and the transfer to and from $\mathbf{x}$ can therefore be replaced by a direct transfer from $\mathbf{y}$ to $\mathbf{z}$. More generally, if a sequence of diminishing bilateral transfers uses intermediate elements of X, then, as just described, it is always possible to replace one or more of these diminishing bilateral transfers with a direct diminishing bilateral transfer. Consequently, no outcome needs to both send and receive probability mass. We let $C = \{(\mathbf{p}, \mathbf{r}) \in P \times R | \mathbf{r} \in L(\mathbf{p})\}$ be the pairs of outcomes which correspond to possible (direct) diminishing bilateral transfers.

3 Basic insights

With the definitions given in section 2 we can formulate the problem of checking first order dominance between two finite multivariate distributions as a bipartite network problem. It is essentially a transportation, problem where the "suppliers" from P have to transport a required amount to the "customers" from R. If it is possible to transport probability mass from $\mathbf{p} \in P$ to $\mathbf{r} \in R$, i.e. $(\mathbf{p}, \mathbf{r}) \in C$, we incur a unit cost of zero of transportation from $\mathbf{p}$ to $\mathbf{r}$, whereas if it is not possible to transport probability mass from $\mathbf{p}$ to $\mathbf{r}$, then a unit cost of one is incurred for transporting probability mass from $\mathbf{p}$ to $\mathbf{r}$. Solving the resulting transportation problem either yields a zero value objective, in which case we have identified a finite set of diminishing bilateral transfers, or a strictly positive objective showing that it is necessary to send something from $\mathbf{p} \in P$ to a $\mathbf{r} \notin L(\mathbf{p}) \cap R$. Hence, if the objective is zero, then we have by (A) that f first order dominates g, whereas if the objective is positive, then f does not first order dominate g. If we let $b = \max\{|P|, |R|\}$ and $d = \min\{|P|, |R|\}$, and k is the number of feasible connections, then this problem can be solved in $O(b \log b(k + d \log d))$ by the method described by Kleinschmidt and Schannath (1995). Furthermore, there is a set-up cost of $\Theta(n)$ for identifying P and R from X and of $O(|P||R|)$ for identifying which arcs are feasible. Note that $|P||R| = bd$, which is part of the algorithm's complexity. Thus, using a transportation problem algorithm directly yields a time complexity of $O(n + b \log b(k + d \log d))$.[7]

Below we present an alternative linear programming model which is based on the observation that we just need to identify a feasible transportation problem solution having objective value equal to zero. Let $z_{\mathbf{pr}} \geq 0$ be the amount of probability mass transferred from $\mathbf{p}$ to $\mathbf{r}$, where $(\mathbf{p}, \mathbf{r}) \in C$. Furthermore, let $c_{\mathbf{p}} \geq 0$ for $\mathbf{p} \in P$ and $d_{\mathbf{r}} \geq 0$ for $\mathbf{r} \in R$ be two sets of auxiliary variables. For given values of the $z_{\mathbf{pr}}$ variables, $c_{\mathbf{p}}$ measures the amount of probability mass not transferred out of $\mathbf{p} \in P$ to elements $\mathbf{r} \in L(\mathbf{p}) \cap R$ (in order to reach $s(\mathbf{p})$). Likewise, $d_{\mathbf{r}}$ measures the excess amount (compared to $-s(\mathbf{r})$) of probability mass transferred to element $\mathbf{r} \in R$ from the elements in $U(\mathbf{r}) \cap P$. Then the problem is to identify a feasible set of transfers such that the amount of probability mass which cannot be transferred out of $\mathbf{p} \in P$ and the amount of probability mass received beyond $-s(\mathbf{r})$ for $\mathbf{r} \in R$ are minimized. Hence, we have to solve the following linear program:

$$Z^* = \min \sum_{\mathbf{p} \in P} c_{\mathbf{p}} + \sum_{\mathbf{r} \in R} d_{\mathbf{r}} \tag{1}$$

$$\text{s.t.} \sum_{\mathbf{r} \in L(\mathbf{p}) \cap R} z_{\mathbf{pr}} + c_{\mathbf{p}} \geq s(\mathbf{p}), \quad \mathbf{p} \in P \tag{2}$$

$$\sum_{\mathbf{p} \in U(\mathbf{r}) \cap P} z_{\mathbf{pr}} - d_{\mathbf{r}} \leq -s(\mathbf{r}), \quad \mathbf{r} \in R \tag{3}$$

$$z_{\mathbf{pr}} \geq 0, \quad (\mathbf{p}, \mathbf{r}) \in C \tag{4}$$

$$c_{\mathbf{p}} \geq 0, \quad \mathbf{p} \in P \tag{5}$$

$$d_{\mathbf{r}} \geq 0, \quad \mathbf{r} \in R \tag{6}$$

The objective (1) of this problem is to minimize the untransferred probability mass from the elements of P as well as the excess probability mass transferred to elements of R. The first constraint (2) states that each element $\mathbf{p} \in P$ either has to transfer to the elements in $L(\mathbf{p}) \cap R$ or leave some of the probability mass untransferred. The second constraint (3) states that an element of $\mathbf{r}$ cannot receive more than $-s(\mathbf{r})$ probability mass from the elements in $U(\mathbf{r}) \cap P$, but if it does, then the excess probability mass is added to $d_{\mathbf{r}}$.[8] Finally,

[7] The worst case complexity of the maximum flow problem described by Preston (1974) or Hansel and Troallic (1978) is $O((|P| + |R|)^2 |C|)$ using the preflow-push algorithm described by Goldberg and Tarjan (1988). The linear programming based approaches described by Mosler and Scarsini (1991) and Dyckerhoff and Mosler (1997) are typically solved by means of pseudo-polynomial algorithms such as the Simplex algorithm; see Schrijver (1987).

[8] To be consistent we use the convention that a sum of no elements is zero, i.e. if $L(\mathbf{p}) \cap R = \emptyset$, then

the constraints (4)-(6) just state the non-negativity of the variables.

A solution for problem (1)-(6) will be denoted $(\mathbf{z}, \mathbf{c}, \mathbf{d})$, where $\mathbf{z} = (z_{\mathbf{pr}})_{(\mathbf{p},\mathbf{r})\in C}$, $\mathbf{c} = (c_{\mathbf{p}})_{\mathbf{p}\in P}$, and $\mathbf{d} = (d_{\mathbf{r}})_{\mathbf{r}\in R}$. A solution is said to be feasible if it satisfies all the constraints (2)-(6). Furthermore, a solution is said to be optimal if it is feasible and minimizes objective (1). If the solution is optimal, then it is denoted $(\mathbf{z}^*, \mathbf{c}^*, \mathbf{d}^*)$. A feasible solution to (1)-(6) with $\sum_{\mathbf{p}\in P} c_{\mathbf{p}} + \sum_{\mathbf{r}\in R} d_{\mathbf{r}} = 0$ has the characteristic that the inequality constraints (2) and (3) have to be binding, which is summarized in Lemma 1. This is an important observation for the direct algorithm described in section 4.2.

Lemma 1. *Let* $(\mathbf{z}, \mathbf{c}, \mathbf{d})$ *be a feasible solution to problem* (1)-(6)*. If* $\sum_{\mathbf{p}\in P} c_{\mathbf{p}} + \sum_{\mathbf{r}\in R} d_{\mathbf{r}} = 0$*, then*

1. $c_{\mathbf{p}} = 0$ *for all* $\mathbf{p} \in P$*,*
2. $d_{\mathbf{r}} = 0$ *for all* $\mathbf{r} \in R$*,*
3. *constraints* (2) *and* (3) *will be binding.*

Proof. Let $(\mathbf{z}, \mathbf{c}, \mathbf{d})$ be a feasible solution to problem (1)-(6) with $\sum_{\mathbf{p}\in P} c_{\mathbf{p}} + \sum_{\mathbf{r}\in R} d_{\mathbf{r}} = 0$. As $\sum_{\mathbf{p}\in P} c_{\mathbf{p}} + \sum_{\mathbf{r}\in R} d_{\mathbf{r}} = 0$, and both $\mathbf{c} \geq \mathbf{0}$ and $\mathbf{d} \geq \mathbf{0}$, we must have that $c_{\mathbf{p}} = 0$ for all $\mathbf{p} \in P$ and $d_{\mathbf{r}} = 0$ for all $\mathbf{r} \in R$ showing parts 1 and 2 of the lemma. Part 3 of the lemma can be realized as follows: For each pair $(\mathbf{p}, \mathbf{r}) \in C$ the variable $z_{\mathbf{pr}}$ is present in exactly one of the constraints (2), and the corresponding coefficient is equal to one. The same holds for constraint (3). Thus, we have the relation

$$\sum_{\mathbf{p}\in P} \sum_{\mathbf{r}\in L(\mathbf{p})\cap R} z_{\mathbf{pr}} = \sum_{\mathbf{r}\in R} \sum_{\mathbf{p}\in U(\mathbf{r})\cap P} z_{\mathbf{pr}}$$

Hence, summarizing constraint (2) yields the following

$$\begin{aligned} \sum_{\mathbf{p}\in P} (s(\mathbf{p}) - c_{\mathbf{p}}) &\leq \sum_{\mathbf{p}\in P} \sum_{\mathbf{r}\in L(\mathbf{p})\cap R} z_{\mathbf{pr}} \\ &= \sum_{\mathbf{r}\in R} \sum_{\mathbf{p}\in U(\mathbf{r})\cap P} z_{\mathbf{pr}} \\ &\leq \sum_{\mathbf{r}\in R} (-s(\mathbf{r}) + d_{\mathbf{r}}) \end{aligned}$$

Note that, as $\sum_{\mathbf{x}\in X} s(\mathbf{x}) = \sum_{\mathbf{x}\in X} f(\mathbf{x}) - \sum_{\mathbf{x}\in X} g(\mathbf{x}) = 0$, we have that $\sum_{\mathbf{p}\in P} s(\mathbf{p}) = \sum_{\mathbf{r}\in R} -s(\mathbf{r})$. Consequently, the above has to hold with equality when $c_{\mathbf{p}} = 0$ and $d_{\mathbf{r}} = 0$. When $c_{\mathbf{p}} = 0$ we have that $\sum_{\mathbf{p}\in P} s(\mathbf{p}) = \sum_{\mathbf{p}\in P} \sum_{\mathbf{r}\in L(\mathbf{p})\cap R} z_{\mathbf{pr}}$. Now suppose that for an element $\mathbf{p}' \in P$ we have that $s(\mathbf{p}') < \sum_{\mathbf{r}\in L(\mathbf{p}')\cap R} z_{\mathbf{p}'\mathbf{r}}$. Then some other $\mathbf{p}'' \in P$ must exist having $s(\mathbf{p}'') > \sum_{\mathbf{r}\in L(\mathbf{p}'')\cap R} z_{\mathbf{p}''\mathbf{r}}$, i.e. requiring that $c_{\mathbf{p}''} > 0$, which forces the

$\sum_{\mathbf{r}\in L(\mathbf{p})\cap R} z_{\mathbf{pr}} = 0$, and if $U(\mathbf{r}) \cap P = \emptyset$, then $\sum_{\mathbf{p}\in U(\mathbf{r})\cap P} z_{\mathbf{pr}} = 0$.

objective to be positive. Hence, for the objective to have value zero we therefore must have $s(\mathbf{p}) = \sum_{\mathbf{r} \in L(\mathbf{p}) \cap R} z_{\mathbf{pr}}$ for all $\mathbf{p} \in P$. Therefore, constraint (2) must be binding for an optimal solution to have value zero. An analog argument can be made for (3), and consequently both constraints (2) and (3) have to be binding. □

Clearly, if the optimal solution for problem (1)-(6) is zero, $Z^* = 0$, then we have a feasible finite series of diminishing bilateral transfers – showing that f first order dominates g – whereas a positive objective corresponds to the case where no feasible series of diminishing bilateral transfers exists.[9] In the latter case we conclude that f does not first order dominate g.

Lemma 1 gives a further characterization of first order dominance. That is, if f first order dominates g, then we know that a solution $(\mathbf{z}^*, \mathbf{c}^*, \mathbf{d}^*)$ having value $Z^* = 0$ exists and consequently having $\mathbf{c}^* = \mathbf{0}$, $\mathbf{d}^* = \mathbf{0}$, $\sum_{\mathbf{r} \in L(\mathbf{p}) \cap R} z^*_{\mathbf{pr}} = s(\mathbf{p})$ for all $\mathbf{p} \in P$, and $\sum_{\mathbf{p} \in U(\mathbf{r}) \cap P} z^*_{\mathbf{pr}} = -s(\mathbf{r})$ for all $\mathbf{r} \in R$. The following theorem states this observation explicitly:

Theorem 1. *f first order dominates g if and only if a vector $\mathbf{z} \in \mathbb{R}^{|C|}$ exists with $\mathbf{z} \geq \mathbf{0}$ and*

$$\sum_{\mathbf{r} \in L(\mathbf{p}) \cap R} z_{\mathbf{pr}} = f(\mathbf{p}) - g(\mathbf{p}), \quad \forall \mathbf{p} \in P \tag{7}$$

$$\sum_{\mathbf{p} \in U(\mathbf{r}) \cap P} z_{\mathbf{pr}} = g(\mathbf{r}) - f(\mathbf{r}), \quad \forall \mathbf{r} \in R \tag{8}$$

Proof. If f first order dominates g, then a finite sequence of diminishing bilateral transfers exists. Consequently, a feasible set of transfers between elements of C exists yielding a solution value of zero for problem (1)-(6). Then directly by Lemma 1 we have that (7) and (8) hold. On the other hand, if $\mathbf{z} \geq \mathbf{0}$ exists such that (7) and (8) hold, then it is a feasible solution for problem (1)-(6) with $\mathbf{c} = \mathbf{0}$ and $\mathbf{d} = \mathbf{0}$. The values of $z_{\mathbf{pr}}$ then constitute a finite sequence of diminishing bilateral transfers. □

The linear programming model (1)-(6) may have alternative solutions yielding the same objective value. We make the following observation for a feasible solution.

Lemma 2 (Equivalence of transfer)**.** *Let $(\overline{\mathbf{z}}, \overline{\mathbf{c}}, \overline{\mathbf{d}})$ be a feasible solution for* (1)-(6). *Let $\mathbf{x}, \mathbf{y} \in P$ and $\mathbf{v}, \mathbf{w} \in L(\mathbf{x}) \cap L(\mathbf{y}) \cap R$ and put*

$$\beta = \min\{\overline{z}_{\mathbf{xv}}, \overline{z}_{\mathbf{yw}}\} \tag{9}$$

[9] The number of constraints, not counting the non-negativity constraints, in problem (1)-(6) is exactly $|P| + |R|$. Hence, the basis of the LP will consist of $|P| + |R|$ variables which can attain a non-negative value, while the remaining variables are non-basic at a lower bound of zero. Thus, we need at most $|P| + |R|$ diminishing bilateral transfers.

If $\beta > 0$, then we can construct an alternative solution $(\mathbf{z}', \mathbf{c}', \mathbf{d}')$ having $\mathbf{c}' = \overline{\mathbf{c}}$, $\mathbf{d}' = \overline{\mathbf{d}}$, and all elements of $\mathbf{z}'$ equal to the corresponding elements of $\overline{\mathbf{z}}$ except for

$$z'_{\mathbf{xv}} = \overline{z}_{\mathbf{xv}} - \beta \tag{10}$$

$$z'_{\mathbf{yw}} = \overline{z}_{\mathbf{yw}} - \beta \tag{11}$$

$$z'_{\mathbf{xw}} = \overline{z}_{\mathbf{xw}} + \beta \tag{12}$$

$$z'_{\mathbf{yv}} = \overline{z}_{\mathbf{yv}} + \beta \tag{13}$$

with the same objective value.

Proof. If the two solutions have $\mathbf{c}' = \overline{\mathbf{c}}$ and $\mathbf{d}' = \overline{\mathbf{d}}$, then they have the same objective value. Thus, we have to show that altering $\overline{\mathbf{z}}$ to $\mathbf{z}'$ maintains $\mathbf{c}' = \overline{\mathbf{c}}$ and $\mathbf{d}' = \overline{\mathbf{d}}$. Both for $\mathbf{x}$ and $\mathbf{y}$ we have added and subtracted β in constraint (2), thus not changing the values $c_{\mathbf{x}}$ and $c_{\mathbf{y}}$. Furthermore, we have for both $\mathbf{v}$ and $\mathbf{w}$ added and subtracted β in the constraint (3) thereby not changing $d_{\mathbf{v}}$ or $d_{\mathbf{w}}$ either. Consequently, as solution $\overline{\mathbf{z}}$ was feasible so will $\mathbf{z}'$ be, and they will have the same objective value. □

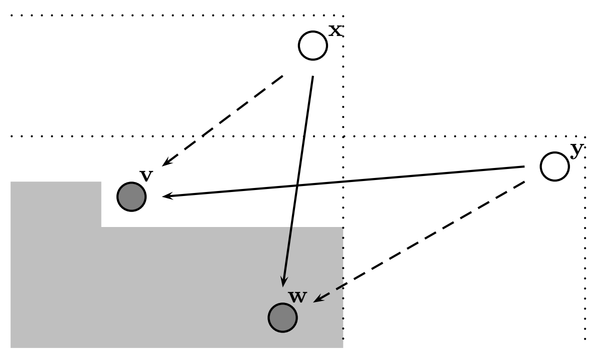

Figure 1: Swap of probability mass transfer.

The observation from Lemma 2 is illustrated for the bivariate case in Figure 1. The full arrows are the bilateral transfers which are decreased, whereas the dashed arrows are the bilateral transfers which are increased. Lemma 2 will be used to show that given any solution to (1)-(6) with objective value zero we can construct an alternative solution following a specific pattern. As a consequence, it will only be necessary to search for this pattern when checking first order dominance in the bivariate case.

4 Algorithms for the Bivariate Case

First we introduce a constructive $O(n^2)$ algorithm for identifying a finite sequence of diminishing bilateral transfers. Then we observe that it is not necessary to construct the finite sequence of diminishing bilateral transfers directly to show that such a sequence exists, and we use this observation as a base for an $O(n)$ algorithm determining whether or not f first order dominates g.

For both algorithms the elements of X are searched in a specific order. This ordering is described in section 4.1. Then in section 4.2 we introduce the direct algorithm, and finally in section 4.3 we introduce the indirect algorithm for checking first order dominance.

4.1 Ordering elements

The elements of X can be completely ordered in many ways – in this paper we use two specific complete orderings. In what follows we let $a, b \in \{1, 2\}$ with $a \neq b$. We say that an element $\mathbf{x} = (x_1, x_2) \in X$ has a lower (a, b)-order than $\mathbf{y} = (y_1, y_2) \in X$ if $x_a < y_a$ or if $x_a = y_a$ and $x_b > y_b$. In case $\mathbf{x}$ has a lower (a, b)-order than $\mathbf{y}$ then we write $o_{ab}(\mathbf{x}) < o_{ab}(\mathbf{y})$. Hence, the ordering increases with increasing element indices from X_a subsequently decreasing elements of X_b.

4.2 A direct approach

First we identify a sequence of diminishing bilateral transfers which may give a solution where $Z^* = 0$, and then we show that if a solution having value $Z^* = 0$ exists, then we can transform this solution into the solution found by this sequence. Consequently, we have that the solution found by the given sequence will identify a solution having value $Z^* = 0$ if and only if such a solution exists.

In the direct algorithm we manipulate the variables of the formulation (1)-(6) in such a way that we keep $d_{\mathbf{r}} = 0$ for all $\mathbf{r} \in R$ while keeping constraint (2) binding. Hence, we have that $c_{\mathbf{p}} = s(\mathbf{p}) - \sum_{\mathbf{r} \in L(\mathbf{p}) \cap R} z_{\mathbf{pr}}$ for all $\mathbf{p} \in P$, and we put $\rho_{\mathbf{r}} = -s(\mathbf{r}) - \sum_{\mathbf{p} \in U(\mathbf{r}) \cap P} z_{\mathbf{pr}}$. Clearly, $\rho_{\mathbf{r}}$ corresponds to the slack variable of constraint (3) and has to be non-negative. Implicitly, we initialize $z_{\mathbf{pr}} = 0$ for all pairs $(\mathbf{p}, \mathbf{r}) \in P \times R$.[10] Thus, increasing $z_{\mathbf{pr}}$ will decrease both $c_{\mathbf{p}}$ and $\rho_{\mathbf{r}}$, and the aim is to identify a sequence of increases of $z_{\mathbf{pr}}$ such that both $c_{\mathbf{p}}$ and $\rho_{\mathbf{r}}$ become zero for all $\mathbf{p}$ and $\mathbf{r}$.

The direct approach is given in Algorithm 1. It initializes the variables and then repeats three steps, which results in either a solution to the problem (1)-(6) with a value zero (terminating in step 1) or a positive objective value (terminating in step 2).

[10] It is not necessary to initialize $z_{\mathbf{pr}}$ explicitly as we just need to keep track of which pairs $(\mathbf{p}, \mathbf{r})$ have increased values of $z_{\mathbf{pr}}$.

Step 0 Initialize $c_{\mathbf{p}} = s(\mathbf{p})$ for all $\mathbf{p} \in P$ and $\rho_{\mathbf{r}} = -s(\mathbf{r})$ for all $\mathbf{r} \in R$.

Step 1 Select $\mathbf{p} \in P$ with minimal $(1,2)$-order, $o_{12}(\mathbf{p})$, such that $c_{\mathbf{p}} > 0$. If no such $\mathbf{p}$ exists, then terminate, continue otherwise.

Step 2 Select $\mathbf{r} \in L(\mathbf{p}) \cap R$ with maximal $(2,1)$-order, $o_{21}(\mathbf{r})$, such that $\rho_{\mathbf{r}} > 0$. If no such $\mathbf{r}$ exists, then terminate else continue.

Step 3 Update

$$\begin{aligned} z_{\mathbf{pr}} &= \min\{c_{\mathbf{p}}, \rho_{\mathbf{r}}\} \\ c_{\mathbf{p}} &= c_{\mathbf{p}} - z_{\mathbf{pr}} \\ \rho_{\mathbf{r}} &= \rho_{\mathbf{r}} - z_{\mathbf{pr}} \end{aligned} \tag{14}$$

Go to step 1.

Algorithm 1: Direct bilateral transfer algorithm

An example of the progression in the three steps is given in the left part of Figure 2. It shows the sequence of selections of elements of P and R. Elements of P are white nodes, whereas elements of R are black nodes. The first element of P encountered is $\mathbf{p}^1$, and it transfers probability mass to $\mathbf{r}^6$, $\mathbf{r}^5$, $\mathbf{r}^4$, and $\mathbf{r}^3$. While $\mathbf{r}^6$, $\mathbf{r}^5$, and $\mathbf{r}^4$ are fully saturated, $\mathbf{r}^3$ only received a fraction of $-s(\mathbf{r}^3)$, and it can therefore receive more later in the algorithm. The sequence of diminishing transfers of probability mass away from $\mathbf{p}^1$ is illustrated by the full black arrows in the figure. The same is then done for $\mathbf{p}^2$ and $\mathbf{p}^3$, where the gray arrows represent the sequence of diminishing transfers from $\mathbf{p}^2$ and the dashed arrows show the sequence of diminishing transfers from $\mathbf{p}^3$.

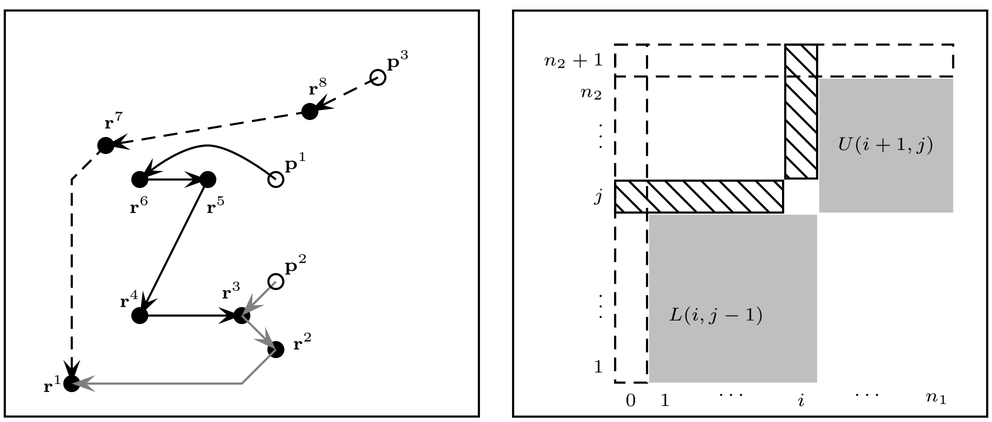

Figure 2: Left: The process of the $O(n^2)$ algorithm. Right: The base sets for the $O(n)$ algorithm.

Clearly, if a solution with value $Z^* = 0$ is found then we have identified a feasible solution to problem (1)-(6), and we can therefore conclude that f first order dominates g. On the other hand, if a solution value greater than zero is found by the approach above, then we need to guarantee that no zero value solution actually exists. We do this in Theorem 2 by showing that if a zero value solution exists, then the solution found by the above approach will also have a value of zero. We need the following Lemma:

Lemma 3. *Let $X = X_1 \times X_2$. Given four elements $\mathbf{x}, \mathbf{y}, \mathbf{v}, \mathbf{w} \in X$ having $\mathbf{w} \in L(\mathbf{x})$ and $\mathbf{v} \in L(\mathbf{y})$, then if $x_1 \leq y_1$ and $v_2 \geq w_2$, then $\mathbf{w} \in L(\mathbf{y})$.*

Proof. This can be realized as follows: As $\mathbf{w} \in L(\mathbf{x})$ then $w_1 \leq x_1$, and by assumption we have that $x_1 \leq y_1$. Furthermore, we have that $\mathbf{v} \in L(\mathbf{y})$, then $v_2 \leq y_2$, and by assumption we have that $w_2 \leq v_2$. Hence, $\mathbf{w} \in L(\mathbf{y})$. □

Theorem 2. *Algorithm 1 terminates in at most $O(n^2)$ iterations and either terminates with a finite sequence of diminishing bilateral transfers or shows that no such sequence exists.*

Proof. If Algorithm 1 terminates in step 1, then a sequence of diminishing bilateral transfers has been found satisfying all of the constraints (2) and (3). This has been done such that all $c_{\mathbf{p}}$ values have been decreased to zero while keeping the $d_{\mathbf{r}}$ values at zero. Hence, a finite sequence of diminishing bilateral transfers exists.

Suppose that we have a solution $\mathbf{z}^*$ having value $Z^* = 0$, and $\mathbf{z}^*$ is not equal to a solution found by the approach above. Let the solution found by the approach above be $\overline{\mathbf{z}}$ with $\mathbf{z}^* \neq \overline{\mathbf{z}}$. As these are not identical, an element $\mathbf{x} \in P$ must exist such that the corresponding $\mathbf{v} \in R$ found by the approach above is not transferring as much probability mass from $\mathbf{x}$ to $\mathbf{v}$ in the optimal solution $\mathbf{z}^*$ as in the solution $\overline{\mathbf{z}}$. Hence $z^*_{\mathbf{xv}} < \overline{z}_{\mathbf{xv}}$. Note that Algorithm 1 is greedy and transfers as much as possible to elements with a lower $(1,2)$-order before transferring probability mass to elements with a higher $(1,2)$-order. Therefore, the lowest $(1,2)$-order element $\mathbf{x}$ having $z^*_{\mathbf{xv}} \neq \overline{z}_{\mathbf{xv}}$ will have $z^*_{\mathbf{xv}} < \overline{z}_{\mathbf{xv}}$. Select first pair $(\mathbf{x}, \mathbf{v})$ in the $(1,2)$-order where $z^*_{\mathbf{xv}} < \overline{z}_{\mathbf{xv}}$. Hence, a pair $(\mathbf{x}, \mathbf{w})$ must exist where $\mathbf{w} \in R$ would be selected later by step 2 with value $z^*_{\mathbf{xw}} > 0$. As $\mathbf{w}$ is selected later, we must have that $o_{21}(\mathbf{v}) > o_{21}(\mathbf{w})$, and consequently $v_2 \geq w_2$. Note that the gray area in Figure 1 is the area of possible selections of the element $\mathbf{w}$. On the other hand, an element $\mathbf{y} \in P$ found later in the $(1,2)$-ordering in step 1 must transfer probability mass to $\mathbf{v}$, because constraint (3) holds with equality in the optimal solution with value $Z^* = 0$, and therefore $z^*_{\mathbf{yv}} > 0$. Because $\mathbf{y}$ is selected later than $\mathbf{x}$ in step 1, we must have that $o_{12}(\mathbf{x}) < o_{12}(\mathbf{y})$ and therefore that $x_1 \leq y_1$. Using Lemma 3 we can conclude that $\mathbf{v}, \mathbf{w} \in L(\mathbf{x}) \cap L(\mathbf{y}) \cap R$. Hence, we can do the exchange given by Lemma 2 increasing $z^*_{\mathbf{xv}}$ and $z^*_{\mathbf{yw}}$ while decreasing $z^*_{\mathbf{xw}}$ and $z^*_{\mathbf{yv}}$ correspondingly. This process can be continued until the two solutions are identical. Thus, if we have a solution $\mathbf{z}^*$ with value $Z^* = 0$, we can always transform it into an alternative solution which can be constructed by Algorithm 1. Consequently, if it is not possible to identify a zero-value solution by Algorithm 1, then no zero value solution exists, and we can therefore conclude that no finite sequence of diminishing bilateral transfers exists.

The approach using the three steps above has a time complexity of $O(n^2)$, as for each element $\mathbf{p} \in P$ we have to search through the elements of $L(\mathbf{p})$ to identify a suitable element

$\mathbf{r} \in L(\mathbf{p}) \cap R$. □

4.3 An indirect approach

We now present a more efficient algorithm for the first order dominance. Like Algorithm 1 it is based on iterating through the elements of X in increasing $(1,2)$-order, but it only records how much probability mass is transferred between specific aggregated subsets without specifying the bilateral transfers directly.

To ease the notation, we expand the sets X_1 and X_2 to $\overline{X}_1 = X_1 \cup \{0\}$ and $\overline{X}_2 = X_2 \cup \{n_2+1\}$. Now we use the simpler notation $\mathbf{x} = (i,j)$ for $i \in \overline{X}_1$ and $j \in \overline{X}_2$. Furthermore, we let $f_{ij} = f(\mathbf{x})$, $g_{ij} = g(\mathbf{x})$, and $s_{ij} = s(\mathbf{x})$. We let $f_{0j} = g_{0j} = s_{0j} = 0$ for all $j \in \overline{X}_2$, and $f_{i,n_2+1} = g_{i,n_2+1} = s_{i,n_2+1} = 0$ for all $i \in \overline{X}_1$. We say that elements $(i,\cdot)$ are the column of i, and the elements $(\cdot,j)$ are the row of j. See Figure 2 for an illustration of the set-up, where the dashed boxes correspond to the artificial elements added, and the hatched boxes correspond to the row and column elements in row j and column i having a lower $(2,1)$-order and higher $(1,2)$-order, respectively, compared to the element (i,j).

For each element $(i,j) \in X$ we associate a variable $e_{ij} \geq 0$, which is the excess probability mass left in element (i,j) after processing this element and all elements having lower $(1,2)$-order. The value of e_{ij} is the amount of probability mass which can be transferred to row j by some unprocessed element $\mathbf{y} \in U(i+1,j)$. We also associate the variable $u_{ij} \geq 0$ with the untransferred probability mass from the set of elements (i,h) in column i with $h \geq j$ after processing element (i,j) and all elements having a lower $(1,2)$-order. The value u_{ij} is the amount of probability mass which is necessary to transfer to some element in $L(i,j-1)$. Again we let the boundary values be $e_{0j} = 0$ for all $j \in \overline{X}_2$ and $u_{i,n_2+1} = 0$ for all $i \in \overline{X}_1$. It is important to note that a value of $e_{ij} = 0$ means that no more probability mass can be sent to elements (h,j) with $h = 1,\ldots,i$ in the remaining iterations. Furthermore, if $u_{ij} > 0$, then some of the elements (i,k) for $k = j,\ldots,n_2$ have untransferred probability mass. In this case we need to transfer at least u_{ij} into the set $L(i,j-1)$. On the other hand, if $u_{ij} = 0$, then all required probability mass has been transferred. The variables e_{ij} and u_{ij} will never be positive simultaneously. If they were, then we could send $\delta = \min\{e_{ij}, u_{ij}\}$ from (i,j) to the elements of row i and then decrease both variables by δ. Clearly, δ is zero when either e_{ij} or u_{ij} is zero, in which case it is not possible to reduce these further.

Let $t_{ij} = u_{i,j+1} - e_{i-1,j} + s_{ij}$ be net required transfer after processing element (i,j). If $t_{ij} > 0$, then t_{ij} units still have to be transferred to $L(i,j-1)$, whereas if $t_{ij} < 0$, then we can still transfer $-t_{ij}$ units to the elements in (h,j) for $h = 1,\ldots,i$ from $U(i+1,j)$. We can calculate the values of e_{ij} and u_{ij} by the following recursions:

$$u_{ij} = \max\{0, t_{ij}\}, \quad (i,j) \in X \tag{15}$$

and

$$e_{ij} = \max\{0, -t_{ij}\}, \quad (i,j) \in X \tag{16}$$

By this recursion we transfer as much as possible from the elements (i,k) with $k \geq j$ to the elements (h,j) with $h \leq i$ before proceeding to the next element with higher $(1,2)$-order. The indirect algorithm can be stated as in Algorithm 2.

Step 0 Let $f_{0j} = g_{0j} = e_{0j} = 0$ for all $j \in \overline{X}_2$ and $f_{i,n_2+1} = g_{i,n_2+1} = u_{i,n_2+1} = 0$ for all $i \in \overline{X}_1$.
Let $i = 1$, $j = n_2$.

Step 1 Calculate

$$\begin{aligned} s_{ij} &= f_{ij} - g_{ij} \\ t_{ij} &= u_{i,j+1} - e_{i-1,j} + s_{ij} \\ u_{ij} &= \max\{0, t_{ij}\} \\ e_{ij} &= \max\{0, -t_{ij}\} \end{aligned}$$

Step 2 Choose one of the following

- If $j > 1$, then put $j = j - 1$ and goto step 1.
- If $j = 1$, $i < n_1$, and $u_{i1} = 0$, then put $i = i + 1$, $j = n_2$ and goto step 1.
- If $j = 1$ and $u_{i1} > 0$, then return FALSE.
- If $j = 1$, $i = n_1$, and $u_{i1} = 0$, then return TRUE.

Algorithm 2: Indirect bilateral transfer algorithm

The algorithm iterates through the elements of X in the $(1,2)$-order. Each element is processed at most once, and processing an element requires a constant number of calculations. Hence, the time complexity is $O(n)$. Intuitively, if for some $\mathbf{x} = (i,1)$ we achieve $u_{i1} > 0$, then it is required to transfer probability mass from an element in (i,h) with $h \geq 1$ to an element in $L(i,0)$, but as the element $(i,0) \notin X$ does not exist, we have that $L(i,0) = \emptyset$ and we can therefore not accommodate the required transfer. Hence, we can conclude that it is not possible to make a sequence of diminishing bilateral transfers, and the algorithm terminates, which implies that f does not first order dominate g. On the other hand, if the algorithm iterates through all the elements of X without this happening, then we have found a feasible sequence of diminishing bilateral transfers, and the algorithm concludes that f first order dominates g.

Lemma 4 states the case where the algorithm returns that f first order dominates g, whereas Lemma 5 describes the case where the algorithm returns that f does not first order dominate g. Finally, Theorem 3 states the correctness and time complexity of the indirect algorithm.

Lemma 4. *If Algorithm 2 terminates with $u_{n_1 1} = 0$, then a finite sequence of diminishing bilateral transfers exists such that g can be obtained from f.*

Proof. We prove this by showing that the sequence of diminishing bilateral transfers obtained by algorithm 1 can be obtained by algorithm 2 as well. Algorithm 2 traverses the elements in increasing $(1,2)$-order and therefore encounters elements of P in the same order as Algorithm 1. Suppose that we maintain a list L of the encountered elements of $\mathbf{p} \in P$ having $c_{\mathbf{p}} > 0$ which is sorted in increasing $(1,2)$-order. These are the elements for which we have not transferred a sufficient amount of probability mass. Furthermore, for each element $j \in X_2$ we maintain a list Q_j of encountered elements of R with $\rho_{\mathbf{r}} > 0$ which is sorted in increasing $(2,1)$-order. When reaching element $(i,j) \in X$ then if $s_{ij} > 0$ we add the (i,j) to the end of L as it has the highest encountered $(1,2)$-order. On the other hand, if $s_{ij} < 0$, then we add the (i,j) to the front of Q_j as it has a lower $(2,1)$-order than the other elements of the list Q_j. The value t_{ij} corresponds to $\sum_{\mathbf{p}\in L} c_{\mathbf{p}} - \sum_{\mathbf{r}\in Q_j} \rho_{\mathbf{r}}$ after doing the above insertions. If either L or Q_j is empty, then it is not possible to transfer any probability mass and therefore the next element from X is selected in increasing $(1,2)$-order. On the other hand, if both lists are not empty, then it is possible to transfer probability mass. Therefore, repeat the following until either L or Q_j becomes empty; select the first element of $\mathbf{p} \in L$ and the last element of $\mathbf{r} \in Q_j$. Put $z_{\mathbf{pr}} = \min\{c_{\mathbf{p}}, \rho_{\mathbf{r}}\}$ and then update $c_{\mathbf{p}} = c_{\mathbf{p}} - z_{\mathbf{pr}}$ and $\rho_{\mathbf{r}} = \rho_{\mathbf{r}} - z_{\mathbf{pr}}$. Either $c_{\mathbf{p}}$ or $\rho_{\mathbf{r}}$ becomes zero. If $c_{\mathbf{p}}$ becomes zero, then $\mathbf{p}$ is removed from L, and if $\rho_{\mathbf{r}}$ becomes zero, then $\mathbf{r}$ is removed from Q_j. In this way, we always allocate the lowest $(1,2)$-order $\mathbf{p} \in P$ having $c_{\mathbf{p}} > 0$ to the highest compatible $(2,1)$-order element of $\mathbf{r} \in R$. If $u_{i1} = 0$, then we have that $t_{ij} \leq 0$, which indicates that $\sum_{\mathbf{p}\in L} c_{\mathbf{p}} \leq \sum_{\mathbf{r}\in Q_1} \rho_{\mathbf{r}}$. If at this point $\sum_{\mathbf{p}\in L} c_{\mathbf{p}} > 0$, then we can still transfer this amount to $L(i,1)$ and decrease $\sum_{\mathbf{p}\in L} c_{\mathbf{p}}$ and $\sum_{\mathbf{r}\in Q_1} \rho_{\mathbf{r}}$ accordingly. Consequently, if $u_{i1} = 0$, then the list L is empty, and we proceed to the next column. When the algorithm reaches element $(n_1, 1)$ with a value of $u_{n_1,1} = 0$, then we have constructed a finite series of diminishing bilateral transfers. This is exactly what Algorithm 1 does, and we can therefore obtain a finite sequence of diminishing bilateral transfers to obtain g from f. □

Lemma 5. *If Algorithm 2 terminates with $u_{i1} > 0$ a comprehensive set $Y \subseteq X$ exists such $\sum_{\mathbf{x}\in Y} g(\mathbf{x}) < \sum_{\mathbf{x}\in Y} f(\mathbf{x})$.*

Proof. We can explicitly identify a comprehensive set which violates (B) if the value of $u_{i1} > 0$ for some $i \in X_1$. First, note that

$$u_{ij} - e_{ij} = \max\{0, t_{ij}\} - \max\{0, -t_{ij}\} = t_{ij} = u_{i,j+1} - e_{i-1,j} + s_{ij}$$

and suppose that we are given a comprehensive set Y. Then we have

$$\sum_{(i,j)\in Y} (u_{ij} - e_{ij}) = \sum_{(i,j)\in Y} (u_{i,j+1} - e_{i-1,j}) + \sum_{(i,j)\in Y} s_{ij} \tag{17}$$

Now let

$$\begin{array}{lclll} H & = & \{ \quad i \in X_1 & \mid \quad (i,1) \in Y \} \\ I & = & \{ \quad j \in X_2 & \mid \quad (1,j) \in Y \} \\ J & = & \{ \quad (i,j) \in Y & \mid \quad (i+1,j) \notin Y & \} \\ K & = & \{ \quad (i,j) \notin Y & \mid \quad (i,j-1) \in Y & \} \end{array}$$

The sets J and K are illustrated in Figure 3, where Y is the gray area (including both shades of gray). The union of the hatched boxes corresponds to J, whereas the union of the dark gray boxes is the set K. Furthermore, the dashed box is the elements $(i,1) \in Y$ with $i \in H$, and the dotted box is the elements $(1,j)$ with $j \in I$. We can then rearrange (17) as

$$\sum_{(i,j)\in Y} s_{ij} = \sum_{j\in I} e_{0j} - \sum_{(i,j)\in J} e_{ij} + \sum_{h\in H} u_{h1} - \sum_{(i,j)\in K} u_{ij} \tag{18}$$

where $\sum_{j\in I} e_{0j} = 0$ by the definition of e_{0j}. Showing that the comprehensive set Y violates condition (B) corresponds to showing that $\sum_{(i,j)\in Y} s_{ij} > 0$, which is equivalent to showing that $\sum_{h\in H} u_{h1} > \sum_{(i,j)\in J} e_{ij} + \sum_{(i,j)\in K} u_{ij}$.

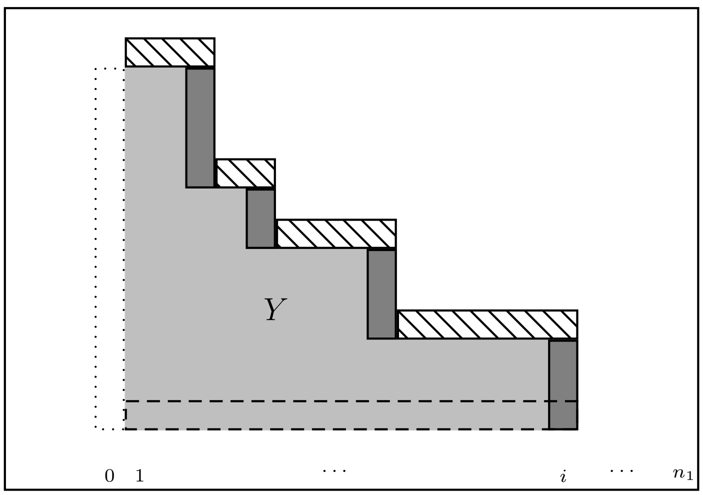

Figure 3: A violated comprehensive set.

Suppose that the Algorithm 2 terminates with $u_{i1} > 0$. Then we know that $u_{h1} = 0$ for $h = 1, \ldots, i-1$ and therefore $\sum_{h\in H} u_{h1} = u_{i1}$. Hence, if we can construct the comprehensive set Y such that $\sum_{(h,j)\in J} e_{hj} + \sum_{(h,j)\in K} u_{hj} = 0$ then we have the violation we are seeking. We construct Y implicitly by constructing J and K explicitly. Each time an element (h,j) is added to J, then all elements (a,j) with $a \le h$ are added to Y. Start with $(h,j) = (i,1)$. Because $u_{i1} > 0$, we have that $e_{i1} = 0$. Therefore add $(i,1)$ to J. Repeat the following until $h = 0$. If $u_{h,j+1} = 0$, add $(h,j+1)$ to K and put $h = h-1$, otherwise $u_{h,j+1} > 0$

and consequently $e_{h,j+1} = 0$ and therefore add $(h, j + 1)$ to J and put $j = j + 1$. When terminating we have only added elements to J having $e_{hj} = 0$ and elements to K having elements $u_{hj} = 0$, thus having $\sum_{(i,j)\in J} e_{ij} + \sum_{(i,j)\in K} u_{ij} = 0 < u_{i1}$ showing that (B) is violated by Y. □

Theorem 3. *Algorithm 2 terminates in $O(n)$ iterations either stating that f first order dominates g or that f does not first order dominate g.*

Proof. If the algorithm terminates with $u_{n_1 1} = 0$, then we have by Lemma 4 and property (A) that f first order dominates g. On the other hand, if the algorithm terminates with $u_{i1} > 0$ for some $i \in X_1$, then by Lemma 5 a violated comprehensive set exists. Consequently, by property (B) f does not first order dominate g. Finally, as each element of X is traversed maximally once and the number of operations for each element is constant, the algorithm terminates in $O(n)$ iterations. □

It is possible to achieve the sequence of diminishing bilateral transfers without increasing the worst case time complexity by augmenting Algorithm 2. The proof of Lemma 4 uses insertion of elements into lists. These lists and the corresponding insertions could be added to an augmented version of Algorithm 2. Insertions and deletions of elements from the lists are only performed in the end of the lists, and we can therefore do this in $O(1)$ time complexity. Throughout the algorithm at most $|P| + |R|$ elements is inserted into the lists, as any element is inserted into one of the lists only once. Hence, the number of insertions and number of deletions are therefore bounded by $|P| + |R|$ and this augmentation will have a complexity of $O(n)$.

Furthermore, it is also possible to derive a violating comprehensive set without increasing the worst case time complexity by applying the constructive method in the proof of Lemma 5. The method uses no more than $\max\{|P|, |R|\}$ iterations of complexity $O(1)$, and Algorithm 2 can therefore be augmented with this construction and still have a worst case time complexity of $O(n)$.

5 Conclusion

In this paper we have obtained a new characterization of first order dominance for the general multivariate case. Furthermore, we have described two algorithms for checking first order dominance in the bivariate case, one of which has linear time worst case complexity. It remains an open topic for further research to develop an efficient approach for checking first order dominance in the multivariate case.

References

Arndt, C., Distante, R., Hussain, M. A., Østerdal, L. P., Huong, P. L., and Ibraimo, M. (2012). Ordinal welfare comparisons with multiple discrete indicators: A first order dominance approach and application to child poverty. *World Development*, 40:2290 – 2301.

Arndt, C., Hussain, M., Salvucci, V., Tarp, F., and Østerdal, L. P. (2013). Advancing small area estimation. WIDER Working Paper, No. 2013/053.

Atkinson, A. and Bourguignon, F. (1982). The comparison of multi-dimensioned distributions of economic status. *Review of Economics Studies*, 49:183–201.

Bawa, V. S. (1982). Stochastic dominance: A research bibliography. *Management Science*, 28(6):698–712.

Dyckerhoff, R. and Mosler, K. (1997). Orthant orderings of discrete random vectors. *Journal of Statistical Planning and Inference*, 62:193–205.

Goldberg, A. V. and Tarjan, R. E. (1988). A new approach to the maximum-flow problem. *Journal of the Association for Computing Machinery*, 35(4):921–940.

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

Hansel, G. and Troallic, J. (1978). Measures marginales et théorème de Ford-Fulkerson. *Zeitschrift für Wahrsheinlichkeitstheorie und verwandte Gebiete*, 43:345–251.

Harder, J. and Russell, W. (1974). Stochastic dominance in choice under uncertainty. In Balch, M., D.L., M., and Wu, S., editors, *Essays on Economic Behavior under Uncertainty*. North-Holland, Amsterdam.

Huang, C., Kira, D., and Vertinsky, I. (1978). Stochastic dominance rules for multi-attribute utility functions. *Review of Economic Studies*, 45:611–615.

Kamae, T., Krengel, U., and O'Brien, G. (1977). Stochastic inequalities on partially ordered spaces. *The Annals of Probability*, 5:899–912.

Kleinschmidt, P. and Schannath, H. (1995). A strongly polynomial algorithm for the transportation problem. *Mathematical Programming*, 68:1–13.

Lehmann, E. (1955). Ordered families of distributions. *The Annals of Mathematical Statistics*, 26:399–419.

Levhari, D., Paroush, J., and Peleg, B. (1975). Efficiency analysis for multivariate distributions. *Review of Economic Studies*, 42:87–91.

Levy, H. and Paroush, J. (1974). Toward multivariate efficiency criteria. *Journal of Economic Theory*, 7:129–142.

Marshall, A. and Olkin, I. (1979). *Inequalitites: Theory of Majorization and Its Applications*, volume 143 of *Mathematics in Science and Engineering*. New York: Academic Press.

Mosler, K. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Science*, 30:1311–1322.

Mosler, K. and Scarsini, M. (1991). Some theory of stochastic dominance. In Mosler, K. and Scarsini, M., editors, *Stochastic Orders and Decision under Risk*. Institute of Mathematical Statistics, Heyward, California.

Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks.* John Wiley and Sons.

Preston, C. (1974). A generalization of the FKG inequalities. *Communications in Mathematical Physics.*

Russell, W. and Seo, T. (1978). Ordering uncertain prospects: The multivariate utility functions case. *Review of Economic Studies*, 45:605–610.

Scarsini, M. (1988). Dominance conditions for multivariate utility functions. *Management Science*, 34:454–460.

Schrijver, A. (1987). *Theory of Linear and Integer Programming.* Discrete Mathematics and Optimization. Wiley-Interscience.

Shaked, M. and Shanthikumar, J. (2007). *Stochastic Orders.* Springer.

Silvapulle, M. J. and Sen, P. K. (2011). *Constrained statistical inference: Order, inequality, and shape constraints.*, volume 912. Wiley-Interscience.

Sriboonchita, S., Dhompongsa, S., Wong, W. K., and Nguyen, H. T. (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*, 36:423439.

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.