Working Paper

2021/26/DSC

Decisions with Several Objectives under Uncertainty: Sufficient Conditions for Multivariate Almost Stochastic Dominance Based on Means and Variances

Alfred Müller
Universität Siegen, mueller@mathematik.uni-siegen.de

Marco Scarsini
Luiss University, mscarsini@luiss.it

Ilia Tsetlin
INSEAD, ilia.tsetlin@insead.edu

Robert L. Winkler
Duke University, rwinkler@duke.edu

May 23, 2021

Important business, public policy, and personal decisions typically involve multiple objectives, which in turn can be represented by multiple attributes, and uncertainty. Assessing both multiattribute utility and multivariate distributions for the attributes can be challenging. Moreover, big decisions are often made by boards or committees with members holding divergent views and preferences and facing pressures from different stakeholders. Thus, a full-blown traditional decision analysis that leads to the computation of expected utility is very difficult at best and often not possible. We develop sufficient conditions for multivariate almost stochastic dominance (MASD) based on marginal distributions of the attributes or just on their means and variances. To apply MASD, one only needs to assess bounds on marginal utilities. Alternatively, preferences can be explained and elicited via transfers. Realistic examples illustrate our results, which provide tools for "fast and frugal" screening and evaluation of the available options, while properly accounting for tradeoffs and riskiness. Such tools, consistent with normative decision analysis, are useful when making important decisions in today's fast-moving and often complex world.

Subject classifications: Decision analysis: stochastic dominance, utility, risk. Probability: distribution comparisons.
Keywords: Multivariate Almost Stochastic Dominance; Transfers; Sufficient Conditions for Dominance; Choice between Lotteries; Mean and Variance

Decisions with Several Objectives under Uncertainty: Sufficient Conditions for Multivariate Almost Stochastic Dominance Based on Means and Variances

Alfred Müller[1], Marco Scarsini[2], Ilia Tsetlin[3], and Robert L. Winkler[4]

[1]Department Mathematik, Universität Siegen, 57072 Siegen, Germany
[2]Dipartimento di Economia e Finanza, Luiss University, 00197 Roma, Italy
[3]INSEAD, Singapore 138676
[4]Fuqua School of Business, Duke University, Durham, North Carolina 27708

May 23, 2021

Abstract

Important business, public policy, and personal decisions typically involve multiple objectives, which in turn can be represented by multiple attributes, and uncertainty. Assessing both multi-attribute utility and multivariate distributions for the attributes can be challenging. Moreover, big decisions are often made by boards or committees with members holding divergent views and preferences and facing pressures from different stakeholders. Thus, a full-blown traditional decision analysis that leads to the computation of expected utility is very difficult at best and often not possible. We develop sufficient conditions for multivariate almost stochastic dominance (MASD) based on marginal distributions of the attributes or just on their means and variances. To apply MASD, one only needs to assess bounds on marginal utilities. Alternatively, preferences can be explained and elicited via transfers. Realistic examples illustrate our results, which provide tools for "fast and frugal" screening and evaluation of the available options, while properly accounting for tradeoffs and riskiness. Such tools, consistent with normative decision analysis, are useful when making important decisions in today's fast-moving and often complex world.

Subject classifications: Decision analysis: stochastic dominance, utility, risk. Probability: distribution comparisons.
Keywords: multivariate almost stochastic dominance, transfers, sufficient conditions for dominance, choice between lotteries, mean and variance

1 Introduction

When faced with an important choice, decision makers are typically interested in more than just one attribute. For example, a company choosing between two risky projects, A and B, might be interested in the net present value (NPV) of profits for the first five years and the market share (MS) at the end of the fifth year. Traditional decision analysis would suggest: a) assessing the bivariate distribution of NPV and MS for each project and b) eliciting the two-attribute utility function of the company. Our approach partially bypasses these steps.

The following scenarios illustrate the type of situations in which our approach can be helpful. We will return to these examples as we present our results.

(a) Alice, the company's chairman of the board, knows her utility function for the company. However, assessments of joint distributions of NPV and MS for either project A or B are not available. Our results will sometimes allow Alice to make a choice even if the only information she has consists of the marginal distributions of NPV and MS under projects A and B, but not their joint distribution, or just the means and variances of these distributions.

(b) The board makes decisions together. Some board members have their own assessed utility functions for the company, and these assessments are not identical. Other board members are not sure about their risk preferences and tradeoffs among the attributes, but they can agree about some constraints on them and about the marginal distributions or at least the means and variances. We provide conditions under which the board can unanimously rank risky projects A and B.

(c) A data analytics startup develops inventory and allocation solutions, with the objective of maximizing profit (P) and net promotion score (NPS), as well as some other attributes if requested by a client. Their current approach is to maximize the expected weighted sum (e.g., P+wNPS, where w represents the NPS/P tradeoff). However, often a client is not satisfied with this solution, arguing that this approach does not consider the risks associated with different options or the client's attitude toward these risks. At the same time, the members of the startup team feel that they cannot apply a full-blown decision analysis approach, which would require assessing a multiattribute utility function and the joint distribution of the attributes under each option. As part of their analysis, they have estimates of means and variances of different options. Using our results, they can narrow down the choice to a few non-dominated options and see how the optimal strategies vary with different parameters.

Because important decisions usually involve multiple attributes, extensions of stochastic dominance (SD) to the multivariate case have received some attention. Such extensions are tricky, as there are many multivariate stochastic orders (see, e.g., Müller and Stoyan, 2002, Shaked and Shanthikumar, 2007). Studies of multivariate stochastic dominance (MSD) include Levy and Paroush (1974), Levhari et al. (1975), Mosler (1984), Scarsini (1988), and Baccelli and Makowski (1989). Denuit et al. (2013) develop MSD, using a stochastic order that is a natural extension of the standard order typically used for univariate SD. The theory of SD has a counterpart in the literature about inequality measurement. Recent multivariate analyses of it can be found in Faure and Gravel (2021) and Mosler (2021).

The SD order provides a partial ranking of distributions that can be helpful when only partial information is known about a decision maker's utility function. There is a big jump from first-degree stochastic dominance (FSD) (with increasing utility) to second-degree stochastic dominance (SSD) (with increasing and concave utility). Many decision makers are mostly risk averse but cannot assert that they would dislike any risk, an indication of convex segments in their utility functions. The almost stochastic dominance (ASD) relation can provide a continuum of SD rules covering preferences from FSD to SSD (Leshno and Levy, 2002, Müller et al., 2017, Huang et al., 2020, Mao and Wang, 2020). The importance of MASD is due to the fact that it allows us to rank multivariate distributions when utility does not satisfy multivariate SSD but is "close" to doing so. Tsetlin and Winkler (2018) develop MASD, considering both concave and convex versions. When multiple decision makers are involved, it helps us rank distributions "by most decision makers" in the multivariate context.

Often only partial information is known about the distributions that we want to rank for decision-making purposes. For example, we might know the means and variances of the distributions but not their shapes. The seminal paper by Markowitz (1952) inspired a strong focus on the mean and variance for decision making in finance. Müller et al. (2021) provide a ranking in the single-attribute case when only means and variances are known by bounding how much marginal utility can change.

Even if the distributions we want to rank are known, integral conditions for MSD and MASD do not exist in most situations. In such cases sufficient conditions, which are easy to check, are helpful. The sufficient conditions that we develop in this paper are especially practical, as they require knowing only the marginal distributions of the attributes, or just their means and variances.

In Section 2 we provide definitions for limiting how much marginal utilities can change, for dominance based on these limitations, and we develop the corresponding transfers. In Section 3 we develop sufficient conditions for the case where the full marginal distributions are known and for the more common case where we only know their means and variances. We also provide bounds on multiattribute utilities that are additive across the multiple attributes, which is important because it allows us to develop sufficient conditions for MASD involving only marginal distributions of the multivariate random variables associated with the alternatives. In Section 4 we develop a path to a complete order and the corresponding transfers. Throughout we provide examples to illustrate our results. Concluding comments are given in Section 5. Appendix A shows a generalization of some characterizations to distributions with nonfinite support. The proofs of our results can be found in Appendix B.

2 Defining $\boldsymbol{\gamma}$-multivariate almost stochastic dominance

We consider a decision maker whose utility function u depends on N attributes $\boldsymbol{x} = (x_1, \ldots, x_N)$. The function $u\colon \mathbb{R}^N \to \mathbb{R}$ is assumed to be differentiable, and u_i' denotes its partial derivative with respect to its i-th argument:

$$u_i'(\boldsymbol{x}) := \frac{\partial u(\boldsymbol{x})}{\partial x_i}.$$

We now define $\boldsymbol{\gamma}$-*multivariate almost stochastic dominance* ($\boldsymbol{\gamma}$-MASD) for N-variate random vectors. In general, given a class $\mathcal{U}$ of utility functions, we say that $\boldsymbol{X} \leq_{\mathcal{U}} \boldsymbol{Y}$ if

$$\mathbb{E}[u(\boldsymbol{X})] \leq \mathbb{E}[u(\boldsymbol{Y})] \quad \text{for all } u \in \mathcal{U}.$$

Definition 1. For a given vector $\boldsymbol{\gamma} := (\gamma_1, \ldots, \gamma_N) \in [0,1]^N$, the symbol $\mathcal{U}_{\boldsymbol{\gamma}}$ denotes the set of utility functions such that, for all $i \in \{1, \ldots, N\}$, we have

$$0 < \gamma_i u_i'(\boldsymbol{y}) \leq u_i'(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N. \tag{2.1}$$

Notice that the condition in inequality (2.1) is equivalent to

$$\frac{\inf u_i'(\boldsymbol{x})}{\sup u_i'(\boldsymbol{x})} \geq \gamma_i. \tag{2.2}$$

For $\boldsymbol{\gamma} \in [0,1]^N$, the random vector $\boldsymbol{X}$ is dominated by the random vector $\boldsymbol{Y}$ in the sense of $\boldsymbol{\gamma}$-MASD if $\boldsymbol{X} \leq_{\mathcal{U}_{\boldsymbol{\gamma}}} \boldsymbol{Y}$. For the sake of simplicity, we will write $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ instead of $\boldsymbol{X} \leq_{\mathcal{U}_{\boldsymbol{\gamma}}} \boldsymbol{Y}$.

Definition 1 corresponds to MASD, as defined by Tsetlin and Winkler (2018), with $\gamma_i = \varepsilon_1/(1-\varepsilon_1)$ for all $i \in \{1, \ldots, N\}$. In the univariate case ($N = 1$), it corresponds to almost first-degree stochastic dominance (AFSD), as defined by Leshno and Levy (2002).

Notice that, if $\boldsymbol{\gamma} \leq \boldsymbol{\lambda}$ componentwise, then $\mathcal{U}_{\boldsymbol{\lambda}} \subset \mathcal{U}_{\boldsymbol{\gamma}}$. Therefore

$$\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y} \implies \boldsymbol{X} \leq_{\boldsymbol{\lambda}} \boldsymbol{Y}.$$

Example 2. (a) A company evaluates a project by focusing on two attributes x_1 and x_2, where x_1 is the NPV of profits for the next five years and x_2 is the MS in percentage at the end of the period. The relevant ranges for these quantities are $-100 \leq x_1 \leq 500$ and $10 \leq x_2 \leq 50$. Brandon, the CEO of the company, is risk seeking with respect to NPV, risk averse with respect to MS, and correlation neutral. His utility function is

$$u(x_1, x_2) = \exp(x_1/600) + w(1 - \exp(-x_2/40)),$$

and he is unsure about the value of w, which reflects the relative importance of x_2 versus x_1. Since

$$\frac{\inf u_1'(\boldsymbol{x})}{\sup u_1'(\boldsymbol{x})} = \exp((-100-500)/600) \approx 0.37, \qquad \frac{\inf u_2'(\boldsymbol{x})}{\sup u_2'(\boldsymbol{x})} = \exp((10-50)/40) \approx 0.37,$$

we have that $u \in \mathcal{U}_{(0.37,0.37)}$ for every positive w.

(b) Alice's utility function is given by

$$u(x_1, x_2) = (x_1 + 200)^{0.7} + 4x_2^{1.1} + 0.04(x_1 + 200)^{0.7} x_2^{1.1}.$$

Alice is risk averse with respect to NPV, risk seeking with respect to MS, and correlation loving. In this case

$$\frac{\inf u_1'(\boldsymbol{x})}{\sup u_1'(\boldsymbol{x})} = \frac{0.7 \times 700^{-0.3}(1 + 0.04 \times 10^{1.1})}{0.7 \times 100^{-0.3}(1 + 0.04 \times 50^{1.1})} \approx 0.21,$$
$$\frac{\inf u_2'(\boldsymbol{x})}{\sup u_2'(\boldsymbol{x})} = \frac{1.1 \times 10^{0.1}(4 + 0.04 \times 100^{0.7})}{1.1 \times 50^{0.1}\,(4 + 0.04 \times 700^{0.7})} \approx 0.53,$$

and $u \in \mathcal{U}_{(0.21,0.53)}$.

(c) Note that $\mathcal{U}_{(0.37,0.37)} \subset \mathcal{U}_{(0.21,0.37)}$ and $\mathcal{U}_{(0.21,0.53)} \subset \mathcal{U}_{(0.21,0.37)}$. Therefore, if two projects can be ranked as $\boldsymbol{X} \leq_{(0.21,0.37)} \boldsymbol{Y}$, then both Brandon and Alice will prefer $\boldsymbol{Y}$ to $\boldsymbol{X}$.

We have defined an SD rule by a set of utility functions with bounded marginal utilities, and illustrated how one can check that a particular utility belongs to this set. The corresponding preferences can be also characterized via transfers, which might be easier to explain and use for elicitation of decision makers' preferences. The idea of using transfers to characterize SD can be traced back to the seminal paper Rothschild and Stiglitz (1970), who have shown that increasing risk can be decomposed into a sequence of mean-preserving spreads. The name *transfer* for such operations like mean-preserving spreads was originally more common in the related literature on inequality measurement, where these transfers have the meaning of real transfers of income or wealth; see Atkinson (1970), a famous companion paper to Rothschild and Stiglitz (1970). It can be shown for many types of SD that, in the case of distributions assuming only a finite number of values, the dominance rule holds if and only if one distribution can be obtained from the other by a sequence of simple transfers. For multivariate FSD Østerdal (2010) shows that this holds for increasing transfers, i.e., transfers that shift some probability mass from some point $\boldsymbol{x}$ to some point $\boldsymbol{y} > \boldsymbol{x}$, meaning that we have a *good transfer* to a better situation. For first-degree or second-degree ASD one typically also allows for decreasing transfers shifting some probability mass from some point $\boldsymbol{x}$ to some point $\boldsymbol{y} < \boldsymbol{x}$ as long as this is compensated or overcompensated by corresponding inreasing transfers. See, e.g., Müller et al. (2017) for the univariate case or Müller and Scarsini (2012) for the multivariate case of inframodular transfers. We will show now that similar characterizations hold for the multivariate versions of SD considered in this paper. The proofs will be based on Müller (2013), where a general theory of such transfers and their relation to SD rules induced by classes of utility functions is developed.

Given two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N$ we use the notation $\boldsymbol{x} < \boldsymbol{y}$ to indicate

$$x_i \leq y_i, \quad \text{for } i = 1, \ldots, N, \quad \text{and} \quad \boldsymbol{x} \neq \boldsymbol{y}.$$

The symbol $\boldsymbol{e}_i$ denotes the i-th vector of the canonical basis.

Definition 3. Consider two discrete cumulative distribution functions F and G with respective mass functions f and g.

(a) We say that G is obtained from F via an *increasing transfer* if there exist $\boldsymbol{x}_1 < \boldsymbol{x}_2$ and $\eta > 0$ such that

$$\begin{aligned} g(\boldsymbol{x}_1) &= f(\boldsymbol{x}_1) - \eta, \\ g(\boldsymbol{x}_2) &= f(\boldsymbol{x}_2) + \eta, \\ g(\boldsymbol{z}) &= f(\boldsymbol{z}) \quad \text{for all other values } \boldsymbol{z}. \end{aligned}$$

(b) We say that G is obtained from F via a γ_i*-transfer along dimension* i if there exist $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4 \in \mathbb{R}^N$, $h, \eta_1, \eta_2 > 0$ such that

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 + h\boldsymbol{e}_i, \quad \eta_2(\boldsymbol{x}_4 - \boldsymbol{x}_3) = \gamma_i \eta_1 (\boldsymbol{x}_2 - \boldsymbol{x}_1), \tag{2.3}$$

and

$$\begin{aligned} g(\boldsymbol{x}_1) &= f(\boldsymbol{x}_1) - \eta_1, \\ g(\boldsymbol{x}_2) &= f(\boldsymbol{x}_2) + \eta_1, \\ g(\boldsymbol{x}_3) &= f(\boldsymbol{x}_3) + \eta_2, \\ g(\boldsymbol{x}_4) &= f(\boldsymbol{x}_4) - \eta_2, \\ g(\boldsymbol{z}) &= f(\boldsymbol{z}) \quad \text{for all other values } \boldsymbol{z}. \end{aligned}$$

We say that G is obtained from F via a $\boldsymbol{\gamma}$*-transfer* if G is obtained from F via a γ_i-transfer along some dimension $i \in \{1, \dots, N\}$.

Fig. 1 gives an example of γ_1-transfer with $N = 2$, $\gamma_1 = 2/3$, $\eta_1 = \eta_2$. This multivariate transfer is the natural generalization of the univariate (convex or concave) γ-transfer (or equivalently the univariate AFSD transfer (Müller et al., 2017)). It simply consists of a decreasing transfer from $\boldsymbol{x}_4$ to $\boldsymbol{x}_3$ which is compensated by an increasing transfer from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ concerning the same component i. It leads to a univariate γ-transfer of the i-th marginal as described in Müller et al. (2017)) and does not affect any of the other marginals.

We now characterize the order $\leq_{\boldsymbol{\gamma}}$ in terms of probability transfers.

Theorem 4. *Let the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ assume only a finite number of values. Then $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ if and only if the distribution of $\boldsymbol{Y}$ can be obtained from the distribution of $\boldsymbol{X}$ by a finite number of increasing transfers and $\boldsymbol{\gamma}$-transfers.*

Theorem 4 illustrates that preferences consistent with $\boldsymbol{\gamma}$-MASD can be thought of as preferences for multivariate $\boldsymbol{\gamma}$-transfers. Later we state a similar result in Theorem 19 and discuss a generalization to distributions with nonfinite support in Theorem 24.

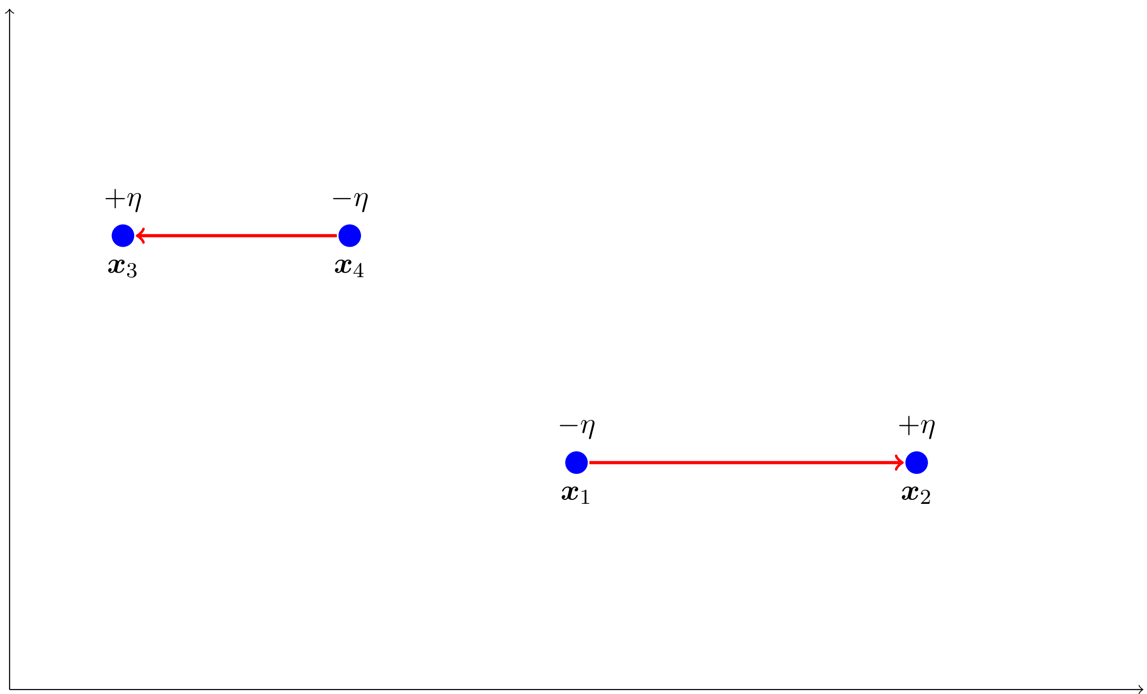

Figure 1: Example of $\boldsymbol{\gamma}$-transfer with $\gamma_1 = 2/3$, $\eta_1 = \eta_2 = \eta$.

3 Sufficient dominance conditions

The random vectors $\boldsymbol{X}, \boldsymbol{Y}$ are assumed to have components with finite means and variances:

$$\mu_{X_i} \coloneqq \mathbb{E}[X_i], \quad \mu_{Y_i} \coloneqq \mathbb{E}[Y_i], \quad \sigma^2_{X_i} \coloneqq \mathbb{V}[X_i], \quad \sigma^2_{Y_i} \coloneqq \mathbb{V}[Y_i], \tag{3.1}$$

and the marginal distributions will be denoted by

$$F_i(t) \coloneqq \mathbb{P}(X_i \le t), \quad G_i(t) \coloneqq \mathbb{P}(Y_i \le t),$$

for all $i \in \{1, \ldots, N\}$.

3.1 Conditions for $\boldsymbol{\gamma}$-dominance

We now provide various sufficient conditions for $\boldsymbol{X} \le_{\boldsymbol{\gamma}} \boldsymbol{Y}$.

Theorem 5. *Assume that the marginal distributions of the components of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are known and that $\mu_{X_i} \le \mu_{Y_i}$ for all $i = 1, \ldots, N$. Let $\delta_i \coloneqq \inf\{x \colon F_i(x) + G_i(x) \ge 1\}$ and let*

$$\gamma_i \coloneqq \frac{\mathbb{E}[(\delta_i - Y_i)_+] + \mathbb{E}[(X_i - \delta_i)_+]}{\mathbb{E}[(Y_i - \delta_i)_+] + \mathbb{E}[(\delta_i - X_i)_+]}, \tag{3.2}$$

for $i = 1, \ldots, N$. Then $\boldsymbol{X} \le_{\boldsymbol{\gamma}} \boldsymbol{Y}$.

The proof of the above theorem is based on the following lemma, which establishes upper and lower bounds for utility functions $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ that are additive. This allows us to develop sufficient conditions for MASD involving only marginal distributions of random vectors instead of having to deal with the full joint distribution of these variables. For example, we do not need to know anything about the dependence among the variables.

Lemma 6. *Let*

$$v_U(x;\gamma) := \begin{cases} \gamma x & \text{if } x \leq 0, \\ x & \text{if } x > 0, \end{cases}$$

$$v_L(x;\gamma) := \begin{cases} x & \text{if } x \leq 0, \\ \gamma x & \text{if } x > 0. \end{cases}$$

For any $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, let $b_i := \sup_{\boldsymbol{x} \in \mathbb{R}^N} u_i'(\boldsymbol{x})$ and fix some $\boldsymbol{z} \in \mathbb{R}^N$. Then, for any $\boldsymbol{x} \in \mathbb{R}^N$, we have

$$\sum_{i=1}^{N} b_i v_L(x_i - z_i; \gamma_i) \leq u(\boldsymbol{x}) - u(\boldsymbol{z}) \leq \sum_{i=1}^{N} b_i v_U(x_i - z_i; \gamma_i). \tag{3.3}$$

If the marginal distributions are symmetric and location-scale, such as normal, then the sufficient bounds in Theorem 5 are easier to compute, as shown in Proposition 7. A univariate distribution function F is said to belong to the symmetric location-scale H-family if

$$F(x) = H\left(\frac{x-\mu}{\sigma}\right), \quad \text{with } H(x) = 1 - H(-x).$$

In other words, H is the distribution function of a random variable Z as well as of $-Z$, and F is the distribution function of $\mu + \sigma Z$.

Proposition 7. *Let F_i and G_i belong to the same symmetric location-scale H-family and let*

$$\eta(t) := \frac{\mathbb{E}[(t-Z)_+]}{\mathbb{E}[(Z-t)_+]},$$

where Z has distribution function H. If

$$\tau_i = \frac{\mu_{X_i} - \mu_{Y_i}}{\sigma_{X_i} + \sigma_{Y_i}},$$

then, in (3.2), we have $\gamma_i = \eta(\tau_i)$.

We now consider the case where the marginal distributions of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not completely specified, but only the means and variances are known.

Define

$$\zeta(t) := \frac{1}{1 + 2t\left(t + \sqrt{t^2 + 1}\right)}. \tag{3.4}$$

Theorem 8. *Let the two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ have finite means and variances. Moreover, for all $i = 1, \dots, N$, let $\mu_{X_i} \le \mu_{Y_i}$ and let*

$$\tau_i = \frac{\mu_{Y_i} - \mu_{X_i}}{\sigma_{X_i} + \sigma_{Y_i}}.$$

If $\gamma_i = \zeta(\tau_i)$, $i = 1, \dots, N$, then $\boldsymbol{X} \le_\gamma \boldsymbol{Y}$.

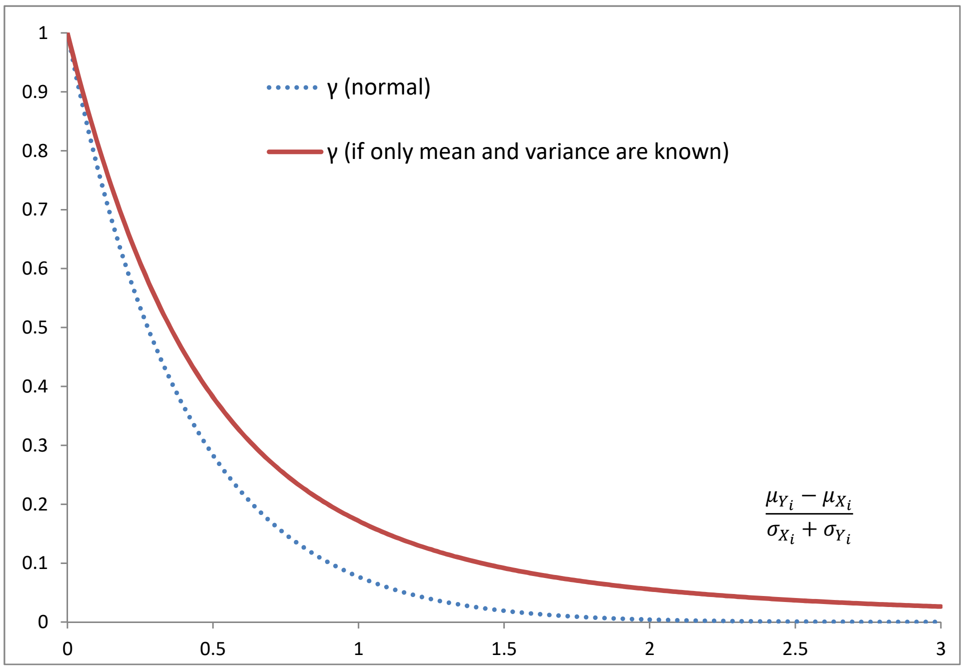

Figure 2: γ_i as a function of $(\mu_{Y_i} - \mu_{X_i})/(\sigma_{X_i} + \sigma_{Y_i})$.

Fig. 2 shows the values of γ_i as functions of $(\mu_{Y_i} - \mu_{X_i})/(\sigma_{X_i} + \sigma_{Y_i})$ when the distributions of X_i and Y_i are normal (according to Proposition 7) and when only their means and variances are known (Theorem 8). Fig. 2 extends Figure 3 in Müller et al. (2021), which deals with the univariate case when the dominated distribution is degenerate.

Example 9. To continue Example 2(c), let $\boldsymbol{X}$ denote the return from project A and $\boldsymbol{Y}$ the return from project B, and recall that Alice and Brandon's utility functions belong to $\mathcal{U}_{(0.21,0.37)}$. According to Theorem 8, $\boldsymbol{X} \le_{(0.21,0.37)} \boldsymbol{Y}$ if

$$\frac{\mu_{Y_1} - \mu_{X_1}}{\sigma_{X_1} + \sigma_{Y_1}} \ge 0.87 \quad \text{and} \quad \frac{\mu_{Y_2} - \mu_{X_2}}{\sigma_{X_2} + \sigma_{Y_2}} \ge 0.52.$$

In this case they both can make the decision knowing only the projects' means and variances. Furthermore, if other board members feel that the utility class $\mathcal{U}_{(0.21,0.37)}$ includes their preferences, then the choice of project B over A is unanimous. If, in addition to the means and variances, it is known that the marginal distributions are normal, then by Proposition 7 project B is preferred if

$$\frac{\mu_{Y_1} - \mu_{X_1}}{\sigma_{X_1} + \sigma_{Y_1}} \geq 0.62 \quad \text{and} \quad \frac{\mu_{Y_2} - \mu_{X_2}}{\sigma_{X_2} + \sigma_{Y_2}} \geq 0.4.$$

If one of the alternatives is characterized by a degenerate distribution (i.e., a sure payoff vector), then the sufficient conditions simplify further, as illustrated in Propositions 10 and 11.

Proposition 10. *Assume that the marginal distributions of the components of $\boldsymbol{X}$ are known and that $\boldsymbol{c}$ is a sure payoff vector.*

(a) *Let $c_i \leq \mu_{X_i}$ for all $i = 1, \ldots, N$. If*

$$\gamma_i = \frac{\mathbb{E}[(c_i - X_i)_+]}{\mathbb{E}[(X_i - c_i)_+]}, \quad i = 1, \ldots, N, \tag{3.5}$$

then $\boldsymbol{c} \leq_{\boldsymbol{\gamma}} \boldsymbol{X}$.

(b) *Let $\mu_{X_i} \leq c_i$ for all $i = 1, \ldots, N$. If*

$$\gamma_i = \frac{\mathbb{E}[(X_i - c_i)_+]}{\mathbb{E}[(c_i - X_i)_+]}, \quad i = 1, \ldots, N, \tag{3.6}$$

then $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{c}$.

Proof. If a random vector $\boldsymbol{Y}$ is degenerate in $\boldsymbol{c}$, then for each $i \in \{1, \ldots, N\}$, we have $\mathbb{E}[(Y_i - c_i)_+] = \mathbb{E}[(c_i - Y_i)_+] = 0$. Therefore the results follow directly from Theorem 5. □

Proposition 11. *Assume that the random vector $\boldsymbol{X}$ has finite means and variances and that $\boldsymbol{c}$ is a sure payoff. Define*

$$t_i = \frac{\mu_{X_i} - c_i}{\sigma_{X_i}}. \tag{3.7}$$

(a) *Let $c_i \leq \mu_{X_i}$ for all $i = 1, \ldots, N$. If $\gamma_i = \zeta(t_i)$, as defined in Eq. (3.4), then $\boldsymbol{c} \leq_{\boldsymbol{\gamma}} \boldsymbol{X}$.*

(b) *Let $\mu_{X_i} \leq c_i$ for all $i = 1, \ldots, N$. If $\gamma_i = \zeta(-t_i)$, as defined in Eq. (3.4), then $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{c}$.*

Proof. The result is an immediate corollary of Theorem 8. □

Remark 12. Notice that in Eq. (3.6) the right hand side is equal to the Omega ratio $\Omega_{X_i}(c_i)$, as defined in Shadwick and Keating (2002), whereas in Eq. (3.5) the right hand side is $1/\Omega_{X_i}(c_i)$. Note that the right hand side of Eq. (3.7) can be interpreted as the Sharpe ratio. The connection between univariate ASD, the Omega ratio, and the Sharpe ratio is discussed in Müller et al. (2021).

In this subsection we established sufficient conditions for $\boldsymbol{\gamma}$-MASD. These conditions are based on marginal distributions, which make them especially easy to implement. In the next subsection we discuss a bit more how powerful this is and the corresponding intuition, given that usually the comparison of marginal distributions provides only necessary conditions for MSD.

3.2 Joint and marginal dominance relations

Given that the proof of Theorem 5 is based on Lemma 6, which provides separable bounds for the utility functions in $\mathcal{U}_{\boldsymbol{\gamma}}$, one may suspect that checking whether $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ holds is equivalent to separately checking whether $X_i \leq_{\gamma_i} Y_i$ for each $i \in \{1, \dots N\}$. The following counterexample shows that this is not the case.

Example 13. Let $N = 2$ and $\boldsymbol{\gamma} = (1/2, 1/2)$. Consider the binary random vectors $\boldsymbol{X}, \boldsymbol{X}', \boldsymbol{Y}, \boldsymbol{Y}'$ having the following distributions:

$$\mathbb{P}(\boldsymbol{X} = (0,0)) = \mathbb{P}(\boldsymbol{X} = (5,2)) = \frac{1}{2}, \quad \mathbb{P}(\boldsymbol{X}' = (0,2)) = \mathbb{P}(\boldsymbol{X}' = (5,0)) = \frac{1}{2},$$
$$\mathbb{P}(\boldsymbol{Y} = (2,0)) = \mathbb{P}(\boldsymbol{Y} = (4,2)) = \frac{1}{2}, \quad \mathbb{P}(\boldsymbol{Y}' = (2,2)) = \mathbb{P}(\boldsymbol{Y}' = (4,0)) = \frac{1}{2}.$$

Then $\boldsymbol{X}$ and $\boldsymbol{X}'$ have the same marginal distributions as well as $\boldsymbol{Y}$ and $\boldsymbol{Y}'$. With the characterizations via transfers one can easily see that

$$\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y} \quad \text{and} \quad \boldsymbol{X}' \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}',$$

but

$$\boldsymbol{X} \not\leq_{\boldsymbol{\gamma}} \boldsymbol{Y}'.$$

For a proof of the last statement consider the following utility function u:

$$u(x_1, x_2) = x_1 + x_2 + \max\{x_1 + x_2 - 4, 0\}.$$

All partial derivatives of this function u are bounded between 1 and 2, so we have $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, but

$$\mathbb{E}[u(\boldsymbol{X})] = 10 > 8 = \mathbb{E}\big[u(\boldsymbol{Y}')\big].$$

This shows that the ordering $\leq_{\boldsymbol{\gamma}}$ depends not only on the marginal distributions, but on the whole joint distributions of the random vectors.

Remark 14. There exist various necessary conditions for SD based on moments both in the univariate and the multivariate case (see, e.g., Fishburn, 1980, O'Brien, 1984, O'Brien and Scarsini, 1991). The perspective we take here is completely different, since we provide sufficient conditions.

4 From partial to complete ordering

In the univariate case, when $\gamma = 1$, we have

$$X \leq_\gamma Y \iff \mathbb{E}[X] \leq \mathbb{E}[Y].$$

This means that, in this case, there exists a complete order on the set of random variables with finite expectation. In the multivariate case the situation is more complicated, due to the fact that $\mathbb{R}^N$ is not completely ordered, so there is no natural way to order random vectors by their expectations. One possible way would be to consider weighted expectations $\mathbb{E}\Big[\sum_{i=1}^N w_i X_i\Big]$, as in portfolio analysis. The case where dominance holds for any possible choice of positive weights has been studied by Muliere and Scarsini (1989).

4.1 Defining $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$-dominance

To achieve a complete order of random vectors, we consider a new class of utility functions defined in terms of two parameters: a scalar γ and a vector $\boldsymbol{\beta}$. Then we define the corresponding SD relation, $(\gamma, \boldsymbol{\beta})$-*multivariate almost stochastic dominance* ($(\gamma, \boldsymbol{\beta})$-MASD).

Definition 15. For $\gamma \in [0, 1]$ and $\boldsymbol{\beta} \in \mathbb{R}_+^N$, let $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$ be the class of utility functions $u \colon \mathbb{R}^N \to \mathbb{R}$ such that

$$0 < \gamma\beta_i \leq u_i'(\boldsymbol{x}) \leq \beta_i \quad \text{for all } i \in \{1, \dots, N\}. \tag{4.1}$$

The random vector $\boldsymbol{X}$ is dominated by the random vector $\boldsymbol{Y}$ in the sense of $(\gamma, \boldsymbol{\beta})$-MASD ($\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$) if

$$\mathbb{E}[u(\boldsymbol{X})] \leq \mathbb{E}[u(\boldsymbol{Y})], \quad \text{for all } u \in \mathcal{U}_{\gamma,\boldsymbol{\beta}}.$$

Notice that, for any $\alpha > 0$, we have $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$ iff $\boldsymbol{X} \leq_{\gamma,\alpha\boldsymbol{\beta}} \boldsymbol{Y}$. This is coherent with the fact that two utility functions represent the same preferences if one is obtained from the other via a positive affine transformation.

4.2 Characterization via $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$-transfers

We now consider the class $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$ of utility functions defined in Definition 15. Notice that β_i is a scale factor that depends on the units that are used. Indeed, if $\tilde{u} : \mathbb{R}^N \to \mathbb{R}$ is a function such that

$$0 < \gamma \leq \tilde{u}_i'(\boldsymbol{x}) \leq 1 \quad \text{for all } i \in \{1, \dots, N\}, \tag{4.2}$$

then the function

$$u(x_1, \dots, x_N) \coloneqq \tilde{u}(\beta_1 x_1, \dots, \beta_N x_N)$$

fulfills (4.1). Thus, by changing units we can assume without loss of generality that u is a function with the property (4.2), i.e., with the property that all marginal utilities are bounded between γ and 1.

A function u that satisfies (4.2) also satisfies

$$\gamma u_i'(\boldsymbol{x}) \le u_j'(\boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \text{ and for all } i, j. \tag{4.3}$$

Vice versa, if a function satisfies (4.3), we can define

$$\beta \coloneqq \sup_{i,\boldsymbol{x}} u_i'(\boldsymbol{x})$$

and then

$$\gamma\beta \le u_j'(\boldsymbol{y}) \le \beta \quad \text{for all } \boldsymbol{y} \text{ and for all } j;$$

thus u/β satisfies (4.2). Hence the functions satisfying (4.3) build the convex cone generated by the functions satisfying (4.2) and therefore define the same SD rule.

Similarly, the convex cone generated by the functions in $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$ is given by the functions satisfying

$$\gamma\beta_j u_i'(\boldsymbol{x}) \le \beta_i u_j'(\boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \text{ and for all } i, j \in \{1, \ldots, N\}.$$

In the following discussion of transfers we will first restrict our attention to the class $\mathcal{U}_{\gamma,\mathbf{1}}$, i.e., the functions that satisfy property (4.2). In contrast to the $\boldsymbol{\gamma}$-transfer we will now allow that the decreasing transfer from $\boldsymbol{x}_4$ to $\boldsymbol{x}_3$ concerning component i can also be compensated by an increasing transfer from $\boldsymbol{x}_1$ to $\boldsymbol{x}_2$ concerning some other component j.

Definition 16. Consider two discrete cumulative distribution functions F and G with respective mass functions f and g. We say that G is obtained from F via a $(\gamma, \mathbf{1})$-*transfer* (along dimensions i, j) if there exist $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$, $\varepsilon_1, \varepsilon_2 > 0$ and $\eta_1, \eta_2 > 0$ such that, for some $i, j \in \{1, \ldots, N\}$,

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 + \varepsilon_1 \boldsymbol{e}_i, \quad \boldsymbol{x}_4 = \boldsymbol{x}_3 + \varepsilon_2 \boldsymbol{e}_j, \quad \eta_2\varepsilon_2 = \gamma\eta_1\varepsilon_1,$$

and

$$\begin{aligned}
g(\boldsymbol{x}_1) &= f(\boldsymbol{x}_1) - \eta_1, \\
g(\boldsymbol{x}_2) &= f(\boldsymbol{x}_2) + \eta_1, \\
g(\boldsymbol{x}_3) &= f(\boldsymbol{x}_3) + \eta_2, \\
g(\boldsymbol{x}_4) &= f(\boldsymbol{x}_4) - \eta_2, \\
g(\boldsymbol{z}) &= f(\boldsymbol{z}) \quad \text{for all other values } \boldsymbol{z}.
\end{aligned}$$

Fig. 3 shows an example of a $(\gamma, \mathbf{1})$-transfer with $N = 2$, $\varepsilon_1 = 1.5$, $\varepsilon_2 = 1$, $\gamma = 2/3$, $\eta_1 = \eta_2 = \eta$.

With a proof similar to the proof of Theorem 4, we get the following result.

Theorem 17. *Let the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ assume a finite number of values. Then $\boldsymbol{X} \le_{\gamma,\mathbf{1}} \boldsymbol{Y}$ if and only if the distribution of $\boldsymbol{Y}$ can be obtained from the distribution of $\boldsymbol{X}$ by a finite number of*

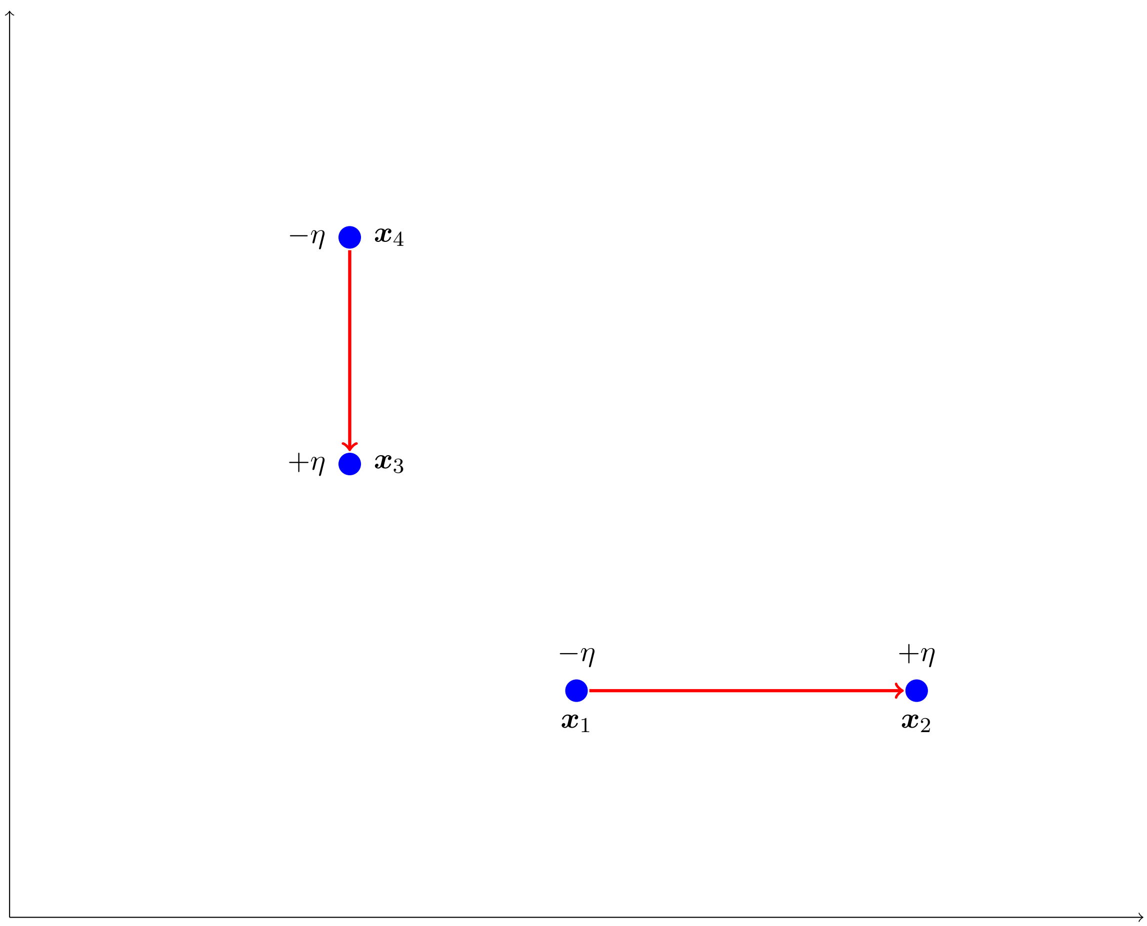

Figure 3: Example of $(\gamma, \mathbf{1})$-transfer with $\varepsilon_1 = 1.5$, $\varepsilon_2 = 1$, $\gamma = 2/3$, $\eta_1 = \eta_2 = \eta$.

increasing transfers and $(\gamma, \mathbf{1})$*-transfers.*

Notice that

$$\mathbb{E}[u(\boldsymbol{X})] \leq \mathbb{E}[u(\boldsymbol{Y})] \quad \text{ for all } u \in \mathcal{U}_{(\gamma,\boldsymbol{\beta})}$$

is equivalent to

$$\mathbb{E}[\tilde{u}(\beta_1 X_1, \ldots, \beta_N X_N)] \leq \mathbb{E}[\tilde{u}(\beta_1 X_1, \ldots, \beta_N X_N)] \quad \text{ for all } \tilde{u} \in \mathcal{U}_{(\gamma,\mathbf{1})}.$$

From this equivalence we get the general $(\gamma, \boldsymbol{\beta})$-transfers as follows.

Definition 18. Consider two discrete cumulative distribution functions F and G with respective mass functions f and g. We say that G is obtained from F via a $(\gamma, \boldsymbol{\beta})$*-transfer* if there are $i, j \in \{1, \ldots, N\}$ and exist $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4 \in \mathbb{R}^N$, $\varepsilon_1, \varepsilon_2, \eta_1, \eta_2 > 0$ such that, for some $i, j \in \{1, \ldots, N\}$,

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 + \varepsilon_1 \boldsymbol{e}_i, \quad \boldsymbol{x}_4 = \boldsymbol{x}_3 + \varepsilon_2 \boldsymbol{e}_j, \quad \eta_2 \varepsilon_2 \beta_j = \gamma \eta_1 \varepsilon_1 \beta_i,$$

and

$$
\begin{aligned}
g(\boldsymbol{x}_1) &= f(\boldsymbol{x}_1) - \eta_1, \\
g(\boldsymbol{x}_2) &= f(\boldsymbol{x}_2) + \eta_1, \\
g(\boldsymbol{x}_3) &= f(\boldsymbol{x}_3) + \eta_2, \\
g(\boldsymbol{x}_4) &= f(\boldsymbol{x}_4) - \eta_2, \\
g(\boldsymbol{z}) &= f(\boldsymbol{z}) \quad \text{for all other values } \boldsymbol{z}.
\end{aligned}
$$

Theorem 19. *Let the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ assume a finite number of values. Then $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$ if and only if the distribution of $\boldsymbol{Y}$ can be obtained from the distribution of $\boldsymbol{X}$ by a finite number of increasing transfers and $(\gamma, \boldsymbol{\beta})$-transfers.*

4.3 Sufficient conditions for $(\gamma, \boldsymbol{\beta})$-dominance

Theorem 20. *Assume that the marginal distributions of the components of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are known. Let $\delta_i \coloneqq \inf\{x \colon F_i(x) + G_i(x) \geq 1\}$ and let*

$$
\gamma \coloneqq \frac{\sum_{i=1}^{N} \beta_i(\mathbb{E}[(\delta_i - Y_i)_+] + \mathbb{E}[(X_i - \delta_i)_+])}{\sum_{i=1}^{N} \beta_i(\mathbb{E}[(Y_i - \delta_i)_+] + \mathbb{E}[(\delta_i = X_i)_+])}.
$$

If

$$
\sum_{i=1}^{N} \beta_i \mu_{X_i} \leq \sum_{i=1}^{N} \beta_i \mu_{Y_i}, \tag{4.4}
$$

then $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$.

We next address the case where only means and variances are known.

Theorem 21. *Let the two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ have finite means and variances. Let*

$$
\gamma = \frac{\sum_{i=1}^{N} \beta_i \Big(\sqrt{(\sigma_{X_i} + \sigma_{Y_i})^2 + (\mu_{Y_i} - \mu_{X_i})^2} - (\mu_{Y_i} - \mu_{X_i}) \Big)}{\sum_{i=1}^{N} \beta_i \Big(\sqrt{(\sigma_{X_i} + \sigma_{Y_i})^2 + (\mu_{Y_i} - \mu_{X_i})^2} + (\mu_{Y_i} - \mu_{X_i}) \Big)}.
$$

If (4.4) holds, then $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$.

Example 22. Returning to the setting of Example 2, suppose the board has to decide whether project $\boldsymbol{Y}$ (with $N = 2$, Y_1 being NPV and Y_2 being MS) is worth undertaking, thus comparing $\boldsymbol{Y}$ to the status quo $\boldsymbol{0}$. It is hard to assess the joint distribution of $\boldsymbol{Y}$, but estimates of the means and variances are available:

$$
\mu_{Y_1} = -5, \quad \sigma_{Y_1} = 4, \quad \mu_{Y_2} = 2, \quad \sigma_{Y_2} = 1.
$$

In expectation, this risky project will decrease NPV by \$5 million and increase MS by 2%. To apply

Theorem 21 we choose the parameters β_1, β_2 as the sup of the partial derivatives of the utility function:

$$\frac{\beta_2}{\beta_1} = \frac{\sup(u_2'(\boldsymbol{x}))}{\sup(u_1'(\boldsymbol{x}))}.$$

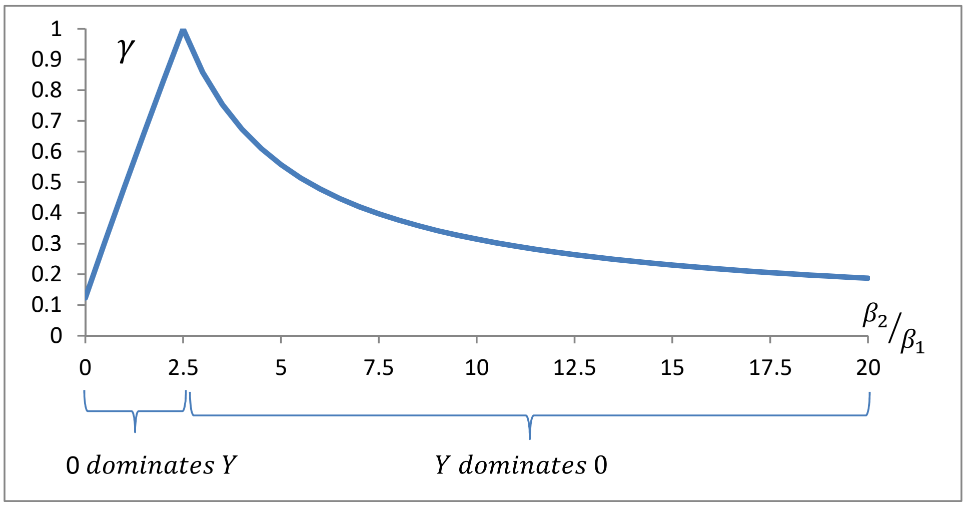

Figure 4: γ as a function of β_2/β_1.

Fig. 4 plots γ from Theorem 21 as a function of β_2/β_1. As we can see, for $\beta_2/\beta_1 < 2.5$, it is better not to undertake the project. For Alice (Example 2(b)),

$$\frac{\beta_2}{\beta_1} = \frac{1.1 \times 50^{0.1}\left(4 + 0.04 \times 700^{0.7}\right)}{0.7 \times 100^{-0.3}(1 + 0.04 \times 50^{1.1})} \approx 18 \quad (\$ \text{ million per } 1\% \text{ of MS}).$$

For Brandon (Example 2(a))

$$\frac{\beta_2}{\beta_1} = w\frac{(1/40) \times \exp\left(-10/40\right)}{(1/600) \times \exp\left(500/600\right)} \approx 5w.$$

Brandon thinks that the parameter w is at least 4, which gives $\beta_2/\beta_1 \geq 20$. Other board members agree that 1% of MS is worth at least \$20 million of NPV. Therefore, for all board members, $\beta_2/\beta_1 \geq 18$. From Fig. 4, if $\beta_2/\beta_1 = 18$, then $\gamma = 0.2$, which implies that $\boldsymbol{Y}$ dominates $\mathbf{0}$ with respect to the utility class $\mathcal{U}_{0.2,\boldsymbol{\beta}}$ with $\beta_2/\beta_1 \geq 18$. If, as in Example 9, all board members agree that $\mathcal{U}_{(0.21,0.37)}$ includes their preferences, then the support for the project is unanimous, and further information on the joint distribution of the components of $\boldsymbol{Y}$ is not needed.

Example 23. Returning to the data analytics startup from the scenario (c) in the Introduction, the current approach is to maximize $\mathbb{E}\left[\sum_{i=1}^{N} w_i X_i\right]$, as mentioned at the beginning of Section 4. At the same time, some bounds on variances might be available (Arlotto et al., 2014). To take into account the riskiness of different solutions, the startup team can apply Theorem 21 for $\beta_i = w_i$, $i = 1, \ldots, N$. For any two contending solutions that will give γ; by varying γ, they can identify dominated and dominating sets of the available alternatives, and see how these sets change as $\boldsymbol{\beta}$ varies. Then these results can be discussed with a client.

5 Conclusions

SD is a useful concept, especially in a multivariate context, where assessing multiattribute utility is challenging and different stakeholders might have divergent views. However, applying multivariate SD is difficult for two reasons: First, often distributions cannot be ranked (e.g., by FSD); this can be overcome by using $\boldsymbol{\gamma}$-MASD. Second, integral conditions for multivariate SD do not exist; to overcome this challenge, we develop sufficient conditions for $\boldsymbol{\gamma}$-MASD that are based on marginal distributions of the compared alternatives or just on their means and variances. This makes our conditions very practical, as full assessments of joint multivariate distributions are usually difficult. In the framework of portfolio analysis, Arvanitis et al. (2021) studied stochastic bounding of a portfolio by another, i.e., they look at conditions under which a set of portfolios contains one portfolio that stochastically dominates all portfolios in another set. When these conditions are not satisfied, they look for approximate bounds, in the spirit of ASD.

Another distinction of the multivariate case, compared to the univariate case, is that a real coordinate space is not completely ordered. To attain a path to a complete order, we need to constrain maximal marginal utilities for different attributes. Section 4 presents the corresponding definition of $(\gamma, \boldsymbol{\beta})$-MASD, its characterization via transfers, and sufficient conditions for comparing two risky alternatives.

Within the expected utility framework, $\boldsymbol{\gamma}$-MASD and $(\gamma, \boldsymbol{\beta})$-MASD translate into bounds on marginal utilities (Definitions 1 and 15). Alternatively, these preferences can be characterized via transfers (Definitions 3 and 18 and Theorems 4 and 19). Such transfers might be easier to explain to decision makers and use for elicitation of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$.

Examples 2, 9 and 22 illustrate our approach in a classical decision-making setting, where one needs to choose between two alternatives. We also discuss broader potential applications (Example 23) for screening of the most promising solutions in (potentially large-scale) optimization problems.

There is always a tension between a careful comparison and evaluation of available alternatives and a search for new solutions. With multiple attributes, the former is difficult and laborious. Our results provide tools for "fast and frugal" screening and evaluation, while properly accounting for tradeoffs and riskiness. As the world moves toward decisions with multiple objectives (e.g., many environmental, social and governance (ESG) criteria in addition to the financial performance of a company), such tools, consistent with normative decision analysis, should become even more in demand.

A Note on general distribution functions

The following theorem shows how the previous results on transfers can be adapted to the case of random variables that are not finite.

Theorem 24. *Let $B \subset \mathbb{R}^N$ be bounded and let $\mathcal{U}$ be a class of continuous increasing functions $u : B \to \mathbb{R}$. Let the random vectors $\boldsymbol{X}, \boldsymbol{Y}$ take values in B. Then $\boldsymbol{X} \leq_{\mathcal{U}} \boldsymbol{Y}$ if and only if there exist two sequences $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{Y}_n)_{n\in\mathbb{N}}$ such that*

$$\boldsymbol{X}_n \to \boldsymbol{X} \quad \textit{a.s.}, \quad \boldsymbol{Y}_n \to \boldsymbol{Y} \quad \textit{a.s.}, \quad \textit{and} \quad \boldsymbol{X}_n \leq_{\mathcal{U}} \boldsymbol{Y}_n \quad \textit{for all } n \in \mathbb{N}.$$

Proof. For bounded univariate random variables X, Y we can construct sequences $(X_n), (Y_n)$ with $X_n \leq X$, $Y_n \geq Y$, and $X_n \to X, Y_n \to Y$ a.s.. A concrete construction is given in the proof of Theorem 2.8 in Müller et al. (2017). We can apply this procedure componentwise to bounded random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$. Thus we get sequences such that, almost surely, $\boldsymbol{X}_n \leq \boldsymbol{X}$, $\boldsymbol{Y}_n \geq \boldsymbol{Y}$, $\boldsymbol{X}_n \to \boldsymbol{X}$, and $\boldsymbol{Y}_n \to \boldsymbol{Y}$. Since $\boldsymbol{X} \leq \boldsymbol{Y}$ a.s. implies $\boldsymbol{X} \leq_{\mathcal{U}} \boldsymbol{Y}$ we thus get sequences with

$$\boldsymbol{X}_n \leq_{\mathcal{U}} \boldsymbol{X} \leq_{\mathcal{U}} \boldsymbol{Y} \leq_{\mathcal{U}} \boldsymbol{Y}_n.$$

This shows the only-if-part. The if-part follows from the fact that any SD relation $\leq_{\mathcal{U}}$ is closed under convergence in distribution if $\mathcal{U}$ consists only of bounded continuous functions, see, e.g., Müller (1997, Theorem 4.2). □

B Proofs

Proofs of Section 2

The proof of Theorem 4 requires the following lemma.

Lemma 25. *Let $u : \mathbb{R}^N \to \mathbb{R}$ be continuously differentiable. Then $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ if and only if*

$$\eta_2(u(\boldsymbol{x}_4) - u(\boldsymbol{x}_3)) \leq \eta_1(u(\boldsymbol{x}_2) - u(\boldsymbol{x}_1)) \tag{B.1}$$

for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ satisfying (2.3) for some i and γ_i.

Proof. If part: Assume that u fulfills (B.1) for some i and γ_i. Then

$$\eta_2(\boldsymbol{x}_4 - \boldsymbol{x}_3) = \gamma_i\eta_1(\boldsymbol{x}_2 - \boldsymbol{x}_1) \implies \boldsymbol{x}_3 = \boldsymbol{x}_4 - \gamma_i\eta_1\boldsymbol{e}_i$$

and so (B.1) implies

$$\gamma_i \frac{\partial}{\partial x_i} u(\boldsymbol{x}_4) = \gamma_i \lim_{\eta_1 \to 0} \frac{u(\boldsymbol{x}_4) - u(\boldsymbol{x}_3)}{\gamma_i \eta_1} \leq \lim_{\eta_2 \to 0} \frac{u(\boldsymbol{x}_2) - u(\boldsymbol{x}_1)}{\eta_2} = \frac{\partial}{\partial x_i} u(\boldsymbol{x}_1).$$

As this holds for arbitrary $\boldsymbol{x}_1, \boldsymbol{x}_4$ and the derivatives are assumed to be continuous, by (2.1), we get $u \in \mathcal{U}_{\boldsymbol{\gamma}}$.

Only if part: Now assume that $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ is continuously differentiable. Let $\boldsymbol{h} := \boldsymbol{x}_2 - \boldsymbol{x}_1$. For $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ satisfying (2.3) for some i and γ_i, from $\eta_2(\boldsymbol{x}_4 - \boldsymbol{x}_3) = \gamma_i \eta_1 (\boldsymbol{x}_2 - \boldsymbol{x}_1)$, we get that

$$\boldsymbol{x}_4 - \boldsymbol{x}_3 = \frac{\gamma_i \eta_1}{\eta_2} (\boldsymbol{x}_2 - \boldsymbol{x}_1).$$

Thus, from (B.1) we can deduce

$$\begin{aligned}
\eta_1(u(\boldsymbol{x}_2) - u(\boldsymbol{x}_1)) &= \int_0^1 \frac{\partial}{\partial x_i} u(\boldsymbol{x}_1 + t\boldsymbol{h}) dt \\
&\geq \eta_1 \gamma_i \int_0^1 \frac{\partial}{\partial x_i} u\left(\boldsymbol{x}_3 + t \frac{\gamma_i \eta_1}{\eta_2} \boldsymbol{h}\right) \mathrm{d}t \\
&= \eta_2 \frac{\gamma_i \eta_1}{\eta_2} \int_0^1 \frac{\partial}{\partial x_i} u\left(\boldsymbol{x}_3 + t \frac{\gamma_i \eta_1}{\eta_2} \boldsymbol{h}\right) \mathrm{d}t \\
&= \eta_2(u(\boldsymbol{x}_4) - u(\boldsymbol{x}_3)).
\end{aligned}$$

□

Proof of Theorem 4. The proof follows from Müller (2013, Theorem 2.4.1), when applied to the class of functions generated by the transfers, i.e., the functions defined in Eq. (B.1). □

Proofs of Section 3

Proof of Lemma 6. Note that $u'_i(\boldsymbol{x}) \leq \sup(u'_i(\boldsymbol{x})) = b_i$ and that by inequality (2.2) we have $u'_i(\boldsymbol{x}) \geq \gamma_i b_i$. By a multivariate first-order Taylor expansion, $u(\boldsymbol{x}) - u(\boldsymbol{z}) = \sum_{i=1}^N u'_i(\boldsymbol{y})(x_i - z_i)$, where y_i is between x_i and z_i. Then, using $u'_i(\boldsymbol{y}) \leq b_i$ if $x_i > z_i$ and $u'_i(\boldsymbol{y}) \geq \gamma_i b_i$ if $x_i < z_i$ provides an upper bound, while using $u'_i(\boldsymbol{y}) \geq \gamma_i b_i$ if $x_i > z_i$ and $u'_i(\boldsymbol{y}) \leq b_i$ if $x_i < z_i$ provides a lower bound. □

Proof of Theorem 5. Given $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, let $b_i = \sup(u'_i(\boldsymbol{x}))$, and without loss of generality, assume $u(\boldsymbol{\delta}) = 0$. By Lemma 6 we have

$$\sum_{i=1}^N b_i v_L(x_i - \delta_i; \gamma_i) \leq u(\boldsymbol{x}) \leq \sum_{i=1}^N b_i v_U(x_i - \delta_i; \gamma_i).$$

First, we show that, for $i = 1, \ldots, N$, for any δ_i we have

$$\mathbb{E}[v_L(Y_i - \delta_i; \gamma_i)] = \mathbb{E}[v_U(X_i - \delta_i; \gamma_i)]$$

for γ_i defined as in Eq. (3.2). This follows from

$$\mathbb{E}[v_L(Y_i - \delta_i; \gamma_i)] = -\mathbb{E}[(\delta_i - Y_i)_+] + \gamma_i \mathbb{E}[(Y_i - \delta_i)_+)],$$
$$\mathbb{E}[v_U(X_i - \delta_i; \gamma_i)] = -\gamma_i \mathbb{E}[(\delta_i - X_i)_+] + \mathbb{E}[(X_i - \delta_i)_+],$$

and the definition of γ_i.

Therefore, from inequality (3.3) it follows that

$$\mathbb{E}[u(\boldsymbol{Y})] \geq \sum_{i=1}^{N} b_i \mathbb{E}[v_L(Y_i - \delta_i; \gamma_i)] = \sum_{i=1}^{N} b_i \mathbb{E}[v_U(X_i - \delta_i; \gamma_i)] \geq \mathbb{E}[u(\boldsymbol{X})]$$

holds for arbitrary δ_i. We want to choose δ_i such that γ_i is as small as possible. As

$$\gamma_i = \frac{\mathbb{E}\big[(\delta_i - Y_i)_+\big] + \mathbb{E}\big[(X_i - \delta_i)_+\big]}{\mathbb{E}\big[(Y_i - \delta_i)_+\big] + \mathbb{E}\big[(\delta_i - X_i)_+\big]} = \frac{\mathbb{E}\big[(\delta_i - Y_i)_+\big] + \mathbb{E}\big[(X_i - \delta_i)_+\big]}{\mu_{Y_i} - \delta_i + \mathbb{E}\big[(\delta_i - Y_i)_+\big] + \delta_i - \mu_{X_i} + \mathbb{E}\big[(X_i - \delta_i)_+\big]},$$

we have to minimize $\mathbb{E}\big[(\delta_i - Y_i)_+\big] + \mathbb{E}\big[(X_i - \delta_i)_+\big]$ with respect to δ_i. The right derivative is

$$\frac{\partial^+}{\partial \delta_i}\big(\mathbb{E}\big[(\delta_i - Y_i)_+\big] + \mathbb{E}\big[(X_i - \delta_i)_+\big]\big) = \mathbb{E}\big[1_{[\delta_i - Y_i \geq 0]}\big] - \mathbb{E}\big[1_{[X_i - \delta_i \geq 0]}\big] = G_i(\delta_i) - 1 + F_i(\delta_i).$$

Therefore, δ_i is minimized for $\delta_i = \inf\{x \colon F_i(x) + G_i(x) \geq 1\}$. □

Proof of Proposition 7. In this case we can solve for δ_i from Theorem 5:

$$\begin{aligned}
F_i(\delta_i) + G_i(\delta_i) = 1 \iff & H\left(\frac{\delta_i - \mu_{X_i}}{\sigma_{X_i}}\right) + H\left(\frac{\delta_i - \mu_{Y_i}}{\sigma_{Y_i}}\right) = 1 \\
\iff & H\left(\frac{\delta_i - \mu_{X_i}}{\sigma_{X_i}}\right) = H\left(\frac{\mu_{Y_i} - \delta_i}{\sigma_{Y_i}}\right) \\
\iff & \frac{\delta_i - \mu_{X_i}}{\sigma_{X_i}} = \frac{\mu_{Y_i} - \delta_i}{\sigma_{Y_i}} \\
\iff & \quad \delta_i = \frac{\mu_{X_i}\sigma_{Y_i} + \mu_{Y_i}\sigma_{X_i}}{\sigma_{X_i} + \sigma_{Y_i}}.
\end{aligned}$$

Hence

$$\gamma_i = \frac{\mathbb{E}\big[(\delta_i - Y_i)_+\big] + \mathbb{E}\big[(X_i - \delta_i)_+\big]}{\mathbb{E}\big[(Y_i - \delta_i)_+\big] + \mathbb{E}\big[(\delta_i - X_i)_+\big]} = \frac{\sigma_{Y_i} \mathbb{E}\big[(\tau_i - Z)_+\big] + \sigma_{X_i} \mathbb{E}\big[(\tau_i - Z)_+\big]}{\sigma_{Y_i} \mathbb{E}\big[(Z - \tau_i)_+\big] + \sigma_{X_i} \mathbb{E}\big[(Z - \tau_i)_+\big]} = \eta(\tau_i). \qquad \square$$

Proof of Theorem 8. The proof is similar to the proof of Theorem 3 in Müller et al. (2021). Fix arbitrary $\boldsymbol{\delta}$, consider $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, and let $b_i = \sup(u_i'(\boldsymbol{x}))$. Without loss of generality assume $u(\boldsymbol{\delta}) = 0$. By Lemma 6,

$$\sum_{i=1}^{N} b_i v_L(x_i - \delta_i; \gamma_i) \leq u(\boldsymbol{x}) \leq \sum_{i=1}^{N} b_i v_U(x_i - \delta_i; \gamma_i).$$

We need to show that, for some δ_i, $\mathbb{E}[v_L(Y_i-\delta_i;\gamma_i)] \geq \mathbb{E}[v_U(X_i-\delta_i;\gamma_i)]$ for $i=1,\ldots,N$. As established in the proof of Theorem 3 in Müller et al. (2021), this holds for

$$\delta_i = \frac{\mu_{X_i}\sigma_{Y_i} + \mu_{Y_i}\sigma_{X_i}}{\sigma_{X_i}+\sigma_{Y_i}}. \qquad \square$$

Proofs of Section 4

Proof of Theorem 20. As in Lemma 6, we get for $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$

$$\sum_{i=1}^{N} \beta_i v_L(x_i-\delta_i;\gamma) \leq u(\boldsymbol{x}) - u(\boldsymbol{\delta}) \leq \sum_{i=1}^{N} \beta_i v_U(x_i-\delta_i;\gamma).$$

Therefore we can derive as in Theorem 5 that a sufficient condition for $\mathbb{E}[u(\boldsymbol{Y})] \geq \mathbb{E}[u(\boldsymbol{X})]$ is

$$\sum_{i=1}^{N} \beta_i \,\mathbb{E}[v_L(Y_i-\delta_i;\gamma)] \geq \sum_{i=1}^{N} \beta_i \,\mathbb{E}[v_U(X_i-\delta_i;\gamma)],$$

which is equivalent to

$$\gamma \geq \frac{\sum_{i=1}^{N} \beta_i \big(\mathbb{E}\big[(X_i-\delta_i)_+\big] + \mathbb{E}\big[(\delta_i-Y_i)_+\big]\big)}{\sum_{i=1}^{N} \beta_i \big(\mathbb{E}\big[(\delta_i-X_i)_+\big] + \mathbb{E}\big[(Y_i-\delta_i)_+\big]\big)}. \qquad \square$$

Proof of Theorem 21. Assume that (4.4) holds. Fix arbitrary $\boldsymbol{\delta}$, consider $u \in \mathcal{U}_{\gamma,\boldsymbol{\beta}}$, and without loss of generality set $u(\boldsymbol{\delta}) = 0$. As in Lemma 6, it follows that

$$\sum_{i=1}^{N} \beta_i v_L(x_i-\delta_i;\gamma) \leq u(\boldsymbol{x}) \leq \sum_{i=1}^{N} \beta_i v_U(x_i-\delta_i;\gamma).$$

It is sufficient to show that for some $\boldsymbol{\delta}$ we have

$$\sum_{i=1}^{N} \beta_i \,\mathbb{E}[v_L(Y_i-\delta_i;\gamma)] \geq \sum_{i=1}^{N} \beta_i \,\mathbb{E}[v_U(X_i-\delta_i;\gamma)]$$

for any $\boldsymbol{X}$ and $\boldsymbol{Y}$ such that (3.1) holds. As in the proof of Theorem 3 in Müller et al. (2021), we get

$$\mathbb{E}[v_L(Y_i-\delta_i;\gamma)] \geq \gamma\left(\mu_{Y_i}-\delta_i\right) - (1-\gamma)\frac{1}{2}\left(\delta_i - \mu_{Y_i} + \sqrt{\sigma_{Y_i}^2 + \left(\mu_{Y_i}-\delta_i\right)^2}\right)$$

and

$$\mathbb{E}[v_U(X_i-\delta_i;\gamma)] \leq \gamma\left(\mu_{X_i}-\delta_i\right) + (1-\gamma)\frac{1}{2}\left(\mu_{X_i} - \delta_i + \sqrt{\sigma_{X_i}^2 + \left(\mu_{X_i}-\delta_i\right)^2}\right).$$

Thus, we need to find some γ such that

$$\sum_{i=1}^{N}\beta_i\left(\gamma\left(\mu_{Y_i}-\delta_i\right)-(1-\gamma)\frac{1}{2}\left(\delta_i-\mu_{Y_i}+\sqrt{\sigma_{Y_i}^2+\left(\mu_{Y_i}-\delta_i\right)^2}\right)\right)$$
$$\geq\sum_{i=1}^{N}\beta_i\left(\gamma\left(\mu_{X_i}-\delta_i\right)+(1-\gamma)\frac{1}{2}\left(\mu_{X_i}-\delta_i+\sqrt{\sigma_{X_i}^2+\left(\mu_{X_i}-\delta_i\right)^2}\right)\right)$$

for some $\boldsymbol{\delta}$. Following Müller et al. (2021, Theorem 3), we choose

$$\delta_i=\frac{\mu_{X_i}\sigma_{Y_i}+\mu_{Y_i}\sigma_{X_i}}{\sigma_{Y_i}+\sigma_{X_i}},$$

so that

$$\frac{\mu_{Y_i}-\delta_i}{\sigma_{Y_i}}=t_i\quad\text{and}\quad\frac{\mu_{X_i}-\delta_i}{\sigma_{X_i}}=-t_i,\quad\text{where}\quad t_i=\frac{\mu_{Y_i}-\mu_{X_i}}{\sigma_{X_i}+\sigma_{Y_i}}.$$

Then the equation for γ becomes

$$\sum_{i=1}^{N}\beta_i\left(\gamma\sigma_{Y_i}t_i-(1-\gamma)\frac{1}{2}\left(-\sigma_{Y_i}t_i+\sigma_{Y_i}\sqrt{1+t_i^2}\right)\right)$$
$$=\sum_{i=1}^{N}\beta_i\left(\gamma\left(-\sigma_{X_i}t_i\right)+(1-\gamma)\frac{1}{2}\left(-\sigma_{X_i}t_i+\sigma_{X_i}\sqrt{1+t_i^2}\right)\right),$$

which is equivalent to

$$\gamma\sum_{i=1}^{N}\beta_it_i\left(\sigma_{Y_i}+\sigma_{X_i}\right)=(1-\gamma)\frac{1}{2}\sum_{i=1}^{N}\beta_i\left(-\sigma_{X_i}t_i-\sigma_{Y_i}t_i+\left(\sigma_{X_i}+\sigma_{Y_i}\right)\sqrt{1+t_i^2}\right).$$

Define

$$\Delta=\sum_{i=1}^{N}\beta_it_i\left(\sigma_{Y_i}+\sigma_{X_i}\right)=\sum_{i=1}^{N}\beta_i(\mu_{Y_i}-\mu_{X_i}).$$

Then

$$\left(\gamma+(1-\gamma)\frac{1}{2}\right)\Delta=(1-\gamma)\frac{1}{2}\sum_{i=1}^{N}\beta_i\left(\sigma_{X_i}+\sigma_{Y_i}\right)\sqrt{1+t_i^2},$$

or equivalently,

$$(1+\gamma)\,\Delta=(1-\gamma)\sum_{i=1}^{N}\beta_i\left(\sigma_{X_i}+\sigma_{Y_i}\right)\sqrt{1+t_i^2}.$$

This yields

$$\gamma=\frac{\sum_{i=1}^{N}\beta_i\left(\sigma_{X_i}+\sigma_{Y_i}\right)\sqrt{1+t_i^2}-\Delta}{\Delta+\sum_{i=1}^{N}\beta_i\left(\sigma_{X_i}+\sigma_{Y_i}\right)\sqrt{1+t_i^2}}.$$

Alternatively, we can express γ as

$$\gamma = \frac{\sum_{i=1}^{N} \beta_i \left(\sqrt{(\sigma_{X_i} + \sigma_{Y_i})^2 + (\mu_{Y_i} - \mu_{X_i})^2} - (\mu_{Y_i} - \mu_{X_i}) \right)}{\sum_{i=1}^{N} \beta_i \left(\sqrt{(\sigma_{X_i} + \sigma_{Y_i})^2 + (\mu_{Y_i} - \mu_{X_i})^2} + (\mu_{Y_i} - \mu_{X_i}) \right)}. \qquad \square$$

Acknowledgments

Marco Scarsini is a member of GNAMPA-INdAM. His work was partially supported by the the INdAM-GNAMPA Project 2020 "Random walks on random games," the Italian MIUR PRIN 2017 Project ALGADIMAR "Algorithms, Games, and Digital Markets," and the COST action GAMENET.

References

Arlotto, A., Gans, N., and Steele, J. M. (2014). Markov decision problems where means bound variances. *Oper. Res.*, 62(4):864–875.

Arvanitis, S., Post, T., and Topaloglou, N. (2021). Stochastic bounds for reference sets in portfolio analysis. *Management Sci.*, forthcoming.

Atkinson, A. B. (1970). On the measurement of inequality. *J. Econom. Theory*, 2(3):244–263.

Baccelli, F. and Makowski, A. M. (1989). Multidimensional stochastic ordering and associated random variables. *Oper. Res.*, 37(3):478–487.

Denuit, M., Eeckhoudt, L., Tsetlin, I., and Winkler, R. L. (2013). Multivariate concave and convex stochastic dominance. In *Risk Measures and Attitudes*, EAA Ser., pages 11–32. Springer, London.

Faure, M. and Gravel, N. (2021). Reducing inequalities among unequals. *Internat. Econom. Rev.*, 62(1):357–404.

Fishburn, P. C. (1980). Stochastic dominance and moments of distributions. *Math. Oper. Res.*, 5(1):94–100.

Huang, R. J., Tzeng, L. Y., and Zhao, L. (2020). Fractional degree stochastic dominance. *Management Sci.*, 66(10):4630–4647.

Leshno, M. and Levy, H. (2002). Preferred by "all" and preferred by "most" decision makers: almost stochastic dominance. *Management Sci.*, 48(8):1074–1085.

Levhari, D., Paroush, J., and Peleg, B. (1975). Efficiency analysis for multivariate distributions. *Rev. Econ. Stud.*, 42(1):87–91.

Levy, H. and Paroush, J. (1974). Toward multivariate efficiency criteria. *J. Econom. Theory*, 7(2):129–142.

Mao, T. and Wang, R. (2020). A critical comparison of three notions of fractional stochastic dominance. Technical report, SSRN.

Markowitz, H. (1952). Portfolio selection. *J. Finance*, 7(1):77–91.

Mosler, K. (2021). Representative endowments and uniform Gini orderings of multi-attribute inequality. Technical report, arXiv 2103.17030.

Mosler, K. C. (1984). Stochastic dominance decision rules when the attributes are utility independent. *Management Sci.*, 30(11):1311–1322.

Muliere, P. and Scarsini, M. (1989). Multivariate decisions with unknown price vector. *Econom. Lett.*, 29(1):13–19.

Müller, A. (1997). Stochastic orders generated by integrals: a unified study. *Adv. in Appl. Probab.*, 29(2):414–428.

Müller, A. (2013). Duality theory and transfers for stochastic order relations. In *Stochastic Orders in Reliability and Risk*, volume 208 of *Lect. Notes Stat.*, pages 41–57. Springer, New York.

Müller, A. and Scarsini, M. (2012). Fear of loss, inframodularity, and transfers. *J. Econom. Theory*, 147(4):1490–1500.

Müller, A., Scarsini, M., Tsetlin, I., and Winkler, R. L. (2017). Between first and second-order stochastic dominance. *Management Sci.*, 63(9):2933–2947.

Müller, A., Scarsini, M., Tsetlin, I., and Winkler, R. L. (2021). Ranking distributions when only means and variances are known. *Oper. Res.*, forthcoming.

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

O'Brien, G. L. (1984). Stochastic dominance and moment inequalities. *Math. Oper. Res.*, 9(3):475–477.

O'Brien, G. L. and Scarsini, M. (1991). Multivariate stochastic dominance and moments. *Math. Oper. Res.*, 16(2):382–389.

Østerdal, L. P. (2010). The mass transfer approach to multivariate discrete first order stochastic dominance: direct proof and implications. *J. Math. Econom.*, 46(6):1222–1228.

Rothschild, M. and Stiglitz, J. E. (1970). Increasing risk. I. A definition. *J. Econom. Theory*, 2:225–243.

Scarsini, M. (1988). Dominance conditions for multivariate utility functions. *Management Sci.*, 34(4):454–460.

Shadwick, W. F. and Keating, C. (2002). A universal performance measure. *J. Perform. Measurement*, 6(3):59–84.

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

Tsetlin, I. and Winkler, R. L. (2018). Multivariate almost stochastic dominance. *J. Risk Insurance*, 85(2):431–445.