Concepts and Applications of Stochastically Weighted Stochastic Dominance

Jian Hu
Department of Industrial Engineering and Management Sciences
Northwestern University
jianhu@northwestern.edu

Tito Homem-de-Mello
Department of Mechanical and Industrial Engineering
University of Illinois at Chicago
thmello@uic.edu

Sanjay Mehrotra
Department of Industrial Engineering and Management Sciences
Northwestern University
mehrotra@iems.northwestern.edu

April 8, 2011

Abstract

Stochastic dominance theory provides tools to compare random entities. When comparing random vectors (say X and Y), the problem can be viewed as one of multi-criterion decision making under uncertainty. One approach is to compare weighted sums of the components of these random vectors using univariate dominance. In this paper we propose new concepts of *stochastically weighted dominance*. The main idea is to treat the vector of weights as a random vector V. We show that such an approach is much less restrictive than the deterministic weighted approach. We further show that the proposed new concepts of stochastic dominance are representable by a finite number of (mixed-integer) linear inequalities when the distributions of X, Y and V have finite support. We discuss two applications to illustrate the usefulness of the stochastically weighted dominance concept. The first application discusses the effect of this notion on the feasibility regions of optimization problems. The second application presents a multi-criterion staffing problem where the goal is to decide the allocation of servers between two $M/M/c$ queues based on waiting times. The latter example illustrates the use of stochastically weighted dominance concept for a ranking of the alternatives.

Key Words: Stochastic Programming, Stochastic Dominance, Risk Management, Chance Constraint, Integer Programming

1 Introduction

Stochastic dominance provides a way to compare the distributions of two random variables or vectors, which is useful not only from a statistical but also from economics and decision-making perspectives (see, e.g., Shaked and Shanthikumar 1994, Müller and Stoyan 2002 and Levy 2006 for comprehensive treatments of the topic). The concept of stochastic dominance can be framed in terms of utility theory (von Neumann and Morgenstern 1947), which hypothesizes that for each rational decision maker there exists a utility function u such that random gains X are preferred to random benchmarks Y if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$. Often the exact form of the utility function is not known but we can still model risk attitudes; for example, we can model risk-averseness by saying that X is preferred to Y if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all non-decreasing and concave utility functions u for which the expectations are finite. Different notions of stochastic dominance can be defined, corresponding to different classes of functions u.

When X and Y are random vectors, stochastic dominance can be used as a tool for *multi-criterion decision making*, since each component of X and Y can be interpreted as the uncertain outcome of a given criterion. Zaras and Martel (1994) and Nowak (2004), however, expressed concerns that the notion of stochastic dominance defined above, when applied to the case where X and Y are random vectors, may be too conservative in practice, since it requires that $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all multidimensional functions in a certain class $\mathcal{U}$ (e.g., nondecreasing concave functions). Moreover, testing for this notion of dominance can be difficult, although recent work by Armbruster and Luedtke (2010) provides some new tools for that. Under the assumptions that the random outcomes of the various alternatives are probabilistically independent and also satisfy additive independence conditions for different criteria, Zaras and Martel (1994) and Nowak (2004) studied decomposition and relaxation of the multivariate stochastic dominance relation. Unfortunately, those assumptions are not easily satisfied in a real world, where alternatives usually have correlated attributes under different criteria.

A different approach to compare random vectors, which is less restrictive than the multivariate dominance condition defined above, uses the notion of multivariate *linear* stochastic dominance, which compares linear combinations of the random vectors. In the context of multi-criterion decision making, the coefficients of the linear combinations can be interpreted as criterion weights. Given a vector of weights v and a vector of outcomes X, the information is summarized into a scalar by a linear value function $\Psi_v(X) := v^T X$. Then, $\Psi_v(X)$ is compared with $\Psi_v(Y)$, using univariate stochastic dominance, for all v in a certain set of weights $\mathcal{V}$. More precisely, X dominates Y in the multivariate linear sense if $\mathbb{E}[u(\Psi_v(X))] \geq \mathbb{E}[u(\Psi_v(Y))]$ for all $v \in \mathcal{V}$ and $u \in \mathcal{U}$. The standard notion of multivariate linear dominance assumes that $\mathcal{V} = \mathbb{R}^n_+$ (see, e.g., Müller and Stoyan 2002), but Homem-de-Mello and Mehrotra (2009) and Hu et al. (2010) allow $\mathcal{V}$ to be an arbitrary polyhedral and a convex set, respectively. The advantage of allowing the set $\mathcal{V}$ to be defined by the user is that it allows for the representation of the evaluation of the importance of each criterion by different decision makers. In that sense, $\mathcal{V}$ can be viewed as a robust weight region (Hu et al. 2011; Hu and Mehrotra 2010). Hu et al. (2010) discuss two natural ways to construct a weight region — one that specifies the set $\mathcal{V}$ as a convex hull of a given set of weight vectors, and another whereby the weights are treated as independent and identically distributed samples from an unknown distribution, in which case $\mathcal{V}$ is an ellipsoidal confidence region for the expected weight.

The use of a robust weight region $\mathcal{V}$ implicitly assumes that the evaluations of all decision makers are treated equally. In many practical settings, though, such a condition is not satisfied; a hierarchical group decision is a typical example, since the decisions of the group leaders are often emphasized. Thus, it is important to incorporate a measure of the *relative importance* of each criterion weight vector, which represents the influence of each decision maker.

Another drawback of using a robust weight region $\mathcal{V}$ is that, even though it imposes less than the notion of full multivariate stochastic dominance, it may still be too conservative. This is particularly troublesome when the dominance relation is included as a constraint in an optimization problem — a type of model has recently received great attention in the literature as it allows for control of risk within an optimization framework, see for instance Dentcheva and Ruszczyński (2003, 2006, 2009); Roman et al. (2006); Drapkin and Schultz (2007); Luedtke (2008); Homem-de-Mello and Mehrotra (2009); Hu et al. (2010, 2011). In the optimization context, an overly conservative constraint may lead to a small or even empty feasibility region. This was observed by Hu et al. (2011), who proposed a stochastic dominance-constrained optimization model for a homeland security problem using real data. Their computation results showed that the model had a small feasibility region. In fact, the problem was infeasible when the notion of multivariate positive linear dominance (i.e., $\mathcal{V} = \mathbb{R}^n_+$) was used.

In this paper we propose alternative notions of stochastic dominance that address the shortcomings discussed above. This is accomplished by introducing a probability distribution indicating the relative importance of each vector of criterion weights, which is then viewed as a random vector. In the case of the hierarchical group decision, the probability associated with a particular vector of criterion weights may represent the influence of a member of the group. The introduction of such a distribution motivates us to propose three new concepts of *stochastically weighted dominance*, where the vector of weights is random. The new concepts yield more flexibility to the modeling process in two ways: first, by allowing for a representation of the importance of each vector of criterion weights; second, by providing a relaxation on the notion of multivariate linear stochastic dominance, which in the context of optimization problems means enlarging the feasibility region. In addition, we show that the proposed new concepts of stochastic dominance can be represented by a finite number of (mixed-integer) linear inequalities when the distributions of X, Y and of the criterion weight vector have finite support, which allows for easy verification of the dominance conditions both in case of stand-alone comparisons as well as when dominance is used as a constraint of an optimization problem. We also illustrate, by means of an example, how the proposed new concepts can be valuable for multi-criterion decision making by allowing for a (partial) ranking of the alternatives.

The remainder of this paper is organized as follows. In Section 2 we review some concepts of stochastic dominance, including the notion of multivariate linear dominance under a robust weight region, and provide some basic results. In Section 3 we describe new concepts of stochastically weighted dominance and provide their equivalent utility representations. A numerical example is presented in Section 4 to show the relationship between the previously studied robust approach and the concepts of stochastically weighted dominance. In Section 5 we present linear and mixed-integer linear systems of inequalities corresponding to the new dominance relationships. In Section 6 we present two applications: the first one illustrates the effect of the various concepts of stochastic dominance on the feasibility region of an optimization problem, whereas the second one describes a staffing problem where the goal is to allocate workers between two $M/M/c$ queueing stations to minimize waiting times. We show how the new concepts of stochastic dominance proposed in this paper can be used to compare the various possible configurations.

2 (Robust Weighted) Stochastic Dominance

We start by reviewing the concept of nth order stochastic dominance (see e.g. Dentcheva and Ruszczyński 2003). Consider an underlying probability space $(\Omega, \mathcal{F}, P)$ and denote by $\mathcal{L}_n^m$ the space of m-dimensional random vectors X such that $\mathbb{E}[\|X\|^n]$ is finite (we omit the superscript if $m = 1$). Consider a real-valued random variable ξ . We write the cumulative distribution function (cdf) of ξ as

$$F_1(\xi; \eta) := P(\xi \ \leq \ \eta).$$

Furthermore, for $\xi \in \mathcal{L}_{n-1}$ where $n \geq 2$, define recursively the functions

$$F_j(\xi; \eta) := \int_{-\infty}^{\eta} F_{j-1}(\xi; t)\, dt, \quad j = 2, \ldots, n.$$

It is useful to note the equivalence (see Rachev et al. 2008)

$$F_j(\xi; \eta) \ = \ \frac{1}{(j-1)!}\mathbb{E}\left[((\eta - \xi)_+)^{j-1}\right], \quad j = 2, \ldots, n, \tag{2.1}$$

which will be helpful later. To see why (2.1) holds, consider the well-known identity $F_2(\xi, \eta) = \mathbb{E}[(\eta - \xi)_+]$, which follows by changing the order of integration (see, e.g., Müller and Stoyan 2002). Suppose $F_j(\xi, \eta) = \frac{1}{(j-1)!}\mathbb{E}[((\eta - \xi)_+)^{j-1}]$. It then follows that

$$\begin{aligned}
F_{j+1}(\xi, \eta) \ &= \ \int_{-\infty}^{\eta} \frac{1}{(j-1)!}\mathbb{E}\left[((t - \xi)_+)^{j-1}\right]\, dt \\
&= \ \frac{1}{(j-1)!}\mathbb{E}\left[\int_{-\infty}^{\eta} [(t - \xi)_+]^{j-1}\, dt\right] \\
&= \ \frac{1}{(j-1)!}\mathbb{E}\left[\mathbf{1}\{\xi \leq \eta\} \int_{\xi}^{\eta} (t - \xi)^{j-1}\, dt\right] \\
&= \ \frac{1}{j!}\mathbb{E}\left[((\eta - \xi)_+)^{j}\right].
\end{aligned}$$

It follows from (2.1) that, for any n, the condition $\xi \in \mathcal{L}_{n-1}$ suffices to ensure that $F_n(\xi; \eta) < \infty$ for all η.

A random variable $\xi \in \mathcal{L}_{n-1}$ is said to stochastically dominate another random variable $\zeta \in \mathcal{L}_{n-1}$ in nth order if

$$F_n(\xi; \eta) \ \leq \ F_n(\zeta; \eta), \quad \forall \eta \in \mathbb{R}. \tag{2.2}$$

We use the notation $\xi \trianglerighteq_{(n)} \zeta$ to indicate this relationship.

At this point it is useful to establish the connection between the notion of the first (second) stochastic dominance just defined (for $n = 1, 2$) and the interpretation via utility functions described in Section 1. It is well known that

$$\xi \trianglerighteq_{(n)} \zeta \iff \mathbb{E}[u(\xi)] \geq \mathbb{E}[u(\zeta)], \quad \forall u \in \mathcal{U}_n, \tag{2.3}$$

whenever the expectations exist. In the above statement, $\mathcal{U}_1$ is the set of all nondecreasing functions $u : \mathbb{R} \mapsto \mathbb{R}$ and $\mathcal{U}_2$ is the set of all nondecreasing concave functions $u : \mathbb{R} \mapsto \mathbb{R}$. We refer to Müller and Stoyan (2002), Dentcheva and Ruszczyński (2003) and Rachev et al. (2008) for a more detailed discussion of stochastic dominance. To avoid repetition of the condition "whenever the expectations exist" we shall always understand (2.3) in that context.

The concept of nth order dominance can be readily extended to the multivariate linear case as defined below.

Definition 2.1 *A random vector $X \in \mathcal{L}_{n-1}^m$ dominates another random vector $Y \in \mathcal{L}_{n-1}^m$ in the linear nth order with respect to set $\mathcal{V} \subseteq \mathbb{R}_+^m$ if*

$$F_n(v^T X; \eta) \ \leq \ F_n(v^T Y; \eta) \quad \forall \eta \in \mathbb{R}, \ \forall v \in \mathcal{V}. \tag{RSD}$$

Since the above inequality must hold for all $v \in \mathcal{V}$, we shall sometimes refer to this relationship as "robust weighted" stochastic dominance.

In Theorem 2.4 below we show that without loss of generality we can assume that $\mathcal{V} \subseteq \Delta := \{v \in \mathbb{R}^m : \|v\|_1 \leq 1\}$. This property generalizes the results in Homem-de-Mello and Mehrotra (2009) who consider the case of second order dominance. Before we state the theorem, we prove some auxiliary results.

Lemma 2.2 *For any $\alpha > 0$, any $v \in \mathcal{V}$ and $X \in \mathcal{L}_{n-1}^m$,*

$$F_n(\alpha v^T X; \eta) = \alpha^{n-1} F_n(v^T X; \eta/\alpha), \quad \forall \eta \in \mathbb{R}.$$

Proof: It is trivial to confirm that the equality holds for first order dominance. Suppose that the relationship is satisfied for the nth order dominance. We now prove this for the $(n+1)$th order dominance. Note that

$$\begin{aligned} F_{n+1}(\alpha v^T X; \eta) =& \int_{-\infty}^{\eta} F_n(\alpha v^T X; t) dt \\ =& \alpha^{n-1} \int_{-\infty}^{\eta} F_n(v^T X; t/\alpha) dt \\ =& \alpha^n \int_{-\infty}^{\eta/\alpha} F_n(v^T X; t) dt \\ =& \alpha^n F_{n+1}(v^T X; \eta/\alpha). \end{aligned}$$

□

Lemma 2.3 *For $n \geq 2$, let $X \in \mathcal{L}_{n-1}^m$ and let $\{v^k\} \subset \Delta$ be a convergent sequence. Let $\bar{v}$ denote the limit point. Then for all $\eta \in \mathbb{R}$,*

$$\lim_{k \to \infty} F_n\left((v^k)^T X; \eta\right) \ = \ F_n\left(\bar{v}^T X; \eta\right).$$

Proof: For any $v \in \widetilde{\mathcal{V}}$, we have $|v^T X| \leq \|v\|_1 \|X\|_\infty = \|X\|_\infty$ a.s. and thus $v^T X \geq -\|X\|_\infty$ a.s.. It further follows that $0 \leq F_j(v^T X; t) \leq F_j(-\|X\|_\infty; t)$ $(j = 1, \ldots, n-1)$ for all $t \in \mathbb{R}$. Note that, for any fixed $\eta \in \mathbb{R}$, $\int_{-\infty}^{\eta} F_j(-\|X\|_\infty; t)\, dt$ is finite since $X \in \mathcal{L}_{n-1}^m$.

Let us first consider the case with $n = 2$. Since $v^k \to \bar{v}$, it follows that ${v^k}^T X \to \bar{v}^T X$ a.s. and thus ${v^k}^T X$ converge in distribution to $\bar{v}^T X$. Hence, $F_1({v^k}^T X; \eta) \to F_1(\bar{v}^T X; \eta)$ for all continuity points of $F_1(\bar{v}^T X;\cdot)$. It follows from the Lebesgue's dominated convergence theorem (see Billingsley (1995)) that $F_2({v^k}^T X;\eta) \to F_2(\bar{v}^T X;\eta)$ for all $\eta \in \mathbb{R}$.

Next, suppose that $F_j({v^k}^T X;\eta) \to F_j(\bar{v}^T X;\eta)$ for all $\eta \in \mathbb{R}$ for some j, $2 \le j \le n-1$. Using again the dominated convergence theorem, we have that $F_{j+1}({v^k}^T X;\eta) \to F_{j+1}(\bar{v}^T X;\eta)$. The proof is complete by induction. □

Theorem 2.4 *Let $\mathcal{V}$ be a non-empty convex set and $\widetilde{\mathcal{V}} := \mathrm{cone}(\mathcal{V}) \cap \Delta$ (*cone *denotes the conical hull). Then,*

(i). $v^T X \trianglerighteq_{(1)} v^T Y$ for all $v \in \mathcal{V}$ if and only if $v^T X \trianglerighteq_{(1)} v^T Y$ for all $v \in \widetilde{\mathcal{V}}$.

*(ii). For $n \ge 2$, $v^T X \trianglerighteq_{(n)} v^T Y$ for all $v \in \mathcal{V}$ if and only if $v^T X \trianglerighteq_{(n)} v^T Y$ for all $v \in \mathrm{cl}(\widetilde{\mathcal{V}})$ (*cl *denotes the closure of a set).*

Proof: For a given $\alpha > 0$ and $v \in \mathbb{R}^m$, we have $v^T X \trianglerighteq_{(n)} v^T Y$ if and only if $\alpha v^T X \trianglerighteq_{(n)} \alpha v^T Y$ by Lemma 2.2. Since the conical hull of a convex set $\mathcal{V}$ can be written as $\mathrm{cone}(\mathcal{V}) = \{\alpha v : v \in \mathcal{V}, \alpha \in \mathbb{R}_+\}$, it follows that $v^T X \trianglerighteq_{(n)} v^T Y$ for all $v \in \mathcal{V}$ if and only if $v^T X \trianglerighteq_{(n)} v^T Y$ for all $v \in \mathrm{cone}(\mathcal{V})$. Thus, the proof of (i) for $n = 1$ is complete since $\mathrm{cone}(\mathcal{V}) = \mathrm{cone}(\widetilde{\mathcal{V}})$.

When $n \ge 2$ we can strengthen the result. To prove (ii), all we need to show is that $X \trianglerighteq_{(n)} Y$ for all $v \in \widetilde{\mathcal{V}}$ is equivalent to $X \trianglerighteq_{(n)} Y$ for all $v \in \mathrm{cl}(\widetilde{\mathcal{V}})$. If $X \trianglerighteq_{(n)} Y$ for all $v \in \mathrm{cl}(\widetilde{\mathcal{V}})$, it follows trivially that $X \trianglerighteq_{(n)} Y$ for all $v \in \widetilde{\mathcal{V}}$. Let us prove the other direction. Suppose $\bar{v} \in \mathrm{cl}(\widetilde{\mathcal{V}})$. Then there exists a sequence $\{v^k\} \subset \widetilde{\mathcal{V}}$ such that $v^k \to \bar{v}$. It follows that $F_n({v^k}^T X;\eta) \le F_n({v^k}^T Y;\eta)$ for all $\eta \in \mathbb{R}$. By Lemma 2.3, this implies that $F_n(\bar{v}^T X;\eta) \le F_n(\bar{v}^T Y;\eta)$ for all η, i.e., $\bar{v}^T X \trianglerighteq_{(n)} \bar{v}^T Y$.
□

3 Stochastically Weighted Dominance

We now introduce a few alternative notions of stochastically weighted linear stochastic dominance. As mentioned in Section 1, the basic idea is to consider the vector of weights v as a random vector with a known distribution. We shall use the notation V instead of v to emphasize that the weights are random, and assume that V is an m-dimensional random vector defined on the same probability space as X and Y. Define the conditional distribution function

$$F_1(V^T X;\eta \,|\, V) := P(V^T X \le \eta \,|\, V).$$

Also for $X \in \mathcal{L}^m_{n-1}$, denote

$$F_j(V^T X;\eta \,|\, V) := \int_{-\infty}^{\eta} F_{j-1}(V^T X; t \,|\, V)\,dt, \quad j = 2,\ldots,n.$$

Note that $F_j(V^T X;\eta \,|\, V)$ is a nonnegative random variable. Moreover, from (2.1) we have that

$$F_j(V^T X;\eta \,|\, V) \;=\; \mathbb{E}\left[\left((\eta - V^T X)_+\right)^{j-1} \,|\, V\right], \quad \forall \eta \in \mathbb{R} \quad \text{a.s..} \tag{3.1}$$

Let $X, Y \in \mathcal{L}^m_{n-1}$. We first define the following relationship, which we call *strong stochastically* *weighted nth order dominance.* Under this definition, X dominates Y if

$$F_n(V^T X; \eta \,|\, V) \;\le\; F_n(V^T Y; \eta \,|\, V), \quad \forall \eta \in \mathbb{R} \quad \text{a.s..} \tag{3.2}$$

Note that the above concept is similar to the robust notion in Definition 2.1. Indeed, $v^T X \trianglerighteq_{(n)} v^T Y$ for all v in the support set of V implies (3.2). However, in general, the robust condition is not implied by (3.2) unless the support of V is finite. Since

$$\lim_{\eta\to-\infty} F_n(V^T X; \eta \,|\, V) = \lim_{\eta\to-\infty} F_n(V^T Y; \eta \,|\, V) = 0, \quad \text{a.s.},$$

it follows that $\sup_{\eta\in\mathbb{R}} F_n(V^T X; \eta \,|\, V) - F_n(V^T Y; \eta \,|\, V) \ge 0$ a.s. and hence (3.2) can be equivalently represented as

$$\mathbb{E}\left[\sup_{\eta\in\mathbb{R}} F_n(V^T X; \eta \,|\, V) - F_n(V^T Y; \eta \,|\, V)\right] \;\le\; 0. \tag{3.3}$$

The strong stochastically weighted dominance does not provide much relaxation to the robust relationship in Definition 2.1, which was one of the goals laid out in Section 1. This motivates the study of weak versions of (3.2) and its equivalent representation (3.3).

We proceed now in that direction. Note that, by Theorem 2.4, definitions (3.2) and (3.3) remain unchanged with respect to scaling the random vector V. That is, we can normalize V as $V/\|V\|_1$ if $\|V\|_1 > 0$ and zero otherwise so that a nonnegative support of V can be projected onto Δ. Thus, we shall assume henceforth that the support of V is a subset of Δ, which we will indicate by $V \in \mathcal{L}^m(\Delta)$ in the sequel.

3.1 Weak Stochastically Weighted Dominance

The first relaxation of (3.3) we study arises from the observation that the supremum of an expectation is less than or equal to the expectation of the supremum. Thus, we obtain a relaxed version of (3.3) by switching the order of expectation and supremum. This is formally described in the definition below.

Definition 3.1 *A random vector $X \in \mathcal{L}^m_{n-1}$ is said to dominate another random vector $Y \in \mathcal{L}^m_{n-1}$ in* weak stochastically weighted nth order *with respect to a random vector $V \in \mathcal{L}^m(\Delta)$ (written $X \underline{\blacktriangleright}^w_{(n)} Y$) if*

$$\mathbb{E}\left[F_n(V^T X; \eta \,|\, V)\right] \;\le\; \mathbb{E}\left[F_n(V^T Y; \eta \,|\, V)\right], \quad \forall \eta \in \mathbb{R}. \tag{SWD-Weak}$$

At a fixed realization v of V, the condition $F_n(v^T X; \eta) - F_n(v^T Y; \eta) \le 0$ indicates the dominance of $v^T X$ over $v^T Y$. Thus, (SWD-Weak) requires that the dominance relationship hold *on average* with respect to the distribution of V. This is a much weaker condition than (3.3), which requires that $v^T X$ dominate $v^T Y$ for almost all realizations v.

Note that, by applying identity (3.1) into (SWD-Weak) and using the fact that $\mathbb{E}[\mathbb{E}[\cdot \,|\, V]] = \mathbb{E}[\cdot]$, we obtain an important characteristic of the weak stochastically weighted dominance — namely, that this multivariate dominance can be represented by a univariate relationship. For reference, we state this result in Proposition 3.2 below.

Proposition 3.2 *For random vectors $X, Y \in \mathcal{L}^m_{n-1}$ and $V \in \mathcal{L}^m(\Delta)$, let $\xi := V^T X$ and $\zeta := V^T Y$. Then $X \trianglerighteq^w_{(n)} Y$ if and only if $\xi \trianglerighteq_{(n)} \zeta$.*

Consequently, properties of univariate dominance can be used for (SWD-Weak). This is particularly important from an algorithmic perspective, since it is much easier to test for univariate dominance than for robust multivariate linear dominance. We will discuss this issue in more detail in Section 5.

By combining Proposition 3.2 with (2.3), we obtain the following equivalent utility representations of weak stochastically weighted dominance for the first and second order. Recall the definition of $\mathcal{U}_1$ and $\mathcal{U}_2$ following (2.3).

Corollary 3.3 *For random vectors $X, Y \in \mathcal{L}^m_1$ and $V \in \mathcal{L}^m(\Delta)$. Then, for $n \in \{1, 2\}$, $X \trianglerighteq^w_{(n)} Y$ if and only if $\mathbb{E}[u(V^T X)] \leq \mathbb{E}[u(V^T Y)]$ for all $u \in \mathcal{U}_n$.*

3.2 Stochastically Weighted Dominance with Chance

We now discuss another relaxation of (3.3), in which we require that the relationship in (3.2) hold with some probability rather than almost surely. One can draw a parallel between this approach and the traditional notion of chance constraints (also called probabilistic constraints) in stochastic programming where, instead of imposing that a (random) constraint hold for almost every point in the sample space, it is required that the constraint hold on a set of positive probability $1 - \alpha$. Formally, we have the following definition:

Definition 3.4 *A random vector $X \in \mathcal{L}^m_{n-1}$ is said to dominate another random vector $Y \in \mathcal{L}^m_{n-1}$ in* stochastically weighted nth order with chance $\alpha \in [0, 1]$ *with respect to a random vector $V \in \mathcal{L}^m(\Delta)$ (written $X \trianglerighteq^{c(\alpha)}_{(n)} Y$) if*

$$P\left(F_n(V^T X; \eta \,|\, V) \leq F_n(V^T Y; \eta \,|\, V),\ \forall \eta \in \mathbb{R}\right) =: \mathfrak{S}(X, Y) \ \geq \ 1 - \alpha. \qquad \text{(SWD-Chance)}$$

The parameter α allows for the control of the degree of relaxation of (3.3). Indeed, $\alpha = 0$ is the same as (3.2); on the other extreme, $\alpha = 1$ corresponds to complete relaxation, so that (SWD-Chance) trivially holds regardless of the distributions of X, Y or V. In this paper, we call $\mathfrak{S}(X, Y)$ the *satisfaction level* of preference of X over Y, since $X \trianglerighteq^{c(1-\mathfrak{S}(X,Y))}_{(n)} Y$.

The following proposition gives an equivalent representation of stochastically weighted dominance with chance α in terms of expected utilities. The proof follows directly from the equivalence (2.3) applied to the univariate random variables $\hat{v}^T X$ and $\hat{v}^T Y$ for a fixed realization $\hat{v}$ of V.

Proposition 3.5 *For random vectors X, $Y \in \mathcal{L}^m_1$ and $V \in \mathcal{L}^m(\Delta)$. Then, for $n \in \{1, 2\}$, $X \trianglerighteq^{c(\alpha)}_{(n)} Y$ if and only if*

$$P\left(\mathbb{E}[u(V^T X)|V] \geq \mathbb{E}[u(V^T Y)|V],\ \forall u \in \mathcal{U}_n\right) \ \geq \ 1 - \alpha.$$

In Section 5 we will discuss how to test for this type of dominance.

3.3 Relaxed Strong Stochastically Weighted Dominance

The third relaxation of (3.3) we present is obtained by allowing for a tolerance on the right hand side of that inequality, as defined below.

Definition 3.6 *A random vector* $X \in \mathcal{L}^m_{n-1}$ *is said to dominate another random vector* $Y \in \mathcal{L}^m_{n-1}$ *in* relaxed strong stochastically weighted nth order for a given parameter $\gamma \geq 0$ *with respect to a random vector* $V \in \mathcal{L}^m(\Delta)$ *(written* $X \underline{\blacktriangleright}^{s(\gamma)}_{(n)} Y$*) if*

$$\mathbb{E}\left[\sup_{\eta\in\mathbb{R}} F_n(V^T X;\eta\,|\,V) - F_n(V^T Y;\eta\,|\,V)\right] =: \mathfrak{T}(X,Y) \;\leq\; \gamma. \tag{SWD-Relax}$$

We call $\mathfrak{T}(X,Y)$ the *tolerance level* for X to dominate Y since $X \underline{\blacktriangleright}^{s(\mathfrak{T}(X,Y))}_{(n)} Y$.

Further properties of (SWD-Relax) can be derived when both X and Y have bounded support. Suppose this is the case, and without loss of generality assume that $X, Y \in [0,1]^m$ a.s.. Then, the supremum in $\mathfrak{T}(X,Y)$ can be taken over η in $[0,1]$, i.e.,

$$\mathfrak{T}(X,Y) = \mathbb{E}\left[\sup_{\eta\in[0,1]} F_n(V^T X;\eta\,|\,V) - F_n(V^T Y;\eta\,|\,V)\right]. \tag{3.4}$$

Note that $F_n(V^T X;\eta\,|\,V) \leq 1$ for all $\eta \in [0,1]$ a.s.. Thus, we see that, similarly to the case of stochastically weighted dominance with chance α, the parameter γ allows for the control of the degree of relaxation of (3.3). Obviously, $\gamma = 0$ is the same as (3.3), whereas (3.4) is always satisfied with $\gamma = 1$.

In parallel with the other concepts of stochastically weighted linear stochastic dominance introduced above, we now derive an equivalent representation of (3.4) in terms of expected utilities. Theorem 3.7 below states the result.

Theorem 3.7 *Let* X*,* Y *be random vectors with support in* $[0,1]^m$ *and* $V \in \mathcal{L}^m_0(\Delta)$*. Then, for* $n \in \{1,2\}$*,* $X \underline{\blacktriangleright}^{s(\gamma)}_{(n)} Y$ *if and only if*

$$\mathbb{E}\left[\inf_{u\in\widetilde{\mathcal{U}}_n} \mathbb{E}\left[u(V^T X) - u(V^T Y) \mid V\right]\right] \;\geq\; -\gamma, \tag{3.5}$$

where $\widetilde{\mathcal{U}}_1$ *is the set of increasing functions* u *on* $[0,1]$ *with* $u(1) - u(0) \leq 1$ *and* $\widetilde{\mathcal{U}}_2$ *is the set of increasing continuous concave functions* u *on* $[0,1]$ *with* $u'(0_+) \leq 1$ *(*$u'(0_+)$ *is the right derivative of* $u(\cdot)$ *at 0).*

Proof:

Let us first prove the statement for $n = 1$. Note that, for a given realization $\hat{v}$ of V,

$$F_1(V^T X;\eta \mid V = \hat{v}) \;=\; \mathbb{E}\left[\mathbf{1}\{V^T X \leq \eta\} \mid V = \hat{v}\right]. \tag{3.6}$$

Let $\mathbb{E}_{\hat{v}}[\cdot]$ denote the expectation with respect to the conditional distribution $P(\cdot\,|\,V = \hat{v})$, so the term on the right hand side of (3.6) is written $\mathbb{E}_{\hat{v}}\left[\mathbf{1}\{V^T X \leq \eta\}\right]$. We now claim that

$$\sup_{\eta\in[0,1]} \mathbb{E}_{\hat{v}}\left[\mathbf{1}\{\hat{v}^T X \leq \eta\} - \mathbf{1}\{\hat{v}^T Y \leq \eta\}\right] \;=\; \sup_{u\in\widetilde{\mathcal{U}}_1} \mathbb{E}_{\hat{v}}\left[u(\hat{v}^T Y) - u(\hat{v}^T X)\right]. \tag{3.7}$$

It is easy to see that the right hand side of (3.7) is greater than or equal to the left hand side, since the function $u_\eta(x) := 1 - \mathbf{1}\{x \le \eta\}$ for $\eta \in [0,1]$ is in $\widetilde{\mathcal{U}}_1$. We now prove the other direction. Let

$$\tau := \sup_{\eta \in [0,1]} \mathbb{E}_{\hat{v}} \left[\mathbf{1}\{\hat{v}^T X \le \eta\} - \mathbf{1}\{\hat{v}^T Y \le \eta\} \right], \tag{3.8}$$

and note that τ is clearly finite. We need to show

$$\sup_{u \in \widetilde{\mathcal{U}}_1} \mathbb{E}_{\hat{v}}[u(\hat{v}^T Y) - u(\hat{v}^T X)] \;\le\; \tau. \tag{3.9}$$

Now construct a sequence of functions u_n to approximate $u \in \widetilde{\mathcal{U}}_1$. We introduce $n+1$ discrete real points

$$\eta_j := j/n, \quad j = 0, \dots, n,$$

and define

$$\begin{aligned} c_n &:= u(1), \\ a_j &:= u(\eta_j) - u(\eta_{j-1}), \quad j = 1, \dots, n. \end{aligned}$$

Note that $a_j \ge 0$ because $u(\cdot)$ is nondecreasing. Define the function

$$u_n(t) := c_n - \sum_{j=1}^n a_j \mathbf{1}\{t \le \eta_j\}.$$

Then we have

$$u_n(\hat{v}^T X) - u_n(\hat{v}^T Y) = -\sum_{j=1}^n a_j (\mathbf{1}\{\hat{v}^T X - \eta_j\} - \mathbf{1}\{\hat{v}^T Y - \eta_j\}).$$

By (3.8), it follows that

$$\begin{aligned} \mathbb{E}_{\hat{v}} \left[u_n(\hat{v}^T X) \right] &\ge \mathbb{E}_{\hat{v}} \left[u_n(\hat{v}^T Y) \right] - \sum_{j=1}^n a_j \tau \\ &= \mathbb{E}_{\hat{v}} \left[u_n(\hat{v}^T Y) \right] - \tau(u(1) - u(0)) \\ &\ge \mathbb{E}_{\hat{v}} \left[u_n(\hat{v}^T Y) \right] - \tau. \end{aligned}$$

Note that u_n is an approximation of u by the step function with points $(\eta_j, u(\eta_j))$ $(j = 1, \dots, n)$. Thus, it follows from the bounded convergence theorem (see Billingsley (1995)) that

$$\lim_{n \to \infty} \mathbb{E}_{\hat{v}} \left[u_n(\hat{v}^T X) \right] \;=\; \mathbb{E}_{\hat{v}} \left[u(\hat{v}^T X) \right].$$

Therefore, it follows that (3.9) (and hence (3.7)) holds. Since (3.7) holds for each realization $\hat{v}$ of V, we have shown the equivalence between $X \underline{\blacktriangleright}^{s(\gamma)}_{(1)} Y$ and (3.5) for the case $n = 1$.

Next, we show that (3.5) holds for the case $n = 2$. By (2.1), we have, for any realization $\hat{v}$ of V,

$$F_2(V^T X; \eta \,|\, V = \hat{v}) \;=\; \mathbb{E}\left[(\eta - V^T X)_+ \,|\, V = \hat{v} \right] \;=\; \mathbb{E}_{\hat{v}} \left[(\eta - \hat{v}^T X)_+ \right]. \tag{3.10}$$

We claim that

$$\sup_{\eta\in[0,1]} \mathbb{E}_{\hat{v}}\left[(\eta-\hat{v}^TX)_+ - (\eta-\hat{v}^TY)_+\right] \;=\; \sup_{u\in\widetilde{\mathcal{U}}_2} \mathbb{E}_{\hat{v}}\left[u(\hat{v}^TY)-u(\hat{v}^TX)\right]. \tag{3.11}$$

The proof of (3.11) is similar to that of (3.7), except that we use different constructions to approximate the utility functions. It is easy to check that the right hand side of (3.11) is greater than or equal to the left hand side, since the function $u_\eta(x) := -(\eta-x)_+$ is in $\widetilde{\mathcal{U}}_2$ for all η. We prove the other direction. Let

$$\tau := \sup_{\eta\in[0,1]} \mathbb{E}_{\hat{v}}\left[(\eta-\hat{v}^TX)_+ - (\eta-\hat{v}^TY)_+\right]. \tag{3.12}$$

What we need is show

$$\sup_{u\in\widetilde{\mathcal{U}}_2} \mathbb{E}_{\hat{v}}\left[u(\hat{v}^TY)-u(\hat{v}^TX)\right] \;\leq\; \tau. \tag{3.13}$$

Again, to construct a sequence of of functions u_n to approximate $u\in\widetilde{\mathcal{U}}_2$, we introduce $n+1$ discrete real points

$$\eta_j := j/n, \quad j=0,\ldots,n,$$

and define

$$\begin{aligned} c_n &:= u(1),\\ a_i &:= \begin{cases} \frac{u(\eta_j)-u(\eta_{j-1})}{\eta_j-\eta_{j-1}} & j=n,\\ \frac{u(\eta_j)-u(\eta_{j-1})}{\eta_j-\eta_{j-1}} - \frac{u(\eta_{j+1})-u(\eta_j)}{\eta_{j+1}-\eta_j} & j=n-1,\ldots,1. \end{cases} \end{aligned}$$

Define the function

$$u_n(t) := c_n - \sum_{j=1}^{n} a_j(\eta_j-t)_+.$$

It follows from (3.12) that

$$\begin{aligned} \mathbb{E}_{\hat{v}}\left[u_n(\hat{v}^TX)\right] \;&\geq\; \mathbb{E}_{\hat{v}}\left[u_n(\hat{v}^TY)\right] - \sum_{j=1}^{n} a_j\tau \\ &\geq\; \mathbb{E}_{\hat{v}}\left[u_n(\hat{v}^TY)\right] - \tau n(u(1/n)-u(0)). \end{aligned}$$

By construction, u_n is a piecewise linear approximation of u with points $(\eta_j, u(\eta_j))$ $(j=1,\ldots,n)$. Thus, it follows from the bounded convergence theorem that

$$\lim_{n\to\infty} \mathbb{E}_{\hat{v}}\left[u_n(\hat{v}^TX)\right] \;=\; \mathbb{E}_{\hat{v}}\left[u(\hat{v}^TX)\right].$$

Moveover,

$$\lim_{n\to\infty} n(u(1/n)-u(0)) = u'(0^+) \leq 1.$$

Therefore, it follows that (3.13) holds and hence (3.11) holds as well. Thus, we have shown the equivalence between $X \trianglerighteq^{\mathrm{s}(\gamma)}_{(2)} Y$ and (3.5) for $n=2$. $\square$

At first sight, it appears that Theorem 3.7 restricts the class of utility functions that can be considered due to the conditions $u(1)-u(0)\leq 1$ and $u'(0_+)\leq 1$. The corollaries below show that this is not the case, since arbitrary utility functions in $[0,1]$ can be normalized to satisfy those conditions.

Corollary 3.8 *Let X, Y be random vectors with support in $[0,1]^m$ and $V \in \mathcal{L}_0^m(\Delta)$. Let $\mathcal{U}_1(0,1)$ be the set of nondecreasing functions on $[0,1]$. Define a functional l_1 on $\mathcal{U}(0,1)$ as*

$$l_1(u)(\cdot) := \begin{cases} \frac{u(\cdot)}{u(1)-u(0)} & u(1)-u(0) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then, $X \underline{\blacktriangleright}_{(1)}^{s(\gamma)} Y$ is equivalent to

$$\mathbb{E}\left[\inf_{u \in \mathcal{U}_1(0,1)} \mathbb{E}\left[l_1(u)(V^T X) - l_1(u)(V^T Y) \mid V\right]\right] \geq -\gamma. \tag{3.14}$$

Proof: Let $\widetilde{\mathcal{U}}_1$ be as defined in Theorem 3.7. Given any $u \in \mathcal{U}_1(0,1)$, we have that $l_1(u) \in \widetilde{\mathcal{U}}_1$. It follows from Theorem 3.7 that the condition $X \underline{\blacktriangleright}_{(1)}^{s(\gamma)} Y$ guarantees (3.14). Let us again write $X \underline{\blacktriangleright}_{(1)}^{s(\gamma)} Y$ as

$$\mathbb{E}\left[\sup_{\eta \in [0,1]} \mathbb{E}\left[\mathbf{1}\{V^T X \leq \eta\} - \mathbf{1}\{V^T Y \leq \eta\} | V\right]\right] \ \leq \ \gamma,$$

which is satisfied if (3.14) holds, since the function $u_\eta(x) = 1 - \mathbf{1}\{x \leq \eta\}$ for $\eta \in [0,1]$ is in $\mathcal{U}_1(0,1)$ and $l_1(u_\eta)(x) \equiv u_\eta(x)$. □

Corollary 3.9 *Let X, Y be random vectors with support in $[0,1]^m$ and $V \in \mathcal{L}_0^m(\Delta)$. Let $\mathcal{U}_2(0,1)$ be the set of nondecreasing continuous concave functions on $[0,1]$. Define a functional l_2 on $\mathcal{U}_2(0,1)$ as*

$$l_2(u)(\cdot) := \begin{cases} \frac{u(\cdot)}{u'(0^+)} & u'(0^+) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then $X \underline{\blacktriangleright}_{(2)}^{s(\gamma)} Y$ is equivalent to

$$\mathbb{E}\left[\inf_{u \in \mathcal{U}_2(0,1)} \mathbb{E}\left[l_2(u)(V^T X) - l_2(u)(V^T Y) \mid V\right]\right] \ \geq \ -\gamma. \tag{3.15}$$

Proof: Let $\widetilde{\mathcal{U}}_2$ be as defined in Theorem 3.7. Given any $u \in \mathcal{U}_2(0,1)$, we have that $l_2(u) \in \widetilde{\mathcal{U}}_2$. It follows from Theorem 3.7 that the condition $X \underline{\blacktriangleright}_{(2)}^{s(\gamma)} Y$ guarantees (3.15). Let us again write $X \underline{\blacktriangleright}_{(2)}^{s(\gamma)} Y$ as

$$\mathbb{E}\left[\sup_{\eta \in [0,1]} \mathbb{E}\left[(\eta - V^T X)_+ - (\eta - V^T Y)_+ | V\right]\right] \ \leq \ \gamma,$$

which is satisfied if (3.15) holds, since the function $u_\eta(x) = -(\eta - x)_+$ for $\eta \in [0,1]$ is in $\mathcal{U}_2(0,1)$ and $l_2(u_\eta)(x) \equiv u_\eta(x)$. □

4 An Illustrative Example

We present now a simple example to compare the robust weighted dominance in (2.1) with the various stochastically weighted dominance relations proposed in Section 3. We use the first order version of all dominance relationships for the calculations.

Consider a two-dimensional random vector $X(\theta)$ parameterized by $\theta \in [0,1]$. Let its two components, $X_1(\theta)$ and $X_2(\theta)$, be independent random variables uniformly distributed in $(0,\theta)$ and $(\theta,1)$ respectively (for $\theta = 0$, $X_1(\theta)$ has point mass at 0; for $\theta = 1$, $X_2(\theta)$ has point mass at 1). Let Y_1 and Y_2 be i.i.d. random variables with uniform distribution in $(0,1)$. Figure 1 depicts the cumulative distribution functions for the components of $X(\theta)$ and Y, which can be expressed as

$$F_1(X_1(\theta);\eta) = \begin{cases} 0 & \eta < 0 \\ \frac{\eta}{\theta} & 0 \le \eta < \theta \\ 1 & \text{otherwise} \end{cases} \tag{4.1}$$

$$F_1(X_2(\theta);\eta) = \begin{cases} 0 & \eta < \theta \\ \frac{\eta-\theta}{(1-\theta)} & \theta \le \eta < 1 \\ 1 & \text{otherwise} \end{cases} \tag{4.2}$$

$$F_1(Y_1;\eta) = F_1(Y_2;\eta) = \begin{cases} 0 & \eta < 0 \\ \eta & 0 \le \eta < 1 \\ 1 & \text{otherwise} \end{cases} \tag{4.3}$$

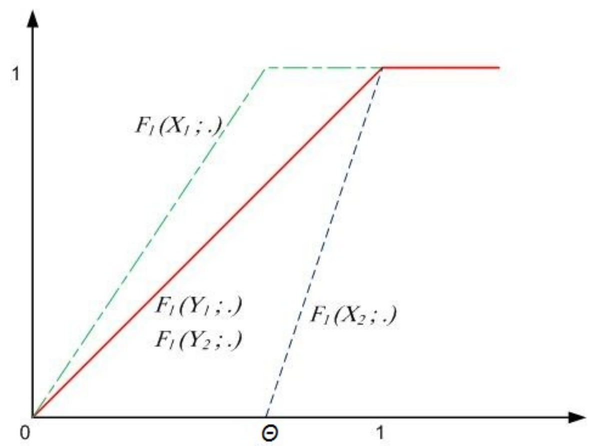


Figure 1: Probability Distribution Functions of X and Y

We are interested in comparing the random vectors $X(\theta)$ and Y. More precisely, the goal of the study is to determine the set Θ of values of θ for which $X(\theta)$ dominates Y in terms of the dominance relationships described in Section 3. By comparing the size of the set Θ obtained for the various dominance concepts we can then illustrate the differences between them.

We start with the notion of robust weighted dominance described in Section 2. Let $\mathcal{V}$ be the set $\{v^1,\ v^2\}$, where $v^1 = (1\ \ 0)^T$ and $v^2 = (0\ \ 1)^T$. It is easy to see that the first order dominance $X(\theta) \trianglerighteq^{\mathcal{V}}_{(1)} Y$ — which translates as $X_1(\theta) \trianglerighteq_{(1)} Y_1$ *and* $X_2(\theta) \trianglerighteq_{(1)} Y_2$ — is satisfied only when $\theta = 1$, since $F_1(X_1(\theta);\eta) > F_1(Y_1;\eta)$ for $\eta \in (0,1)$. Thus, $\Theta = \{1\}$ in this case.

Let us look at the conditions for the stochastically weighted dominance relations proposed in Section 3. Suppose we choose the distribution of the random weight vector V as $V = v^1$ with probability $p \in (0,1)$ and $V = v^2$ with probability $1-p$, where v^1 and v^2 are defined as above.

Consider initially the weak stochastically weighted dominance relation $X(\theta) \underline{\blacktriangleright}^{\mathrm{w}}_{(1)} Y$. When $0 <$

$\theta < 1$, inequality (SWD-Weak) is written as

$$
\begin{aligned}
&\sup_{\eta\in[0,1]} p(F_1(X_1(\theta);\eta) - F_1(Y_1;\eta)) + (1-p)(F_1(X_2(\theta);\eta) - F_1(Y_2;\eta)) \\
&= \sup_{\eta\in[0,1]} \begin{cases} \left(\frac{p}{\theta} - 1\right)\eta & 0 \le \eta < \theta \\ \frac{\theta-p}{1-\theta}\eta - \frac{\theta-p}{1-\theta} & \theta \le \eta \le 1 \end{cases} \\
&\le 0.
\end{aligned}
$$

One can verify that the above inequality holds if and only if $p \le \theta < 1$. Additionally, it is easy to see that $X(\theta) \succeq^{\mathrm{w}}_{(1)} Y$ when $\theta = 1$, since in that case $X_1(\theta)$ has the same distribution as Y_1 and $X_2(\theta) \equiv 1$. Therefore, $X(\theta) \succeq^{\mathrm{w}}_{(1)} Y$ holds if and only if $p \le \theta \le 1$, so $\Theta = [p, 1]$.

Next, we look at the stochastically weighted dominance relation with chance α, denoted $X(\theta) \succeq^{\mathrm{c}(\alpha)}_{(1)} Y$. Then, (SWD-Chance) becomes

$$
p\mathbf{1}\{F_1(X_1(\theta);\eta) \le F_1(Y_1;\eta)\ \forall \eta \in [0,1]\} + (1-p)\mathbf{1}\{F_1(X_2(\theta);\eta) \le F_1(Y_2;\eta)\ \forall \eta \in [0,1]\} \ \ge\ 1-\alpha.
$$

It is easy to see that the first term on left hand side of the above inequality is equal to zero unless $\theta = 1$, in which case it is equal to p. The second term is equal to $1-p$ for any $\theta \in [0,1]$. The parameter α then plays a key role: if $\alpha \ge p$, then $X(\theta) \succeq^{\mathrm{c}(\alpha)}_{(1)} Y$ holds for all $\theta \in [0,1]$; otherwise, $X(\theta) \succeq^{\mathrm{c}(\alpha)}_{(1)} Y$ only for $\theta = 1$. So, we see that, if α is "too small" (i.e., below the threshold p), then we have $\Theta = \{1\}$, so the stochastically weighted dominance relation with chance α does not provide any relaxation to robust weighted dominance. On the other hand, when α is above the threshold p we have $\Theta = [0,1]$, i.e., the dominance constraint does not impose any restriction on θ.

Finally, we discuss the case of the relaxed strong stochastically weighted dominance relation for a given index $\gamma \ge 0$, denoted $X(\theta) \succeq^{\mathrm{s}(\gamma)}_{(1)} Y$. Then, (3.4) becomes

$$
p\left[\max_{\eta\in[0,1]} F_1(X_1(\theta);\eta) - F_1(Y_1;\eta)\right] + (1-p)\left[\max_{\eta\in[0,1]} F_1(X_2(\theta);\eta) - F_1(Y_2;\eta)\right] \ \le\ \gamma.
$$

Note that the maximum of $F_1(X_1(\theta);\eta) - F_1(Y_1;\eta)$ is equal to $1-\theta$ (achieved at $\eta = \theta$), while the maximum of $F_1(X_2(\theta);\eta) - F_1(Y_2;\eta)$ is 0 (at $\eta = 0$ and 1). Hence, the above inequality is equivalent to $p(1-\theta) \le \gamma$, so we conclude that $X(\theta) \succeq^{\mathrm{s}(\gamma)}_{(1)} Y$ if and only $\theta \ge 1-\gamma/p$, i.e., $\Theta = [(1-\gamma/p)_+, 1]$. We see here how the parameter γ directly controls the degree of relaxation — for $\gamma = 0$ we obtain the same set $\Theta = \{1\}$ as in the robust case; for $0 < \gamma < p$ we obtain an enlarged set $\Theta = [1-\gamma/p, 1]$; for $p \le \gamma \le 1$ the dominance constraint does not imposes any restriction on θ. Note that the particular choice $\gamma = p(1-p)$ yields the same set as in the weak stochastically weighted dominance case. Thus, in this particular case the relaxed strong stochastically weighted dominance relation subsumes the other dominance notions; however, as we shall see in the numerical example in Section 6.1 this is not a general property.

5 Formulations for Stochastically Weighted Dominance

The formulation of univariate nth order stochastic dominance, either in terms of integrals of cumulative distribution functions or in terms of expected utilities (as in (2.2) and (2.3)) does not lend itself to a practical method for testing that relationship. For the case when the random variables being compared (say, X and Y) have finite support, Fishburn and Vickson (1978) propose a linear program that can be used to test for second order dominance. Dentcheva and Ruszczyński (2003) describe an alternative representation based on a set on linear inequalities, which is particularly useful in the context of optimization problems with stochastic dominance constraints, which was introduced in that paper. Luedtke (2008) proposes yet another representation, involving a smaller number of linear inequalities. Inequalities for univariate first order dominance are studied in Luedtke (2008), Noyan et al. (2006) and Noyan and Ruszczyński (2008).

The situation becomes much more complicated in the case of multivariate stochastic dominance. Armbruster and Luedtke (2010) propose a representation for "full" multivariate first and second order dominance, i.e., $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all multidimensional functions nondecreasing (concave) functions $u : \mathbb{R}^n \mapsto \mathbb{R}$. The result is a linear program for the case of second order dominance, and an integer program for the case of first order. As mentioned in Section 1, however, full multivariate dominance can be a very conservative notion. Homem-de-Mello and Mehrotra (2009) study the case of robust multivariate linear second order dominance defined in (RSD) (with $n = 2$) and show that there exists a representation with finitely many linear inequalities. Since the number of inequalities may be large, they propose a cut generation method to selectively obtain these inequalities when solving optimization problems. A concave minimization problem is solved to generate the cuts.

In what follows, we discuss formulations for the three concepts of stochastically weighted dominance described in Section 3 for first and second orders. We present easy-to-check inequalities that can be used to test for dominance. These inequalities are convenient when the goal is simply to compare two random vectors.

We also consider the case situation where the dominance relationship appears as a constraint in an optimization problem. This is an attractive approach for managing risks in an optimization setting. While optimizing an objective function, one avoids high risks by choosing options that are preferable to a random benchmark. In general, we write a stochastic dominance constrained optimization model as

$$\begin{aligned} &\min_{z \in \mathcal{Z}} f(z) \\ &\text{s.t. } X(z) \trianglerighteq^{\bullet}_{(n)} Y, \end{aligned} \tag{5.1}$$

where z represents a decision vector, $\mathcal{Z} \subset \mathbb{R}^d$ represents a deterministic feasible region, $f : \mathcal{Z} \to \mathbb{R}$ is the objective to be minimized, X maps the decision vector to a random vector in $\mathbb{R}^m$, Y is a given reference random vector, and $\trianglerighteq^{\bullet}_{(n)}$ indicates some nth order stochastic dominance relationship. We show that the three concepts of stochastically weighted dominance described in Section 3 (for first and second orders) can also be written in terms of systems of linear (integer) inequalities, which are more easily incorporated into an optimization algorithm that can be used to solve (5.1).

We assume throughout this section that the distributions of X, Y, and V are finite as described by

$$P(X = x^i) = p_i, \quad i \in \mathcal{I} := \{1, \ldots, I\}, \tag{5.2}$$

$$P(Y = y^j) = q_j, \quad j \in \mathcal{J} := \{1, \ldots, J\}, \tag{5.3}$$

$$P(V = v^k) = \mu^k, \quad k \in \mathcal{K} := \{1, \ldots, K\}. \tag{5.4}$$

In the context of the optimization problem (5.1), $X(\cdot)$ can be viewed as a discrete function-valued random vector taking values $x^i(\cdot)$, $i = 1, \ldots, I$, where each $x^i(\cdot)$ is a mapping from $\mathcal{Z}$ to $\mathbb{R}^m$. Thus, the random vector $X(z)$ takes on values $x^i(z) \in \mathbb{R}^m$, $i = 1, \ldots, I$. In what follows we shall use the notation X to represent both an arbitrary random vector and the random vector $X(z)$ for a given z.

5.1 Formulations for Weak Stochastically Weighted Dominance

Recall from Theorem 3.2 that the weak stochastically weighted dominance (SWD-Weak) $X \underline{\blacktriangleright}^{\mathrm{w}}_{(n)} Y$ can be reformulated as a univariate stochastic dominance relationship. Define again $\xi := V^T X$ and $\zeta := V^T Y$. Denote the distribution of ξ and ζ as

$$P(\xi = \xi^i) = \nu_i, \quad i \in \mathcal{A} := \{1, \ldots, A\}, \tag{5.5}$$

$$P(\zeta = \zeta^j) = \rho_j, \quad j \in \mathcal{B} := \{1, \ldots, B\}. \tag{5.6}$$

Also suppose without loss of generality that $\zeta_1 < \cdots < \zeta_B$. It follows from Proposition 3.2 that $X \underline{\blacktriangleright}^{\mathrm{w}}_{(n)} Y$ if and only if $\xi \trianglerighteq_{(n)} \zeta$.

Proposition 5.1 below shows that inequality (2.2) only needs to be tested for finitely many values of η. A similar result is shown in Dentcheva and Ruszczyński (2003), but we present it here for reference.

Proposition 5.1 *Let ξ and ζ be random variables with distributions given by* (5.5) *and* (5.6)*. Then for $n = 1$ or 2, the finite set $\{\zeta^1, \ldots, \zeta^B\}$ contains an optimal solution of*

$$\max_{\eta \in \mathbb{R}} F_n(\xi; \eta) - F_n(\zeta; \eta).$$

Proof: Since $F_1(\xi; \cdot)$ and $F_1(\zeta; \cdot)$ are step functions and ζ^k are all break points of $F_1(\zeta; \eta)$, it is obvious that the proposition holds for $n = 1$.

Let us now prove the case for $n = 2$. Let $\eta \in \mathbb{R}$ be chosen arbitrarily.
Case 1: If $\eta \le \zeta^1$, $F_2(\zeta; \eta) = 0$. Then we have

$$F_2(\xi; \eta) - F_2(\zeta; \eta) \;=\; F_2(\xi; \eta) \;\le\; F_2(\xi; \zeta^1) \;=\; F_2(\xi; \zeta^1) - F_2(\zeta; \zeta^1).$$

Case 2: If $\zeta^k \le \eta \le \zeta^{k+1}$ for some k, then it follows from the convexity of $F_2(\xi; \cdot)$ and from the linearity of $F_2(\zeta; \cdot)$ on $[\zeta^k, \zeta^{k+1}]$ that $F_2(\xi; \cdot) - F_2(\zeta; \cdot)$ is convex on $[\zeta^k, \zeta^{k+1}]$. Thus, $\max_{\eta \in [\zeta^k, \zeta^{k+1}]} F_2(\xi; \eta) - F_2(\zeta; \eta)$ is achieved at one of the extremes of the interval $[\zeta^k, \zeta^{k+1}]$.
Case 3: If $\eta \ge \zeta^B$, we have

$$F_2(\xi; \eta) - F_2(\zeta; \eta) \;=\; F_2(\xi; \eta) - (\eta - \zeta^B) - F_2(\zeta; \zeta^B) \;\le\; F_2(\xi; \zeta^B) - F_2(\zeta; \zeta^B).$$

□

Recall that we can write $F_2(\xi; \eta) = \mathbb{E}[(\eta - \xi)_+]$ and $F_1(\xi; \eta) = \mathbb{E}[\mathbf{1}\{\xi \le \eta\}]$. By Proposition 5.1, we have that $\xi \trianglerighteq_{(2)} \zeta$ if and only if

$$\sum_{i \in \mathcal{A}} \nu_i (\zeta^k - \xi^i)_+ \;\le\; \sum_{j \in \mathcal{B}} \rho_j (\zeta^k - \zeta^j)_+, \quad \forall k \in \mathcal{B}, \tag{5.7}$$

and $\xi \trianglerighteq_{(1)} \zeta$ if and only if

$$\sum_{i\in\mathcal{A}} \nu_i \mathbf{1}\Big\{\xi^i \le \zeta^k\Big\} \;\le\; \sum_{j\le k} \rho_j, \quad \forall k \in \mathcal{B}. \tag{5.8}$$

Inequalities (5.7) and (5.8) can easily be checked to test for weak stochastically weighted dominance between two vectors. As discussed above, though, these inequalities are not convenient in the context of the optimization problem (5.1) — where each ξ^i is a function of the decision vector z — as they contain nonlinearities (in the case of (5.7)) and discontinuities (in the case of (5.8)). We can, however, reformulate those inequalities so that we obtain a system of linear (integer) inequalities as a function of the ξ^i's. The complexity of the resulting problem, of course, depends on the particular dependence of ξ on z — which is determined by the dependence of X on z. For example, if each x^i is a linear function of z then problem (5.1) becomes a standard optimization problem with linear (integer) constraints.

We proceed now with the reformulation. By introducing continuous variables t_{ik} to represent the terms $(\zeta^k - \xi^i)_+$, (5.7) can be used to derive a linear formulation of $\xi \trianglerighteq_{(2)} \zeta$ with $|\mathcal{A}||\mathcal{B}|$ variables and $(|\mathcal{A}|+1)|\mathcal{B}|$ constraints (see Dentcheva and Ruszczyński 2003):

$$\begin{aligned} \sum_{i\in\mathcal{A}} \nu_i t_{ik} \;&\le\; \sum_{j\in\mathcal{B}} \rho_j (\zeta^k - \zeta^j)_+, \quad \forall k \in \mathcal{B}, \\ \xi^i(z) + t_{ik} \;&\ge\; \zeta^k, \quad \forall i \in \mathcal{A}, k \in \mathcal{B}, \\ t_{ik} \;&\ge\; 0, \quad \forall i \in \mathcal{A}, k \in \mathcal{B}. \end{aligned} \tag{5.9}$$

Similarly, using binary variables t_{ik} for the terms $\mathbf{1}\{\xi^i \le \zeta^k\}$, we obtain the mixed integer linear formulation from (5.8)

$$\begin{aligned} \sum_{i\in\mathcal{A}} \nu_i t_{ik} \;&\le\; \sum_{j<k} \rho_j, \quad \forall k \in \mathcal{B}, \\ \xi^i(z) + M_{ik} t_{ik} \;&\ge\; \zeta^k, \quad \forall i \in \mathcal{A}, k \in \mathcal{B}, \\ t_{ik} \;&\in\; \{0,1\}, \quad \forall i \in \mathcal{A}, k \in \mathcal{B}. \end{aligned} \tag{5.10}$$

Here, M_{ik} is sufficiently large to guarantee that the corresponding constraints are redundant when $t_{ik} = 1$. Note that the constant M_{ik} must ensure feasibility for all $z \in \mathcal{Z}$. We assume that for each i there exists a constant l_i such that $v^T x^i(z) \ge l_i$ for all $z \in \mathcal{Z}$ and all $v \in \Delta$ (such an assumption is satisfied, for example, if $x^i(\cdot)$ is continuous and $\mathcal{Z}$ is compact). Then, we can take $M_{ik} = \zeta^k - l_i$ for all i and k. Although this formulation was presented in Noyan et al. (2006) and Noyan and Ruszczyński (2008), the authors do not recommend using this formulation for computation, since the linear programming relaxation bounds are too weak.

We introduce next more efficient representations of the weak stochastically weighted first or second order dominance. It follows from Theorem B.1 — due to Armbruster and Luedtke (2010) and stated in Appendix B for completeness — that $\xi \trianglerighteq_{(2)} \zeta$ is equivalent to

$$
\begin{aligned}
\sum_{j\in\mathcal{B}} \zeta^j t_{ij} &\leq \nu_i \xi^i(z), \quad \forall i \in \mathcal{A}, \\
\sum_{j\in\mathcal{B}} t_{ij} &= \nu_i, \quad \forall i \in \mathcal{A}, \\
\sum_{i\in\mathcal{A}} t_{ij} &= \rho_j, \quad \forall j \in \mathcal{B}, \\
t_{ij} &\geq 0, \quad \forall i \in \mathcal{A}, j \in \mathcal{B}.
\end{aligned}
\tag{5.11}
$$

The system (5.11) consists of $|\mathcal{A}||\mathcal{B}|$ variables and $2|\mathcal{A}| + |\mathcal{B}|$ inequalities. Similarly, we rewrite $X \underline{\blacktriangleright}^{\mathrm{w}}_{(1)} Y$ as a mixed integer linear system of inequalities by Theorem B.3 as follows:

$$
\begin{aligned}
\sum_{j\in\mathcal{B}} t_{ij} \zeta^j &\leq \xi^i(z), \quad \forall i \in \mathcal{A}, \\
\sum_{j\in\mathcal{B}} t_{ij} &= 1, \quad \forall i \in \mathcal{A}, \\
\sum_{i\in\mathcal{A}} \nu_i \sum_{j=1}^{k-1} t_{ij} &\leq \sum_{j=1}^{k-1} \rho_j, \quad \forall k \in \mathcal{B}/\{1\}. \\
t_{ij} &\in \{0,1\}, \quad \forall i \in \mathcal{A}, j \in \mathcal{B}.
\end{aligned}
\tag{5.12}
$$

5.2 Formulations for Stochastically Weighted Dominance with Chance

We turn now to the stochastically weighted dominance relationship with chance α (SWD-Chance), denoted $X \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(n)} Y$ ($n = 1$ or 2). To derive an inequality that represents that dominance concept, we use again the identities $F_2(\xi;\eta) = \mathbb{E}[(\eta - \xi)_+]$ and $F_1(\xi;\eta) = \mathbb{E}[\mathbf{1}\{\xi \leq \eta\}]$, together with Proposition 5.1, to write $X \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(2)} Y$ as

$$
\sum_{k\in\mathcal{K}} \mu_k \mathbf{1}\left\{ \sup_{l\in\mathcal{J}} \sum_{i\in\mathcal{I}} p_i ({v^k}^T y^l - {v^k}^T x^i)_+ - \sum_{j\in\mathcal{J}} q_j ({v^k}^T y^l - {v^k}^T y^j)_+ \leq 0 \right\} \geq 1 - \alpha, \tag{5.13}
$$

which can be alternatively expressed as

$$
\sum_{k\in\mathcal{K}} \mu_k \sup_{l\in\mathcal{J}} \mathbf{1}\left\{ \sum_{i\in\mathcal{I}} p_i ({v^k}^T y^l - {v^k}^T x^i)_+ > \sum_{j\in\mathcal{J}} q_j ({v^k}^T y^l - {v^k}^T y^j)_+ \right\} \leq \alpha. \tag{5.14}
$$

Similarly, we write $X \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(1)} Y$ as

$$
\sum_{k\in\mathcal{K}} \mu_k \mathbf{1}\left\{ \sup_{l\in\mathcal{J}} \sum_{i\in\mathcal{I}} p_i \mathbf{1}\left\{ {v^k}^T x^i \leq {v^k}^T y^l \right\} - \sum_{j\in\mathcal{J}} q_j \mathbf{1}\left\{ {v^k}^T y^j \leq {v^k}^T y^l \right\} \leq 0 \right\} \geq 1 - \alpha. \tag{5.15}
$$

which can be alternatively expressed as

$$\sum_{k\in\mathcal{K}} \mu_k \sup_{l\in\mathcal{J}} \mathbf{1}\left\{ \sum_{i\in\mathcal{I}} p_i \mathbf{1}\left\{ {v^k}^T x^i \le {v^k}^T y^l \right\} > \sum_{j\in\mathcal{J}} q_j \mathbf{1}\left\{ {v^k}^T y^j \le {v^k}^T y^l \right\} \right\} \;\le\; \alpha. \tag{5.16}$$

As before, the above inequalities can easily be checked to test for stochastically weighted dominance with chance α. To derive inequalities that are more suitable for the optimization problem (5.1), we use Theorem B.1. By introducing binary variables s_k so that s_k is equal to 0 if $v^k X \trianglerighteq_{(2)} v^k Y$ and equal to 1 otherwise, it follows that (5.14) can be formulated as

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j\right) t_{ijk} \;\le\; p_i {v^k}^T x^i(z) + p_i M_{ik} s_k, \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.17}$$

$$\sum_{j\in\mathcal{J}} t_{ijk} \;=\; p_i, \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.18}$$

$$\sum_{i\in\mathcal{I}} t_{ijk} \;=\; q_j, \quad \forall j \in \mathcal{J}, k \in \mathcal{K}, \tag{5.19}$$

$$\sum_{k\in\mathcal{K}} \mu_k s_k \;\le\; \alpha, \tag{5.20}$$

$$t_{ijk} \;\ge\; 0, \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, \tag{5.21}$$

$$s_k \;\in\; \{0,1\}, \quad \forall k \in \mathcal{K}. \tag{5.22}$$

Again, M_{ik} is a large enough number to ensure that the system (5.17)-(5.22) is feasible when $s_k = 1$. Under the assumption discussed earlier that $x^i(\cdot)$ is bounded from below, we can choose $M_{ik} = \max_{j\in\mathcal{J}} {v^k}^T y^j - \phi_{ik}$, where $\phi_{ik} \le {v^k}^T x^i(z)$ for all $z \in \mathcal{Z}$.

To obtain an alternative formulation to (5.16), we can use Theorem B.3 — due to Luedtke (2008) and stated in Appendix B for completeness — to obtain

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j\right) t_{ijk} \;\le\; {v^k}^T x^i(z) + M_{ik} s_k, \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.23}$$

$$\sum_{j\in\mathcal{J}} t_{ijk} \;=\; 1, \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.24}$$

$$\sum_{i\in\mathcal{I}} p_i \sum_{j=1}^{l-1} t_{ijk} \;\le\; \sum_{j=1}^{l-1} q_j, \quad l \in \mathcal{J}/\{1\}, k \in \mathcal{K}, \tag{5.25}$$

$$\sum_{k\in\mathcal{K}} \mu_k s_k \;\le\; \alpha, \tag{5.26}$$

$$t_{ijk} \;\in\; \{0,1\}, \quad \forall i \in \mathcal{I}, j \in \mathcal{J}, k \in \mathcal{K}, \tag{5.27}$$

$$s_k \;\in\; \{0,1\}, \quad \forall k \in \mathcal{K}. \tag{5.28}$$

A disadvantage of formulations (5.17)-(5.22) and (5.23)-(5.28) is that they involve "big M"-type constraints in (5.17) and (5.23), which are likely to lead to weak linear programming relaxations. We now provide alternative formulations avoiding the "big M" in the following theorems.

Theorem 5.2 *For each $i \in \mathcal{I}$ and $k \in \mathcal{K}$, suppose we can choose ϕ_{ik} such that $\phi_{ik} \le {v^k}^T x^i(z)$ for* *all $z \in \mathcal{Z}$. Then, $X(z) \succeq^{c(\alpha)}_{(2)} Y$ if and only if there exist variables t_{ijk} and s_k for $i \in \mathcal{I}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$ such that* (5.20)-(5.22) *hold and*

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j - \phi_{ik}\right) t_{ijk} \;\leq\; p_i \left({v^k}^T x^i(z) - \phi_{ik}\right), \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.29}$$

$$\sum_{j\in\mathcal{J}} t_{ijk} \;=\; p_i(1 - s_k), \quad \forall i \in \mathcal{I}, k \in \mathcal{K}, \tag{5.30}$$

$$\sum_{i\in\mathcal{I}} t_{ijk} \;=\; q_j(1 - s_k), \quad \forall j \in \mathcal{J}, k \in \mathcal{K}. \tag{5.31}$$

Proof: We claim that, for each $k \in \mathcal{K}$, inequalities (5.17)- (5.19) hold with $s_k = 0$ if and only if (5.29)-(5.31) hold with $s_k = 0$, and similarly for $s_k = 1$. To see that, fix $k \in \mathcal{K}$. From Theorem B.1, we have that ${v^k}^T X(z)$ dominates ${v^k}^T Y$ in the second order if and only if there exist $t_{ijk} \geq 0$ for $i \in \mathcal{I}$ and $j \in \mathcal{J}$ such that (5.18)-(5.19) hold and

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j\right) t_{ijk} \;\leq\; p_i {v^k}^T x^i(z), \quad \forall i \in \mathcal{I}, \tag{5.32}$$

which, by (5.18), is equivalent to

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j - \phi_{ik}\right) t_{ijk} \;\leq\; p_i \left({v^k}^T x^i(z) - \phi_{ik}\right), \quad \forall i \in \mathcal{I}.$$

Suppose $s_k = 0$. Then, since (5.30)-(5.31) are identical to (5.18)-(5.19), and (5.32) coincides with (5.17), the claim follows. Suppose now that (5.17)-(5.19) can only hold with $s_k = 1$, which means that (5.32) cannot be satisfied. Then, by taking $t_{ijk} = 0$ for all i and j, we see that (5.29)-(5.31) hold with $s_k = 1$. Conversely, suppose that (5.29)-(5.31) can only hold with $s_k = 1$. Then, (5.17)-(5.19) can only hold with $s_k = 1$ as well, for if those inequalities held with $s_k = 0$ then as seen above (5.29)-(5.31) would hold with $s_k = 0$ as well.

In summary, s_k is equal to 1 if and only if ${v^k}^T X(z)$ does not dominate ${v^k}^T Y$ in the second order. Constraint (5.20) then completes the formulation. □

Theorem 5.3 *For each $i \in \mathcal{I}$ and $k \in \mathcal{K}$, suppose we can choose ϕ_{ik} such that $\phi_{ik} \leq {v^k}^T x^i(z)$ for all $z \in \mathcal{Z}$. Then, $X(z) \succeq^{c(\alpha)}_{(1)} Y$ if and only if there exist binary variables t_{ijk} and s_k for $i \in \mathcal{I}$, $j \in \mathcal{J}$, and $k \in \mathcal{K}$ such that* (5.26)-(5.28) *hold and*

$$\sum_{j\in\mathcal{J}} \left({v^k}^T y^j - \phi_{ik}\right) t_{ijk} \;\leq\; \left({v^k}^T x^i - \phi_{ik}\right), \quad \forall i \in \mathcal{I}, k \in \mathcal{K} \tag{5.33}$$

$$\sum_{j\in\mathcal{J}} t_{ijk} \;=\; 1 - s_k, \quad \forall i \in \mathcal{I}, k \in \mathcal{K} \tag{5.34}$$

$$\sum_{i\in\mathcal{I}} p_i \sum_{j=1}^{l-1} t_{ijk} \;\leq\; \sum_{j=1}^{l-1} q_j, \quad \forall l \in \mathcal{J}/\{1\}, k \in \mathcal{K}. \tag{5.35}$$

Proof: A proof can be constructed in the same way as in Theorem 5.2. □

Remark 5.4 *If $X(z) \geq 0$ a.s. for all $z \in \mathcal{Z}$, we can let $\phi_{ik} = 0$ in Theorems 5.2 and 5.3.*

5.3 Formulations for Relaxed Strong Stochastically Weighted Dominance

Finally, we discuss the case of relaxed strong stochastically weighted dominance (SWD-Relax), denoted $X \underline{\blacktriangleright}_{(k)}^{\mathrm{s}(\gamma)} Y$. By using again the identities $F_2(\xi;\eta) = \mathbb{E}[(\eta-\xi)_+]$ and $F_1(\xi;\eta) = \mathbb{E}[\mathbf{1}\{\xi \le \eta\}]$, together with Proposition 5.1, we can derive an inequality for that dominance relationship. We write $X \underline{\blacktriangleright}_{(2)}^{\mathrm{s}(\gamma)} Y$ as

$$\sum_{k\in\mathcal{K}} \mu_k \left(\max_{l\in\mathcal{J}} \sum_{i\in\mathcal{I}} p_i \left({v^k}^T y^l - {v^k}^T x^i \right)_+ - \sum_{j\in\mathcal{J}} q_j \left({v^k}^T y^l - {v^k}^T y^j \right)_+ \right) \le \gamma, \tag{5.36}$$

and $X \underline{\blacktriangleright}_{(1)}^{\mathrm{s}(\gamma)} Y$ as

$$\sum_{k\in\mathcal{K}} \mu_k \left(\max_{l\in\mathcal{J}} \sum_{i\in\mathcal{I}} p_i \mathbf{1}\left\{ {v^k}^T x^i < {v^k}^T y^l \right\} - \sum_{j\in\mathcal{J}} q_j \mathbf{1}\left\{ {v^k}^T y^j < {v^k}^T y^l \right\} \right) \le \gamma. \tag{5.37}$$

To derive inequalities that are more suitable for the optimization problem (5.1), we introduce intermediate variables s_k and t_{ilk}, so that (5.36) can be represented as

$$\begin{aligned}
\sum_{k\in\mathcal{K}} \mu_k s_k &\le \gamma, \\
s_k &\ge \sum_{i\in\mathcal{I}} p_i t_{ilk} - \sum_{j\in\mathcal{J}} q_j \left({v^k}^T y^l - {v^k}^T y^j \right)_+, \quad l\in\mathcal{J}, k\in\mathcal{K}, \\
{v^k}^T x^i(z) + t_{ilk} &\ge {v^k}^T y^l, \quad i\in\mathcal{I}, l\in\mathcal{J}, k\in\mathcal{K}, \\
t_{ilk} &\ge 0, \quad i\in\mathcal{I}, l\in\mathcal{J}, k\in\mathcal{K}.
\end{aligned} \tag{5.38}$$

Analogously, using continuous variables s_k and binary variables t_{ilk}, we reformulate (5.37) as

$$\begin{aligned}
\sum_{k\in\mathcal{K}} \mu_k s_k &\le \gamma, \\
s_k &\ge \sum_{i\in\mathcal{I}} p_i t_{ilk} - \sum_{j\in\mathcal{J}} q_j \mathbf{1}\left\{ {v^k}^T y^l - {v^k}^T y^j \right\}, \quad l\in\mathcal{J}, k\in\mathcal{K}, \\
{v^k}^T x^i(z) + M_{ilk} t_{ilk} &\ge {v^k}^T y^l, \quad i\in\mathcal{I}, l\in\mathcal{J}, k\in\mathcal{K}, \\
t_{ilk} &\in \{0,1\}, \quad i\in\mathcal{I}, l\in\mathcal{J}, k\in\mathcal{K}.
\end{aligned} \tag{5.39}$$

As before, the constants M_{ilk} must be large enough numbers to ensure feasibility of the system; for example, we can define $M_{ilk} := {v^k}^T y^l - \phi_{ik}$, where $\phi_{ik} \le {v^k}^T x^i(z)$ for all $z\in\mathcal{Z}$. As mentioned earlier, large constants likely yield weak linear programming relaxations of (5.39). Below, we propose alternative formulations of $X \underline{\blacktriangleright}_{(2)}^{\mathrm{s}(\gamma)} Y$ and $X \underline{\blacktriangleright}_{(1)}^{\mathrm{s}(\gamma)} Y$.

Theorem 5.5 *For each $i\in\mathcal{I}$ and $k\in\mathcal{K}$, suppose we can choose ϕ_{ik} such that $\phi_{ik} \le {v^k}^T x^i(z)$ for all $z\in\mathcal{Z}$. Let $\pi_{k1} \le \cdots \le \pi_{kJ}$ be a sorting permutation of ${v^k}^T y^1, \ldots, {v^k}^T y^J$ and $\pi_{k0} := \min\{\phi_{ik},\ \pi_{k1}\}$.*

Then, $X(z) \underline{\blacktriangleright}_{(2)}^{s(\gamma)} Y$ *if and only if*

$$\begin{aligned}
\sum_{k \in \mathcal{K}} \mu_k s_k \ &\le \ \gamma, && (5.40)\\
\sum_{j=0}^{J} \pi_{kj} t_{ijk} \ &\le \ {v^k}^T x^i(z), \quad i \in \mathcal{I}, k \in \mathcal{K}, && (5.41)\\
\sum_{j=0}^{J} t_{ijk} \ &= \ 1, \quad i \in \mathcal{I}, k \in \mathcal{K}, && (5.42)\\
\sum_{i=1}^{I} p_i \sum_{j=0}^{l-1} (\pi_{kl} - \pi_{kj}) t_{ijk} - s_k \ &\le \ \sum_{j=1}^{l-1} q_j (\pi_{kl} - \pi_{kj}), \quad l \in \mathcal{J}, k \in \mathcal{K}, && (5.43)\\
t_{ijk} \ &\ge \ 0, \quad i \in \mathcal{I}, j \in \mathcal{J} \cup \{0\}, k \in \mathcal{K}. && (5.44)
\end{aligned}$$

Proof: For a fixed $k \in \mathcal{K}$, define $L_0^k := 0$ and $L_l^k := \sum_{j=1}^{l-1} q_j(\pi_{kl} - \pi_{kj}) + s_k$ for all $l \in \mathcal{J}$. Here, s_k is nonnegative. By Theorem B.2 — also due to Luedtke (2008) and stated in Appendix B for completeness — we have that (5.41)-(5.44) hold if and only if

$$\sum_{i \in \mathcal{I}} p_i \left(\pi_{kl} - {v^k}^T x^i(z) \right)_+ \ \le \ L_l^k, \quad l \in \mathcal{J} \cup \{0\},$$

which is equivalent to

$$s_k \ \ge \ \max_{l \in \mathcal{J}} \sum_{i \in \mathcal{I}} p_i \left(\pi_{kl} - {v^k}^T x^i(z) \right)_+ - \sum_{j \in \mathcal{J}} q_j \left(\pi_{kl} - \pi_{kj} \right)_+ . \tag{5.45}$$

It is easy to see that, together, inequalities (5.40) and (5.45) (for all $k \in \mathcal{K}$) are equivalent to (5.36).
□

For the case of first order dominance, we can directly adapt the proof of Theorem B.3 to our setting. For completeness, the proof of the following theorem is given in Appendix C.

Theorem 5.6 *For each* $i \in \mathcal{I}$ *and* $k \in \mathcal{K}$*, suppose we can choose* ϕ_{ik} *such that* $\phi_{ik} \le {v^k}^T x^i(z)$ *for all* $z \in \mathcal{Z}$*. Let* $\pi_{k1} \le \cdots \le \pi_{kJ}$ *be a sorting permutation of* ${v^k}^T y^1, \ldots, {v^k}^T y^J$ *and* $\pi_{k0} := \min\{\phi_{ik}, \ \pi_{k1}\}$*. Then,* $X(z) \underline{\blacktriangleright}_{(1)}^{s(\gamma)} Y$ *if and only if* (5.40)-(5.42) *hold and*

$$\begin{aligned}
\sum_{i \in \mathcal{I}} p_i \sum_{j=0}^{l-1} t_{ijk} - s_k \ &\le \ \sum_{j=1}^{l-1} q_j, \quad l \in \mathcal{J}, k \in \mathcal{K}, && (5.46)\\
t_{ijk} \ &\in \ \{0, 1\}, \quad i \in \mathcal{I}, j \in \mathcal{J} \cup \{0\}, k \in \mathcal{K}. && (5.47)
\end{aligned}$$

6 Applications to Optimization and Multi-Criterion Decision Making

In this section we discuss the application of the ideas laid out in the previous sections to optimization problems. We also illustrate the usefulness of those concepts in multi-criterion decision making by describing a queueing problem where the goal is to decide on an appropriate allocation of servers between two parallel systems, based on the average waiting times of each queue.

6.1 Enlarging Feasibility Regions in Optimization

As mentioned earlier, one of the benefits of three proposed concepts of stochastically weighted dominance is that they provide some relaxation to the notion of robust weighted stochastic dominance, a property that was illustrated by the example given in Section 4. When stochastic dominance is used as a constraint in an optimization problem as in (5.1), such relaxation translates into a reshaping of the feasibility region. To illustrate that impact, we describe next a simple two-dimensional problem so we can graphically depict the feasibility regions under the various dominance relationships. Let

$$A(\omega) := \begin{bmatrix} 1+0.25\omega & 0.5 \\ 0.5 & 0.5-0.25\omega \\ 0.25 & 0.03 \end{bmatrix}, \quad l(\omega) := \begin{bmatrix} 0.5-0.0025\omega \\ 0.4 \\ 0.1+0.013\omega \end{bmatrix}, \quad g(\omega) := \begin{bmatrix} 0.05 \\ 0.2-0.025\omega \\ 0.05+0.013\omega \end{bmatrix},$$

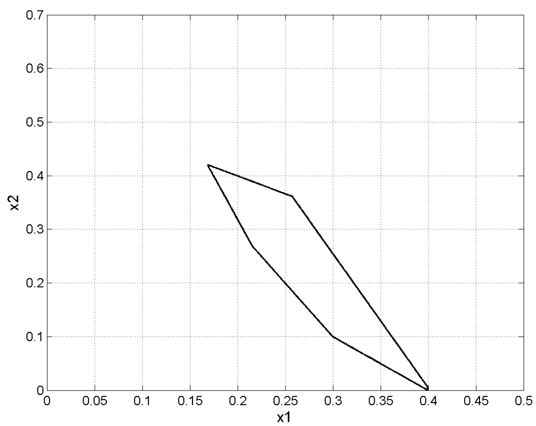
where ω is a random variable with equally likely outcomes ± 1. Consider a feasibility region defined by the following two stochastic dominance constraints:

$$-A(\omega)z \trianglerighteq^{\bullet}_{(2)} -l(\omega), \tag{6.1}$$

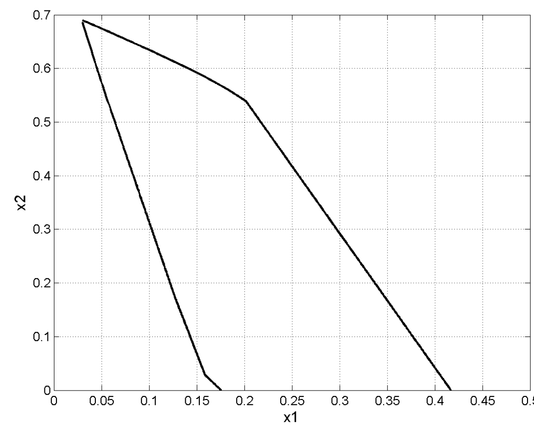
$$A(\omega)z \trianglerighteq^{\bullet}_{(2)} g(\omega), \tag{6.2}$$

where z is a two-dimensional nonnegative decision vector. Recall that $\trianglerighteq^{\bullet}_{(2)}$ means some multivariate linear second order stochastic dominance. The vectors $l(\omega)$ and $g(\omega)$ are two given random benchmarks. For example, $l(\omega)$ can be regarded as losses and $g(\omega)$ as gains. Constraint (6.1) requires that random outcomes of a decision policy be no worse than $l(\omega)$ in a situation of loss, while constraint (6.2) requires the decision policy to be at least as good as $g(\omega)$ when there are gains.

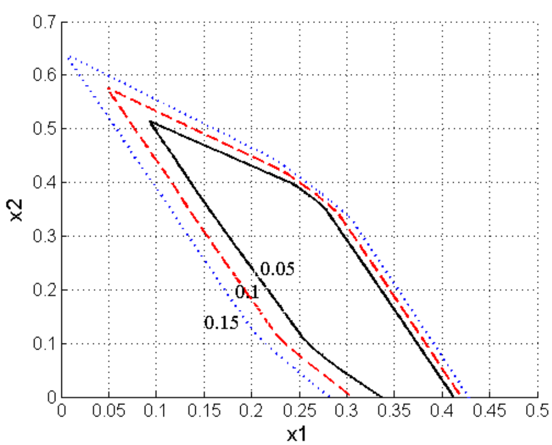
Using the different concepts of stochastic dominance proposed in this paper in constraints (6.1) and (6.2), we obtain feasibility regions shown in Figure 2. Figure 2a describes the case of positive linear stochastic dominance, i.e., $-v^T A(\omega)z \trianglerighteq_{(2)} -v^T l(\omega)$ and $v^T A(\omega)z \trianglerighteq_{(2)} v^T g(\omega)$ for all $v \in \Delta = \{(v_1, v_2) \in \mathbb{R}^2_+ : v_1 + v_2 = 1\}$. Choosing V as a random vector uniformly distributed on Δ, we draw the feasibility regions corresponding to the cases of the three proposed stochastically weighted dominance relationships in Figures 2b-2d. It is clear that all three stochastically weighted dominance relationships yield larger feasibility regions than the one given by positive linear stochastic dominance. In addition, Figure 2c shows expanding regions for the parameter α in the stochastically weighted dominance relationship with chance α, when α increases from 0.05 to 0.1 to 0.15. Figure 2d shows the enlarged regions for the case of relaxed strong stochastically weighted dominance relationship with parameter γ, when γ increases from 0.001 to 0.002, to 0.003. Note that even reasonably small values of α and γ yield significant enlargement of the original feasibility region. Thus, these alternative notions of dominance can potentially lead to solutions with better objective value while still controlling for risk.



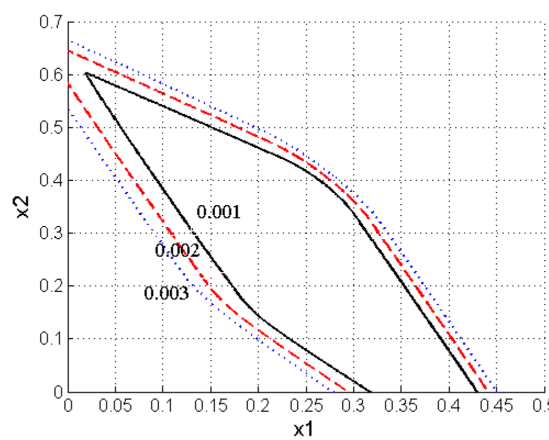
(a) Positive Linear Dominance



(b) Weak Stochastically Weighted Dominance



(c) Stochastically Weighted Dominance with Chance



(d) Relaxed Strong Stochastically Weighted Dominance

Figure 2: Feasible Regions

6.2 A Queueing Staff Allocation Problem

We study now an application of the stochastically weighted dominance in multi-criterion decision making (see Keeney and Raffa (1993) and Dyer and Sarin (1979)). An example of two $M/M/c$ queues in parallel is used to illustrate the effectiveness of using the concepts of stochastically weighted dominance.

Consider a system in which services are provided separately to (a) regular and (b) VIP customers, where the quality of each type of service is measured by the steady-state waiting time of the corresponding queue. Suppose there is a certain number of employees that must be allocated between the two queues. We assume that, due to the need for special training, the workers are dedicated to the queue they are allocated to, i.e., they cannot move to the other queue even if their queue is empty. Also, it is reasonable to assume that the waiting times of regular and VIP customers (denoted W_a and W_b, respectively) have different levels of importance.

How to allocate the workers between the two queues? A standard approach is to define weights v_a and v_b representing the importance of each type of customer, define a value function $\Psi_v(W_a, W_b) := v_a W_a + v_b W_b$, and then find the allocation that minimizes $\mathbb{E}[\Psi_v(W_a, W_b)]$. There are however two issues with such an approach:

(i) How to decide on the values of v_a and v_b? Clearly, the allocation that minimizes $\mathbb{E}[\Psi_v(W_a, W_b)]$ depends on the particular choice of v. Such dependence can be very sensitive, i.e., a small change in v may lead to a different solution;

(ii) The minimization of the expected waiting times does not take into account the risk associated with the uncertainty of waiting times. For example, a solution in which the values of $\Psi_v(W_a, W_b)$ vary wildly may not be desirable, due to the customer dissatisfaction it may generate.

These issues can be addressed by the models proposed in this paper. Risk is taken into account by comparing alternatives in terms of stochastic dominance of $\Psi_v(W_a, W_b)$. For example, second order dominance, which corresponds to risk aversion, can be used. Ideally, we would like to rank the feasible allocations in terms of dominance, though this may not always be possible — indeed, it is possible that no alternative dominates the others. The uncertainty in the weights can be modeled by defining a random vector of weights $V = (V_a, V_b)$. As mentioned earlier, using a random vector of weights allows us to represent not only the existence of multiple decision makers — each corresponding to one value of V — but also the relative importance of each decision maker, which corresponds to the probability of each value of V. The corresponding value function $\Psi_V(W_a, W_b) := V_a W_a + V_b W_b$ is then used for the stochastic dominance comparisons.

We study now the application of the concepts of stochastically weighted dominance to a particular instance of the above problem. Suppose that the total number of workers to be allocated is five. The arrival and service rates for regular customers are $\lambda_a = 15$ and $\mu_a = 8$, while the rates for VIP customers are $\lambda_b = 3.5$ and $\mu_b = 4$. Since the queues must be stable, it follows that there are three alternatives: (4, 1), (3, 2), and (2, 3). The numbers in each pair indicate the number of workers allocated to regular and VIP customers, respectively. The distribution of the random vector of weights V is given in Table 1. Note that the distribution gives more importance to the weight vectors emphasizing the VIP queue as expected. However, it also allows, with low probability, larger weights for the regular queue.

Table 1: Distribution of Random Weight Vector V

Probability	0.05	0.1	0.25	0.35	0.15	0.05	0.03	0.01	0.01
Weight	(0.1, 0.9)	(0.2, 0.8)	(0.3, 0.7)	(0.4, 0.6)	(0.5, 0.5)	(0.6, 0.4)	(0.7, 0.3)	(0.8, 0.2)	(0.9, 0.1)

It is well known (see, e.g., Tijms 2003) that the cumulative distribution function of the steady-state waiting time of an $M/M/c$ queue is given by

$$F_1(W;t) = 1 - P_{\text{wait}} e^{-c\mu(1-\rho)t}, \quad t \geq 0,$$

where c is the number of servers, $\rho := \frac{\lambda}{c\mu}$, and

$$P_{\text{wait}} := \frac{(c\rho)^c}{c!}\left((1-\rho)\sum_{n=0}^{c-1}\frac{(c\rho)^n}{n!} + \frac{(c\rho)^c}{c!}\right)^{-1}$$

is the steady-state probability that an arriving customer must wait in queue because all servers are busy.

In order to use the inequalities presented in Section 5 to compare the waiting times in terms of stochastic dominance, we need to discretize the distributions of W_a and W_b. We can do that by approximating the c.d.f. by a step function:

$$\widehat{F}_1(W_a;t) = \begin{cases} F_1(W_a; iT/N), & iT/N \leq t < (i+1)T/N, \ i = 0, \dots, N-1, \\ 1, & \text{otherwise}, \end{cases}$$

and similarly for W_b. In the numerical experiments reported below, we used $N = 100$ and $T = 30$. Also, we used the second order form for all notions of dominance. Below, we use the notation $W^{(k,l)}$ to denote the random vector $(-W_a,\ -W_b)$ for the alternative with k workers allocated to the regular queue and l workers to the VIP queue (note the minus sign since the notions of dominance defined earlier assume that larger values are preferred).

Results

When using the weak stochastically weighted dominance as the judgement rule, configuration (3, 2) is the best choice since $W^{(3,2)} \underline{\blacktriangleright}^{\mathrm{w}}_{(2)} W^{(4,1)}$ and $W^{(3,2)} \underline{\blacktriangleright}^{\mathrm{w}}_{(2)} W^{(2,3)}$. Although using this dominance relationship suggests a preferred alternative, this approach does not give the decision maker clear guidance on how much better (3, 2) is, compared to (4, 1) and (2, 3).

Recall the quantities $\mathfrak{S}(X,Y)$ and $\mathfrak{T}(X,Y)$, which were defined in (SWD-Chance) and (SWD-Relax) as the satisfaction level and the tolerance level of preference of X over Y, respectively. Tables 2 and 4 list the satisfaction level and the tolerance level for the comparisons. Tables 3 and 5 show the same results but from a different perspective — for fixed values of α and γ, the tables show the results of the comparisons between all pairs. In those tables, we use the notation "N/C" to indicate that the elements of the pair are not comparable, in the sense that neither vector dominates the other. On the other hand, we use the notation $=^{\mathrm{c}(\alpha)}_{(2)}$ and $=^{\mathrm{s}(\gamma)}_{(2)}$ to indicate that the elements of the pair are indistinguishable from the perspective of the corresponding dominance relationship since they dominate each other. We also use the latter notation when a cycle occurs, i.e., $X \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(2)} Y \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(2)} Z \underline{\blacktriangleright}^{\mathrm{c}(\alpha)}_{(2)} X$ and similarly for the $\underline{\blacktriangleright}^{\mathrm{s}(\gamma)}_{(2)}$ order.

The results indicate that an "optimal" tolerance level α or γ exists in both cases — if the tolerance is too low, some pairs are not compared; if it is too high, the pairs become indistinguishable and so the dominance relationship does not convey any information. In this particular example we are able to obtain a complete order when $25\% \leq \alpha < 95\%$ or $0.0159 \leq \gamma < 0.3010$ since in that case (3,2) is preferable to (2,3) which is preferable to (4,1). In general, however, a complete order may not exist for any value of α or γ, although typically a proper choice of α or γ allows for more pairs to become comparable than under the robust weighted dominance relationship (which corresponds to $\alpha = 0$ and $\gamma = 0$).

Table 2: Satisfaction index

	(4, 1)	(3, 2)	(2, 3)
(4, 1)	100%	0	5%
(3, 2)	100%	100%	95%
(2, 3)	75%	0	100%

Table 3: Comparisons for fixed α

	(3,2) vs. (4, 1)	(3, 2) vs. (2,3)	(4,1) vs. (2, 3)
$0 \leq \alpha < 5\%$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(4,1)}$	N/C	N/C
$5\% \leq \alpha < 25\%$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(2,3)}$	N/C
$25\% \leq \alpha < 95\%$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(2,3)}$	$W^{(2,3)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(4,1)}$
$95\% \leq \alpha < 100\%$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}^{c(\alpha)}_{(2)} W^{(2,3)}$	$W^{(2,3)} =^{c(\alpha)}_{(2)} W^{(4,1)}$
$\alpha = 100\%$	$W^{(3,2)} =^{c(\alpha)}_{(2)} W^{(4,1)}$	$W^{(3,2)} =^{c(\alpha)}_{(2)} W^{(2,3)}$	$W^{(2,3)} =^{c(\alpha)}_{(2)} W^{(4,1)}$

Table 4: Least tolerance level

	(4, 1)	(3, 2)	(2, 3)
(4, 1)	0	1.0652	0.7824
(3, 2)	0	0	0.0011
(2, 3)	0.0159	0.3010	0

Table 5: Comparisons for fixed tolerance level γ

	(3,2) vs. (4, 1)	(3, 2) vs. (2,3)	(4,1) vs. (2, 3)
$0 \leq \gamma < 0.0011$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$	N/C	N/C
$0.0011 \leq \gamma < 0.0159$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(2,3)}$	N/C
$0.0159 \leq \gamma < 0.3010$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(2,3)}$	$W^{(2,3)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$
$0.3010 \leq \gamma < 0.7824$	$W^{(3,2)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$	$W^{(3,2)} =^{s(\gamma)}_{(2)} W^{(2,3)}$	$W^{(2,3)} \underline{\blacktriangleright}^{s(\gamma)}_{(2)} W^{(4,1)}$
$0.7824 \leq \gamma$	$W^{(3,2)} =^{s(\gamma)}_{(2)} W^{(4,1)}$	$W^{(3,2)} =^{s(\gamma)}_{(2)} W^{(2,3)}$	$W^{(2,3)} =^{s(\gamma)}_{(2)} W^{(4,1)}$

The analysis based on the satisfaction level and the tolerance level, in addition to the approach using weak stochastically weighted dominance, provides more comprehensive information to help decision makers to understand the preference between alternatives. Both the satisfaction level and the tolerance level can be used in a case study, since these two indices test different aspects of the dominance relationship.

It is worthwhile comparing the above conclusions with that obtained from using the robust weighted dominance relationship, where the robust weight region is defined by the support $\mathcal{V}$ of the random vector V. It is clear from the previous discussion that robust dominance holds if and only if the satisfaction level is 100% and the tolerance level is zero. Hence, under robust weighted dominance, the only conclusion we can draw is that (3, 2) is preferred to (4, 1).

Thus, the above example suggests that the new stochastic dominance relationships introduced in this paper provide valuable tools to compare multiple criteria under uncertainty. In particular, the notions of satisfaction level and tolerance level may allow for a ranking of the alternatives, which is typically not possible under standard concepts of stochastic dominance.

7 Concluding Remarks

We have introduced three new concepts of stochastically weighted dominance — weak stochastically weighted dominance, stochastically weighted dominance with chance, and relaxed strong stochastically weighted dominance. The equivalent utility presentations of the first or second order stochastic dominance and integer and linear programming formulations have been presented. Examples are used to compare these concepts of stochastically weighted dominance with each other, and with the deterministic weighted dominance definitions. These examples show that the use of uncertainly weighted approaches allow larger feasible regions, when used as constraints in optimization. Another application is presented for a multi-criterion staffing problem. In this problem the goal is to decide the allocation of servers between two $M/M/c$ queues based on waiting times. Using this example we illustrate the value of additional concepts of dominance satisfaction level, and dominance tolerance level when using stochastically weighted dominance concepts for a ranking of the alternatives. In conclusion, we have provided decision makers with a valuable tool for comparing multivariate risk, define more tractable optimization problems using stochastic dominance, and a tool to choose among the possible alternative configurations. An interesting extension of the optimization formulations given here would be to develop them for the case where the random vectors have continuous distributions. A possible approach can be developed based on combining sampling in the current framework. This will allow the stochastic weighted dominance based optimization and decision models to become more widely applicable, particularly for use in simulated systems.

References

B. Armbruster and J. Luedtke. Models and formulations for multivariate dominance constrained stochastic programs. `http://www.optimization-online.org/DB_FILE/2010/05/2613.pdf`, 2010.

P. Billingsley. *Probability and Measure*. John Wiley & Sons, 1995.

D. Dentcheva and A. Ruszczyński. Optimization with stochastic dominance constraints. *SIAM J. Optim.*, 14(2):548–566, 2003.

D. Dentcheva and A. Ruszczyński. Portfolio optimization with stochastic dominance constraints. *Journal of Banking Finance*, 30:433–451, 2006.

D. Dentcheva and A. Ruszczyński. Optimization with multivariate stochastic dominance constraints. *Math. Programming*, 117:111–127, 2009.

D. Drapkin and R. Schultz. An algorithm for stochastic programs with first-order dominance constraints induced by linear recourse. Manuscript, Department of Mathematics, University of Duisburg-Essen, Duisburg, Germany, 2007.

J. Dyer and R. Sarin. Measurable value functions. *Operations Research*, 27(4):810–822, 1979.

P. C. Fishburn and R. G. Vickson. Theoretical foundations of stochastic dominance. In G. Whitmore and M. Findlay, editors, *Stochastic Dominance*. Lexington Books, 1978.

T. Homem-de-Mello and S. Mehrotra. A cutting surface method for linear programs with polyhedral stochastic dominance constraints. *SIAM Journal on Optimization*, 20(3):1250–1273, 2009.

J. Hu and S. Mehrotra. Robust and stochastically weighted multi-objective optimization models and reformulations. Manuscript, Northwestern University, available at `http://www.optimization-online.org/DB_HTML/2010/10/2762.html`, 2010.

J. Hu, T. Homem-de-Mello, and S. Mehrotra. Sample average approximation of stochastic dominance constrained programs. To appear in *Mathematical Programming*, 2010.

J. Hu, T. Homem-de-Mello, and S. Mehrotra. Risk adjusted budget allocation models with application in homeland security. To appear in *IIE Transactions.*, 2011.

R. L. Keeney and H. Raffa. *Decisions with multiple objectives: preferences and value tradeoffs*. Cambridge University Press, 1993.

H. Levy. *Stochastic Dominance : Investment Decision Making under Uncertainty*. Springer, New York, 2006.

J. Luedtke. New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, 19(3):1433–1450, 2008.

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

M. Nowak. Interactive approach in multicriteria analysis based on stochastic dominance. *Control and Cybernetics*, 33(3):463–476, 2004.

N. Noyan and A. Ruszczyński. Valid inequalities and restrictions for stochastic programming problems with first order stochastic dominance constraints. *Mathematical Programming*, 114: 249–275, 2008.

N. Noyan, G. Rudolf, and A. Ruszczyński. Relaxations of linear programming problems with first order stochastic dominance constraints. *Operations Research Letters*, 34:653–659, 2006.

S. T. Rachev, S. V. Stoyanov, and F. J. Fabozzi. *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization : The Ideal Risk, Uncertainty, and Performance Measures*. John Wiley & Sons, 2008.

D. Roman, K. Darby-Dowman, and G. Mitra. Portfolio construction based on stochastic dominance and target return distributions. *Math. Programming*, 108:541–569, 2006.

M. Shaked and J. G. Shanthikumar. *Stochastic Orders and their Applications*. Academic Press, Boston, 1994.

H. C. Tijms. *A First Course in Stochastic Models*. John Wiley & Sons, 2003.

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ, 2nd. edition, 1947.

K. Zaras and J. Martel. *Multiattribute analysis based on stochastic dominance*. Kluwer Academic Publishers, Dordrecht, 1994.

A Relaxed Almost Sure Stochastically Weighted Dominance

In this Appendix we discuss another relaxation of robust weighted dominance. It is interesting to notice that, although (3.2) and (3.3) are equivalent, the relaxations of those inequalities yield different results. Indeed, the relaxation of (3.2), which we call the *relaxed almost sure stochastically weighted dominance*, is

$$F_n(V^TX;\eta\,|\,V) - F_n(V^TY;\eta\,|\,V) \;\le\; \gamma, \quad \forall\eta\in\mathbb{R}, \quad \text{a.s.}, \tag{A.1}$$

which is clearly more conservative than the relaxed strong stochastically weighted dominance since in (SWD-Relax) the left hand side of (A.1) is allowed to be greater than γ with positive probability for some η. The relaxed almost sure stochastically weighted dominance is analogous to the relaxation of robust weighted dominance given in Hu et al. (2011) as

$$F_n(v^TX;\eta) - F_n(v^TY;\eta) \;\le\; \gamma, \quad \forall\eta\in\mathbb{R},\ \forall v\in\mathcal{V}. \tag{A.2}$$

Note that (A.1) and (A.2) are the same when the support of V is finite.

For the sake of comparison with the other concepts of stochastically weighed dominance defined in Section 3, we apply (A.1) to the two examples in Section 6, where we find the enlarged feasible region in optimization and the decision policy in the staff allocation problem related to the parameter γ.

Figure 3 shows the feasible regions when the constraints (6.1) and (6.2) represent the relaxed almost sure stochastically weighted dominance relationship with parameter γ. The comparison of Figures 2d and 3 shows the difference between the relaxed strong stochastically weighted dominance and the relaxed almost sure stochastically weighted dominance. When we use the same values of γ (i.e., 0.001, 0.002 and 0.003) used in Figure 2d, the relaxed almost sure stochastically weighted dominance relationship yields little change to the original feasible region.

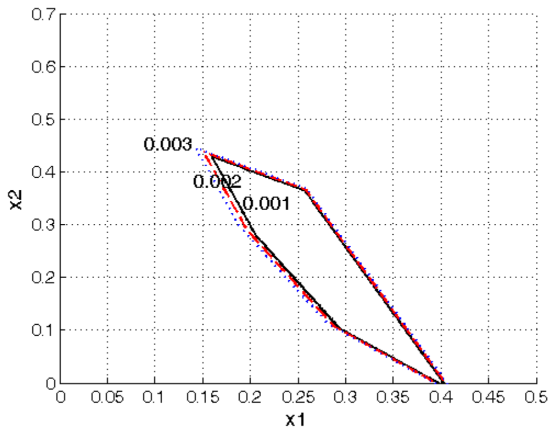


Figure 3: Feasible Region of Relaxed Almost Sure Stochastically Weighted Dominance

We call

$$\mathfrak{R}(X,Y) := \sup_{(\eta,v)\in\mathbb{R}\times\mathcal{V}} F_n(v^TX;\eta) - F_n(v^TY;\eta)$$

the *robust tolerance* level. Table 6 lists the values of $\mathfrak{R}(X,Y)$ for comparisons between all pairs of alternative allocations, whereas Table 7 lists the results of pairwise comparisons for fixed values of that tolerance. We see that, although an "optimal level" of tolerance exists again in this case, the corresponding values ($0.7195 \le \gamma < 0.8243$) are larger than those for tolerance level given in Table 5, i.e., higher tolerance is required if one wants robust weighted dominance to hold.

Table 6: Robust tolerance level

	(4, 1)	(3, 2)	(2, 3)
(4, 1)	0	1.5909	1.5769
(3, 2)	0	0	0.0224
(2, 3)	0.7195	0.8243	0

Table 7: Comparisons for fixed robust tolerance level γ

	(3,2) vs. (4, 1)	(3, 2) vs. (2,3)	(4,1) vs. (2, 3)
$0 \le \gamma < 0.0224$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$	N/C	N/C
$0.0224 \le \gamma < 0.7195$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(2,3)}$	N/C
$0.7195 \le \gamma < 0.8243$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(2,3)}$	$W^{(2,3)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$
$0.8243 \le \gamma < 1.5769$	$W^{(3,2)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$	$W^{(3,2)} =_{(2)}^{s(\gamma)} W^{(2,3)}$	$W^{(2,3)} \underline{\blacktriangleright}_{(2)}^{s(\gamma)} W^{(4,1)}$
$1.5769 \le \gamma$	$W^{(3,2)} =_{(2)}^{s(\gamma)} W^{(4,1)}$	$W^{(3,2)} =_{(2)}^{s(\gamma)} W^{(2,3)}$	$W^{(2,3)} =_{(2)}^{s(\gamma)} W^{(4,1)}$

B Theorems

In this section, we summarize some known theorems which are used in this paper.

Theorem B.1 (Armbruster and Luedtke (2010)) *Let ξ and ζ be random variables with distributions given in* (5.5) *and* (5.6) *respectively.* $\xi \trianglerighteq_{(2)} \zeta$ *if and only if there exists* $t_{ij} \geq 0$ *for* $i \in \mathcal{A}$, $j \in \mathcal{B}$ *such that*

$$\begin{aligned}\sum_{j\in\mathcal{B}} \zeta^j t_{ij} &\leq \nu_i \xi^i, \quad \forall i \in \mathcal{A},\\ \sum_{j\in\mathcal{B}} t_{ij} &= \nu_i, \quad \forall i \in \mathcal{A},\\ \sum_{i\in\mathcal{A}} t_{ij} &= \rho_j, \quad \forall j \in \mathcal{B}.\end{aligned}$$

Theorem B.2 (Luedtke (2008)) *Let ξ be a random variable with distribution given in* (5.5), *let* $\zeta_1 \leq \cdots \leq \zeta_B$ *and let* $0 = L_1 \leq \cdots \leq L_B$. *Denote the set* $\mathcal{B}$ *as* $\{1, \ldots, B\}$. *Then*

$$\mathbb{E}[(\zeta_l - \xi)_+] \leq L_l, \quad \forall l \in \mathcal{B}$$

if and only if there exists $t_{ij} \geq 0$ *for* $i \in \mathcal{A}$, $j \in \mathcal{B}$ *such that*

$$\begin{aligned}\sum_{j\in\mathcal{B}} \zeta_j t_{ij} \leq \xi_i, \quad &\forall i \in \mathcal{A},\\ \sum_{j\in\mathcal{B}} t_{ij} = 1, \quad &\forall i \in \mathcal{A},\\ \sum_{i\in\mathcal{A}} \nu_i \sum_{j=1}^{l-1} (\zeta_l - \zeta_j) t_{ij} \leq L_l, \quad &\forall l \in \mathcal{B}.\end{aligned}$$

Theorem B.3 (Luedtke (2008)) *Let ξ and ζ be random variables with distributions given in* (5.5) *and* (5.6) *respectively.* $\xi \trianglerighteq_{(1)} \zeta$ *if and only if there exists binary variables* t_{ij} *for* $i \in \mathcal{A}$, $j \in \mathcal{B}$ *such that*

$$\begin{aligned}\sum_{j\in\mathcal{B}} \zeta_j t_{ij} \leq \xi_i, \quad &\forall i \in \mathcal{A},\\ \sum_{j\in\mathcal{B}} t_{ij} = 1, \quad &\forall i \in \mathcal{A},\\ \sum_{i\in\mathcal{A}} \nu_i \sum_{j=1}^{l-1} t_{ij} \leq \sum_{j=1}^{l-1} \rho_j, \quad &l = 2, \ldots, B.\end{aligned}$$

C Proof of Theorem 5.6

We now adapt the proof of Theorem B.3 into our case. For the sake of convenience, let $\widehat{\mathcal{J}} := \mathcal{J} \cup \{0\}$ and $v_{ki} := {v^k}^T x^i$. For a fixed $k \in \mathcal{K}$, denote $\Pi_{kl} := \{i \in \mathcal{I} : v_{ki} < \pi_{kl}\}$. By (5.37), $X \blacktriangleright^{s(\gamma)}_{(1)} Y$ can be written as a combination of (5.40) and

$$\mu(\Pi_{kl}) = \sum_{i\in\mathcal{I}} p_i \mathbf{1}\{v_{ki} < \pi_{kl}\} \leq \sum_{j\in\mathcal{J}} q_j \mathbf{1}\{\pi_{kj} < \pi_{kl}\} + s_k = \sum_{j=1}^{l-1} q_j + s_k, \quad \forall l \in \widehat{\mathcal{J}}. \tag{C.1}$$

Note that $s_k \geq 0$ for $l = 0$. Let us now prove that (C.1) holds if and only if (5.41), (5.42), (5.46), and (5.47) are satisfied.

Now suppose that (C.1) is satisfied. For $i \in \mathcal{I}$, denote $l_k^*(i) := \max\{l \in \widehat{\mathcal{J}} : i \notin \Pi_{kl}\}$. $l_k^*(i)$ always exists since $i \notin \Pi_0$ for all i. Note that $l_k^*(i) < l$ if and only if $i \in \Pi_{kl}$. For each $i \in \mathcal{I}$, choose $t_{ijk} = 1$ for $j = l_k^*(i)$ and otherwise $t_{ijk} = 1$. By definition, we have that t satisfies (5.42) and (5.47). For $l \in \mathcal{J}$, $\sum_{j=0}^{l-1} t_{ijk} = 1$ if and only if $l_k^*(i) < l$ and also $\sum_{j=0}^{l-1} t_{ijk} = 1$ if and only if $i \in \Pi_{kl}$. Hence, $\mu(\Pi_{kl}) = \sum_{i\in\mathcal{I}} p_i \sum_{j=0}^{l-1} t_{ijk}$. Thus, by (C.1), we obtain (5.46). As well, for each $i \in \mathcal{I}$, we have $\sum_{j\in\widehat{\mathcal{J}}} \pi_{kj} t_{ijk} = \pi_{kl_k^*(i)} \leq v_{ki}$ since, by definition, $i \notin \Pi_{l_k^*(i)}$, thus establishing that t satisfies (5.41).

Now suppose t satisfies (5.41), (5.42), (5.46), and (5.47). For $l \in \mathcal{J}$, let $\Pi'_{kl} := \{i \in \mathcal{I} : \sum_{j=0}^{l-1} t_{ijk} = 1\}$. Note that if $i \notin \Pi'_{kl}$ then $\sum_{j=0}^{l-1} t_{ijk} = 0$ and by (5.42) $\sum_{j=l}^{J} t_{ijk} = 1$, and so $\sum_{j=0}^{J} \pi_{kj} t_{ijk} \geq \pi_{kl}$. Using (5.41) we have that if $i \in \Pi_{kl}$, then $\sum_{j=0}^{J} \pi_{kj} t_{ijk} \leq v_{ki} < \pi_{kl}$ and hence also $i \in \Pi'_{kl}$. Therefore, by (5.46), we have for $l \in \mathcal{J}$,

$$\mu(\Pi_{kl}) \leq \mu(\Pi'_{kl}) = \sum_{i\in\mathcal{I}} p_i \sum_{j=0}^{l-1} t_{ijk} \leq \sum_{j=1}^{l-1} q_j + s_k.$$