INSEAD The Business School for the World®

Working Paper

2022/52/DSC

(Revised version of 2022/12/DSC)

Sufficient Conditions for Multivariate Almost Stochastic Dominance under Dependence Uncertainty

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

November 7, 2022

Most often important decisions involve several unknown attributes. This produces a double challenge in the sense that both assessing the individual multiattribute preferences and the joint distribution of the attributes can be extremely hard. In this respect, it would be useful to have sufficient conditions for the dominance of one random vector over another under dependence uncertainty, when only partial information on the joint distributions of the two vectors is available, for instance, when only the marginals or just the first two marginal moments are known. Sufficient conditions for multivariate stochastic dominance under dependence uncertainty can be obtained only in very special cases. In this paper we develop sufficient conditions for multivariate almost stochastic dominance based on marginal distributions of the attributes or just on their means and variances. To make use of multivariate almost stochastic dominance, preferences are elicited either in terms of bounds on marginal utilities or via transfers. We apply the theoretical results to comparing the efficiency of photovoltaic plants.

Subject classifications: Decision analysis: criteria, theory; Probability: distribution comparisons.
Keywords: Multivariate Almost Stochastic Dominance; Transfers; Sufficient Conditions for Dominance; Dependence Uncertainty; Mean and Variance.

Sufficient Conditions for Multivariate Almost Stochastic Dominance under Dependence Uncertainty

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

November 7, 2022

Abstract

Most often important decisions involve several unknown attributes. This produces a double challenge in the sense that both assessing the individual multiattribute preferences and the joint distribution of the attributes can be extremely hard. In this respect, it would be useful to have sufficient conditions for the dominance of one random vector over another under dependence uncertainty, when only partial information on the joint distributions of the two vectors is available, for instance, when only the marginals or just the first two marginal moments are known. Sufficient conditions for multivariate stochastic dominance under dependence uncertainty can be obtained only in very special cases. In this paper we develop sufficient conditions for multivariate almost stochastic dominance based on marginal distributions of the attributes or just on their means and variances. To make use of multivariate almost stochastic dominance, preferences are elicited either in terms of bounds on marginal utilities or via transfers. We apply the theoretical results to comparing the efficiency of photovoltaic plants.

Subject classifications: Decision analysis: criteria, theory. Probability: distribution comparisons.
Keywords: multivariate almost stochastic dominance, transfers, sufficient conditions for dominance, dependence uncertainty, mean and variance.

1 Introduction

1.1 The problem

Often a decision maker faces decisions that depend on multiple uncertain attributes. It is typically possible to assess some knowledge of the marginal distributions, whereas one cannot expect to get full knowledge of their joint distribution. In these conditions it is difficult to compare uncertain prospects unless the agent's preferences only depend on the marginal distributions. Moreover, even the exact elicitation of preferences over fully known multivariate distributions is not an easy task. We consider a comparison of multivariate prospects where both the distributions of the prospects and the decision maker's preferences are only partially specified.

Comparison of uncertain prospects under partially specified preferences is the domain of stochastic dominance. For instance, first-degree stochastic dominance (FSD) studies conditions under which a prospect $\boldsymbol{Y}$ is preferred to a prospect $\boldsymbol{X}$ for any decision maker who prefers more to less. This holds if and only if $\mathbb{E}[u(\boldsymbol{X})] \leq \mathbb{E}[u(\boldsymbol{Y})]$ for all weakly increasing utility functions u. Although easy-to-check necessary and sufficient conditions for FSD exist in the univariate case (they amount to a pointwise comparison of distribution functions), the situation is much more complex in the multivariate case. In this respect, it would be helpful to have sufficient dominance conditions that are easy to verify without full knowledge of the multivariate distribution. Such conditions are intrinsically hard to obtain for multivariate FSD. Among other things, this is due to the fact that FSD itself rarely holds for two distributions. For instance, two multivariate normals can be ordered according to FSD only if they have the same covariance matrices.

One possible way out of this conundrum is to restrict the class of admissible preference relations, by imposing some bounds on marginal utilities. This is multivariate almost stochastic dominance (MASD) (Tsetlin and Winkler, 2018), a generalization of the univariate almost stochastic dominance (ASD) defined and studied by Leshno and Levy (2002).

1.2 Our contribution

We consider the problem of comparing two N-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ when only partial information is available concerning the decision maker's preferences and the information about the distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is limited to the marginal distributions of the components of these two vectors. In other words, we consider a stochastic dominance problem under dependence uncertainty, and we provide sufficient conditions for dominance under various information settings concerning the marginal distributions.

Sufficient conditions for ASD based on the first two moments are known in the univariate setting. They are based on bounds of the ratio between the infimum and the supremum of the marginal utility. We consider two possible generalizations of the univariate results. In the first generalization, the decision maker's preferences concern only the various components of the random prospects, separately taken. In the second generalization, the interaction of the various components is also taken into account.

In the first case, we consider stochastic dominance generated by a class $\mathcal{U}_{\boldsymbol{\gamma}}$ of utility functions u parametrized by a vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N) \in [0,1]^N$ where γ_i describes how much the marginal utility of prospect i can vary. More precisely, γ_i is the lower bound of the ratio between the infimum and the supremum of the partial derivative u_i'. When $\boldsymbol{\gamma} = (0, \ldots, 0)$ the class includes all increasing functions and thus reduces to FSD. When $\boldsymbol{\gamma} = (1, \ldots, 1)$ the class includes only affine functions and thus the preference relation reduces to comparing the means of all uncertain prospects. Preferences that are represented by utility functions in this class can also be characterized by preferences for a certain class of probability transfers that depend on $\boldsymbol{\gamma}$. These transfers involve moving probability masses only along specific directions parallel to the main axes, where we allow a probability of a loss in a certain prospect as long as it can be overcompensated by a probability of a gain in the same prospect.

When considering dominance conditions for the class $\mathcal{U}_{\boldsymbol{\gamma}}$ (called $\boldsymbol{\gamma}$-dominance), by varying $\boldsymbol{\gamma} \in [0,1]^N$ we interpolate classical FSD and preferences that are based only on mean vectors. Thus we get a continuum of dominance rules that become weaker and weaker as $\boldsymbol{\gamma}$ increases. Therefore, it is clear that—if all means of the prospects are strictly comparable—there exists a $\boldsymbol{\gamma}$ for which dominance holds.

In this case we provide sufficient conditions for $\boldsymbol{\gamma}$-dominance that only depend on the marginal distributions of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$. These conditions are given for the case where the marginal distributions are completely known and for the case where only their first two moments are known. The rationale for these conditions is to find a sure prospect $\boldsymbol{\delta}$ that $\boldsymbol{\gamma}$-dominates $\boldsymbol{X}$ and is $\boldsymbol{\gamma}$-dominated by $\boldsymbol{Y}$. Clearly the conditions become less strict as $\boldsymbol{\gamma} \to (1,\dots,1)$. When $\gamma = (0,\dots,0)$, i.e., in the case of FSD, the conditions are satisfied only if the support of $\boldsymbol{Y}$ is strictly above the support of $\boldsymbol{X}$.

In many situations even the condition that all means are ordered is not fulfilled, in particular in a high dimensional setting, but it may still be natural to speak of a version of ASD when most of the means are ordered and significantly different. Therefore, we consider a second case parametrized by a scalar $\gamma \in [0,1]$ and a vector $\boldsymbol{\beta} = (\beta_1,\dots,\beta_N) \in \mathbb{R}_+^N$ with the following property: if $\gamma = 0$, then this dominance relation coincides with FSD and if $\gamma = 1$ then the corresponding stochastic dominance rule is the complete ordering obtained by comparing the means of $\sum_i \beta_i X_i$ and $\sum_i \beta_i Y_i$. In general the β_i can be interpreted as attributes' weights. In many circumstances, the β_i correspond to prices with a bid-ask spread such that selling and buying prices are β_i and $\gamma\beta_i$, respectively.

In this case the bounds on the marginal utilities are not imposed separately on each coordinate, but rather in an interconnected way. So the bound on u_i' depends on u_j', for all possible i,j. This translates into a transfer characterization where moving a probability mass in one direction can be suitably compensated by moving some mass in another direction, often with the already mentioned interpretation that we can sell one prospect and buy another one with the money we receive.

As we have a complete ordering now for $\gamma = 1$, there exists a smallest γ for which stochastic dominance holds. Now the sufficient conditions for dominance are expressed in terms of a scalar whose expression involve linear combinations of suitable expectations with weights β_i.

As an illustrative example for the second variant, the results are applied to a comparison of photovoltaic power systems, where the random variables represent the amount of electricity that can be produced at different times of the day in two photovoltaic plants located in two different locations with different weather conditions and different sun angles.

1.3 Related literature

Stochastic dominance (SD) deals with conditions under which a random prospect is preferred to another by all decision makers whose utility function satisfies some properties. We refer the reader to the books by Müller and Stoyan (2002) and Shaked and Shanthikumar (2007) for an extensive analysis of stochastic orders and their properties. Studies of multivariate stochastic dominance (MSD) in an economic decision context include, for instance, Levy and Paroush (1974), Levhari et al. (1975), Mosler (1984), Scarsini (1988), and Baccelli and Makowski (1989).

As conditions for FSD are often very restrictive, Leshno and Levy (2002) developed the concept of univariate ASD where one random prospect is preferred to another by most rational decision makers, rather than by all of them. This is achieved by neglecting agents whose utility functions are in a sense extreme. This idea was carried over to the multivariate setting by Tsetlin and Winkler (2018).

In the univariate case, there are also various attempts to interpolate FSD and second-degree stochastic dominance (SSD) (among them Müller et al., 2017, Huang et al., 2020, Mao and Wang, 2020). The interpolation that we consider here has a different scope and does not refer in any form to risk aversion.

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.

There is a vast literature dealing with problems of uncertainty where only partial knowledge of distributions is known. The partial knowledge of distributions typically involves knowledge of moments and in the multivariate case uncertainty about the dependence between the components, which we call dependence uncertainty. The literature on dependence uncertainty has flourished in the recent years, in particular in the context of bounds for risks in a financial context. Most papers in this stream of literature consider the Fréchet class of multivariate distributions with fixed univariate marginals and either optimize some function over this class, or find lower and upper bounds for some functionals of random vectors whose distribution is in the class. Concerning optimization problems, an early reference is Meilijson and Nádas (1979), who, motivated by critical path analysis on PERT networks, studied optimization problems over the class of multivariate distributions with fixed marginals. Different optimization problems involving dependence uncertainty were studied by Natarajan and coauthors (see, e.g., Natarajan et al., 2009, Doan and Natarajan, 2012, Mishra et al., 2014, Doan et al., 2015, Chen et al., 2022, Natarajan, 2022).

Having in mind financial applications, Puccetti and Rüschendorf (2012) provided an algorithm to compute sharp bounds on the distribution of a function of a random vector in a fixed Fréchet class. Similar problems were considered in Bernard et al. (2014), Embrechts et al. (2015) and Wang et al. (2019). Bernard and Müller (2020) provided dependence uncertainty bounds for the energy score and the Gini mean difference, which play an important role in the context of probabilistic forecasting. Bartl et al. (2022) provided optimal transport duality results for classes of multivariate distribution functions when additional information on the joint distribution is assumed and uncertainty in the marginals is possible. Ghossoub et al. (2022) found bounds for risk measures that can be expressed as nonlinear functions of two factors whose marginal distributions (but not their joint distribution) are known. In the framework of portfolio analysis, Arvanitis et al. (2021) studied stochastic bounding of a portfolio by another. They looked 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 looked for approximate bounds, in the spirit of ASD.

In our work we also consider uncertainty in terms of partial knowledge of moments. Important references for distributionally robust optimization with partial knowledge of moments are, for instance, Delage and Ye (2010) and Wiesemann et al. (2014). Zhu and Fukushima (2009) considered robust portfolio management under various assumptions of uncertainty. Bernard et al. (2017, 2018) combined partial knowledge of moments with dependence uncertainty in a similar context (see also Li et al., 2018).

Once more we emphasize that the existing literature does not deal with sufficient conditions for MASD. Our paper fills this relevant gap and provides interesting characterizations that are well-known for other classes of dominance, such as necessary and sufficient conditions for dominance expressed in terms of transfers. Here we provide such a characterization both when preferences include and do not include substitution among different components of the compared random vectors.

1.4 Organization of the paper

Section 2 defines classes of utility functions by bounding the possible changes of marginal utilities, studies dominance condition based on these classes, and characterizes them in terms of transfers. Section 3 develops sufficient dominance conditions when the full marginal distributions are given and when only their means and variances are known. Section 4 considers the second variant of a parametric class of utility functions, which yields an interpolation between FSD and a complete order; as before, it characterizes dominance based on this class in terms of transfers and provides sufficient conditions for dominance. Section 5 presents a case study concerning a decision on investing in photovoltaic power systems and analyzes real data using the concepts developed in the paper. Concluding comments are given in Section 6. All proofs can be found in Appendix A.

2 Multivariate almost stochastic dominance

In the univariate case ($N = 1$), there exist several concepts of almost stochastic dominance that generalize FSD. One of them is almost first-degree stochastic dominance (AFSD), defined in terms of bounds on marginal utilities (Leshno and Levy, 2002). We generalize this idea to higher dimensions by introducing classes of multivariate utility functions with suitable properties. In this section we consider the most natural generalization of AFSD, which is defined in terms of bounds on marginal utilities for each attribute separately. In Section 4 we will consider a different generalization that takes into account also the interplay between different attributes.

Consider a decision maker who assesses N attributes $\boldsymbol{x} = (x_1, \dots, x_N)$ with a differentiable utility function $u\colon \mathbb{R}^N \to \mathbb{R}$, where u_i' denotes its partial derivative with respect to its i-th argument:

$$u_i'(\boldsymbol{x}) \coloneqq \frac{\partial u(\boldsymbol{x})}{\partial x_i}. \tag{2.1}$$

We now define $\boldsymbol{\gamma}$-*multivariate almost stochastic dominance* ($\boldsymbol{\gamma}$-MASD) for N-variate random vectors. Given any class $\mathcal{U}$ of utility functions, we can define a stochastic dominance relation $\boldsymbol{X} \leq_{\mathcal{U}} \boldsymbol{Y}$ as

$$\mathbb{E}[u(\boldsymbol{X})] \leq \mathbb{E}[u(\boldsymbol{Y})] \quad \text{for all } u \in \mathcal{U}. \tag{2.2}$$

Definition 2.1. For a given vector $\boldsymbol{\gamma} \coloneqq (\gamma_1, \dots, \gamma_N) \in [0,1]^N$, the symbol $\mathcal{U}_{\boldsymbol{\gamma}}$ denotes the set of utility functions u such that, for all $i \in \{1, \dots, N\}$, we have

$$0 \leq \gamma_i u_i'(\boldsymbol{y}) \leq u_i'(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N. \tag{2.3}$$

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 write $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ instead of $\boldsymbol{X} \leq_{\mathcal{U}_{\boldsymbol{\gamma}}} \boldsymbol{Y}$.

Notice that for $\gamma_i > 0$ the condition (2.3) can only hold if the marginal utilities are bounded and then inequality (2.3) is equivalent to

$$\frac{\inf u_i'(\boldsymbol{x})}{\sup u_i'(\boldsymbol{x})} \geq \gamma_i. \tag{2.4}$$

The condition in Eq. (2.4) eliminates utility functions that are too extreme, in the sense that they are neither too flat nor too steep. In particular we assume that marginal utilities are bounded away from zero as well as from infinity. Definition 2.1 corresponds to MASD, as defined by Tsetlin and Winkler (2018), with $\gamma_i = \varepsilon_i/(1-\varepsilon_i)$ for all $i \in \{1, \dots, N\}$. In the univariate case ($N = 1$), it corresponds to 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 $\boldsymbol{X} \leq_{\boldsymbol{\lambda}} \boldsymbol{Y}$ and for $\boldsymbol{\gamma} = \boldsymbol{0}$ we get classical FSD. For the other extreme $\boldsymbol{\gamma} = (1, \dots, 1)$ we get $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ if and only if $\mathbb{E}[X_i] \leq \mathbb{E}[Y_i]$ for all $i = 1, \dots, N$. For all other $\boldsymbol{\gamma}$ we get a stochastic dominance rule that interpolates between these two extreme cases.

We have defined a SD rule by a set of utility functions with bounded marginal utilities. As in the case of FSD, the corresponding preferences can also be characterized via transfers, which may 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 by Rothschild and Stiglitz (1970), who showed that increasing risk can be decomposed into a sequence of mean-preserving spreads. The name *transfer* for operations such as 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). 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) showed 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 concepts of 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 overcompensated by corresponding increasing transfers. See, e.g., Müller et al. (2017) for the univariate case or Müller and Scarsini (2012) for the multivariate case of inframodular transfers. Other related concepts of transfers have been considered in Elton and Hill (1992) and Kamihigashi and Stachurski (2020). A general theory of transfers has been developed in Müller (2013). We now show that such a characterization also holds for the multivariate versions of SD considered in this paper, allowing for some decreasing transfer in some attribute i, as long as this is overcompensated by a corresponding increasing transfer in exactly the same attribute.

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 unit vector of the canonical basis.

Definition 2.2. 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

$$\boldsymbol{x}_2 = \boldsymbol{x}_1 + h\boldsymbol{e}_i, \quad \eta_2(\boldsymbol{x}_4 - \boldsymbol{x}_3) = \gamma_i \eta_1 (\boldsymbol{x}_2 - \boldsymbol{x}_1), \tag{2.5}$$

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 $i \in \{1, \dots, N\}$.

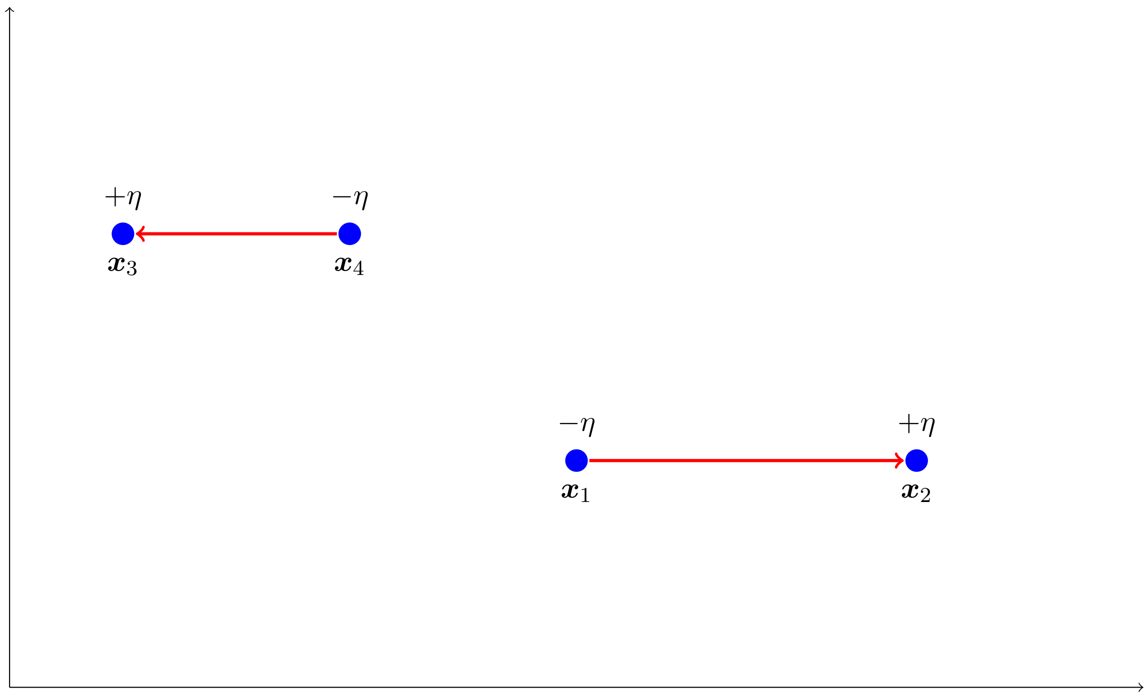

Figure 1: Example of a $\boldsymbol{\gamma}$-transfer with $\gamma_1 = 2/3$, $\eta_1 = \eta_2 = \eta$.

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. Notice, however, that this transfer not only changes a marginal, but also has an effect on the dependence between the components of the random vector.

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

Theorem 2.3. *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 2.3 shows that preferences consistent with $\boldsymbol{\gamma}$-MASD can be thought of as preferences for multivariate $\boldsymbol{\gamma}$-transfers.

One may suspect that Theorem 2.3 means that $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ is equivalent to $X_i \leq_{\gamma_i} Y_i$ for each $i \in \{1, \dots N\}$. The following counterexample shows that this is not the case.

Example 2.4. 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 marginal utilities of this function u are bounded between 1 and 2, so we have $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, but

$$\mathbb{E}[u(\boldsymbol{X})]=5>4=\mathbb{E}\big[u(\boldsymbol{Y}')\big].$$

This shows that the ordering $\leq_{\boldsymbol{\gamma}}$ in general depends not only on the marginal distributions, but on the whole joint distributions of the random vectors. This was to be expected as it is well known that an intricate interplay between marginal dominance and dependence is present already in the classical case of multivariate FSD.

Eliciting a multiattribute utility function is notoriously difficult. However, a decision maker might feel comfortable answering this question: "For any fair lottery (say, a coin flip), would increasing attribute i by one unit if the outcome is heads and reducing this attribute by $t<1$ units if it is tails improve this lottery for you or make it worse for you?" This question can be asked for different values of t. A typical strategy for doing that in decision analysis is to ask the question for a very low value of t (expecting the decision maker will prefer the lottery) and for a high value of t (expecting that the lottery will not be preferred). Then values of t higher than the low value and values lower than the high value can be used to narrow in on an indifference point. This should provide a reasonable estimate of the indifference point, which is the bound γ_i for MASD. The decision maker's preference for γ_i-transfers is, by Theorem 2.3, consistent with $\boldsymbol{\gamma}$-MASD. Note that preferences that are consistent with $\boldsymbol{\gamma}$-transfers extend, beyond the framework of expected utility, to settings with dependent background risk (Section 3.4) and with payoffs that are expressed as suprema of expected utilities (Section 5).

3 Sufficient dominance conditions

In this section we consider sufficient conditions for $\boldsymbol{\gamma}$-dominance. Most of the existing literature on stochastic dominance deals with necessary conditions.

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

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

3.1 Conditions for $\boldsymbol{\gamma}$-dominance when the marginals are known

To better understand the logic that drives our sufficient conditions for $\boldsymbol{\gamma}$-dominance, we start considering the particular case where one of the random vectors is degenerate. In this case the dominance conditions do not depend on the joint distribution of the other random vector, but only on its marginals. Therefore the dependence structure has no role. Moreover, in this special case we can obtain explicit necessary and sufficient dominance conditions.

Proposition 3.1. *Assume that the marginal distributions of the components of $\boldsymbol{X}$ are known and that $\boldsymbol{c}$ is a sure payoff vector.*

(a) *Let $c_i \le \mu_{X_i}$ for all $i = 1, \dots, N$. Then $\boldsymbol{c} \le_{\boldsymbol{\gamma}} \boldsymbol{X}$ if and only if*

$$\gamma_i \ge \frac{\mathbb{E}[(c_i - X_i)_+]}{\mathbb{E}[(X_i - c_i)_+]}, \quad i = 1, \dots, N. \tag{3.2}$$

(b) *Let $\mu_{X_i} \le c_i$ for all $i = 1, \dots, N$. Then $\boldsymbol{X} \le_{\boldsymbol{\gamma}} \boldsymbol{c}$ if and only if*

$$\gamma_i \ge \frac{\mathbb{E}[(X_i - c_i)_+]}{\mathbb{E}[(c_i - X_i)_+]}, \quad i = 1, \dots, N, \tag{3.3}$$

We now provide a sufficient condition for $\boldsymbol{X} \le_{\boldsymbol{\gamma}} \boldsymbol{Y}$ for general $\boldsymbol{X}$ and $\boldsymbol{Y}$ that only uses the marginal distributions and holds for any dependence structures. The basic idea is to find a constant vector $\boldsymbol{\delta}$ that $\boldsymbol{\gamma}$-dominates $\boldsymbol{X}$ and is $\boldsymbol{\gamma}$-dominated by $\boldsymbol{Y}$. According to Proposition 3.1 we can use any $\boldsymbol{\delta}$ between the means of $\boldsymbol{X}$ and $\boldsymbol{Y}$ if these are ordered. Among all these constant vectors $\boldsymbol{\delta}$, we choose the one that produces the smallest $\boldsymbol{\gamma}$. Using this idea, we can derive the following result.

Theorem 3.2. *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 $i = 1, \dots, N$. Let $\delta_i \coloneqq \inf\{x \colon F_i(x) + G_i(x) \ge 1\}$ and let*

$$\gamma_i \coloneqq \frac{\mathbb{E}[(\delta_i - Y_i)_+] + \mathbb{E}[(X_i - \delta_i)_+]}{\mathbb{E}[(Y_i - \delta_i)_+] + \mathbb{E}[(\delta_i - X_i)_+]}, \tag{3.4}$$

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

Notice that δ_i is just a median of the mixture distribution $\frac{1}{2}F_i + \frac{1}{2}G_i$. Also note that if the supports of Y_i and X_i don't overlap, then the corresponding γ_i equals zero.

Remark 3.3. The condition of Theorem 3.2 is obviously not necessary for dominance. If we consider the distributions of Example 2.4, we can see that there does not exist a sure vector $\boldsymbol{\delta}$ that $(1/2, 1/2)$-dominates $\boldsymbol{X}$ and is $(1/2, 1/2)$-dominated by $\boldsymbol{Y}$. The same holds for $\boldsymbol{X}'$ and $\boldsymbol{Y}'$. Nevertheless, we have

$$\boldsymbol{X} \leq_{(1/2,1/2)} \boldsymbol{Y} \quad \text{and} \quad \boldsymbol{X}' \leq_{(1/2,1/2)} \boldsymbol{Y}'. \tag{3.5}$$

If the sufficient condition of Theorem 3.2 were satisfied, we would also have $\boldsymbol{X} \leq_{(1/2,1/2)} \boldsymbol{Y}'$, which actually does not hold.

3.2 Marginal location-scale families

If the marginal distributions of the vectors that we want to compare have nice properties, then the bounds in Theorem 3.2 become easier to compute. In particular, if the marginal distributions are symmetric and belong to a location-scale family, such as normal or uniform, then we can derive easy explicit formulas for the sufficient bounds in Theorem 3.2, as shown in Proposition 3.4. A univariate distribution function F is said to belong to the symmetric location-scale H-family if

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

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 $\mu + \sigma Z$. A particular case of distributions with marginals in a location-scale family is given by elliptical distributions, such as the multivariate normal, the multivariate t-distribution, etc. (see, e.g., Cambanis et al., 1981).

Proposition 3.4. *Let F_i and G_i belong to the same symmetric location-scale H_i-family and let*

$$\eta_i(t) := \frac{\mathbb{E}[(Z_i - t)_+]}{\mathbb{E}[(t - Z_i)_+]}, \tag{3.6}$$

where Z_i has distribution function H_i. If

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

then, in (3.4), we have $\gamma_i = \eta_i(\tau_i)$.

Remark 3.5. If $\boldsymbol{Y} = \boldsymbol{c}$, then $\tau_i = (\mu_{Y_i} - \mu_{X_i})/\sigma_{X_i}$ and the dominance conditions in Proposition 3.4 are necessary and sufficient.

In the next proposition we deal with sufficient conditions for marginal dominance of Y_i over X_i.

Proposition 3.6. *Let F_i and G_i belong to the same symmetric location-scale H_i-family and let*

$$\gamma_i^{\mathrm{M}} := \eta_i\left(\frac{\mu_{Y_i} - \mu_{X_i}}{|\sigma_{Y_i} - \sigma_{X_i}|}\right), \tag{3.7}$$

where η_i is defined as in Eq. (3.6). Then $X_i \leq_{\gamma_i^{\mathrm{M}}} Y_i$.

It is important to notice that γ_i^{M} in Eq. (3.7) is smaller than γ_i in Proposition 3.4. This larger γ_i is the price to pay to have sufficient conditions when the covariance matrices are possibly different. This is relevant, as standard FSD allows the comparison of multinormal random vectors only when they have the same covariance matrix. If one distribution is degenerate, then γ_i^{M} and γ_i are equal.

3.3 Bounds when only means and variances are known

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. For univariate almost stochastic dominance this problem was considered by Müller et al. (2021).

Define

$$\zeta(t) := \frac{1}{1 + 2t\left(t + \sqrt{t^2+1}\right)}. \tag{3.8}$$

Theorem 3.7. *Let the two random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ have finite means and variances. Moreover, for all $i = 1, \ldots, N$, let $\mu_{X_i} \leq \mu_{Y_i}$ and let*

$$\tau_i = \frac{\mu_{Y_i} - \mu_{X_i}}{\sigma_{X_i} + \sigma_{Y_i}}. \tag{3.9}$$

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

As discussed in Müller et al. (2021), these bounds are not sharp. 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 (Proposition 3.4) and when only their means and variances are known (Theorem 3.7). Notice that for normal distributions FSD only holds if the covariance matrices are exactly the same, see Müller (2001).

Remark 3.8. If $\boldsymbol{Y} = \boldsymbol{c}$, where $\boldsymbol{c}$ is a sure vector, then τ_i in Eq. (3.9) becomes $(\mu_{X_i} - c_i)/\sigma_{X_i}$, which is the Sharpe ratio. Notice that in Eq. (3.3) 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. (3.2) the right hand side is $1/\Omega_{X_i}(c_i)$. The connection between univariate ASD, Omega ratio, and Sharpe ratio is discussed in Müller et al. (2021).

3.4 Dependent background risks

In many situations, a decision about a risky project must be made in the presence of other important uncertainties. Pratt (1988, p. 395) makes this point very nicely: "Most real decision makers, unlike those portrayed in our popular texts and theories, confront several uncertainties simultaneously. They must make decisions about some risks when others have been committed to but not resolved. Even when a decision is to be made about only one risk, the presence of others in the background complicates matters."

Many decisions involve some form of background risk that cannot be eliminated. Imagine a department in a company that is contemplating a choice between risks $\boldsymbol{X}$ or $\boldsymbol{Y}$; the background risk $\boldsymbol{Z}$ would

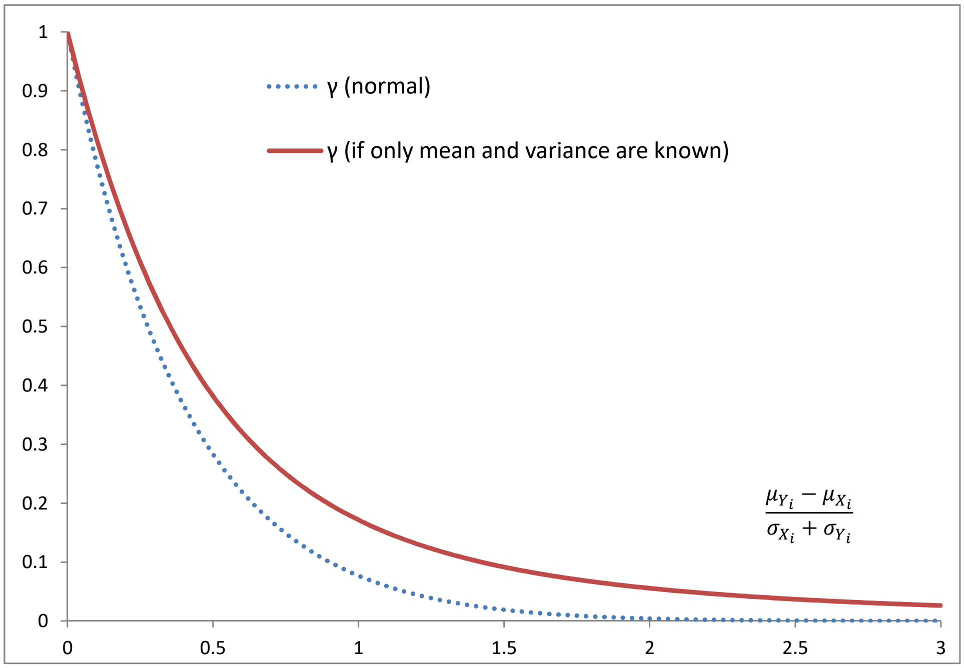

Figure 2: γ_i as a function of $(\mu_{Y_i} - \mu_{X_i})/(\sigma_{X_i} + \sigma_{Y_i})$.

depend on projects undertaken by other departments. For other examples, see Tsetlin and Winkler (2005) and papers cited there. In individual decision making, $\boldsymbol{Z}$ might correspond to health and/or other environmental variables. In many cases, when the random variables that are being compared and the background risk are not stochastically independent, the joint distribution of $(\boldsymbol{X}, \boldsymbol{Z})$ is impossible to estimate, and even the distribution of $\boldsymbol{Z}$ might be hard to assess. Our bounds make it possible to handle such situations.

Theorem 3.9. *Consider $\boldsymbol{X}$ and $\boldsymbol{Y}$ as in Theorem 3.2, and let γ_i be given by Eq. (3.4). Let $\boldsymbol{Z}$ be a K-dimensional multivariate background risk. Let $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_N, 0, \ldots, 0) \in \mathbb{R}_+^{N+K}$. For any $u(\cdot, \cdot) \in \mathcal{U}_{\boldsymbol{\gamma}}$ we have that $\mathbb{E}[u(\boldsymbol{X}, \boldsymbol{Z})] \leq \mathbb{E}[u(\boldsymbol{Y}, \boldsymbol{Z})]$.*

As an illustration, consider univariate X, Y and Z having joint normal distributions with $\mu_Y > \mu_X$. Then

$$\gamma_1^M = \eta\left(\frac{\mu_Y - \mu_X}{|\sigma_X - \sigma_Y|}\right)$$

is given in Proposition 3.6, and dominance with this γ_1^M holds for $X + Z$ and $Y + Z$ if X, Y and Z are independent. In particular, if $\sigma_X = \sigma_Y$, then $Y + Z$ first-order dominates $X + Z$, but this result fails under dependence.

In general, dominance for $X+Z$ and $Y+Z$ holds with

$$\gamma = \eta\left(\frac{\mu_{Y_i} - \mu_{X_i}}{|\sigma_{X+Z} - \sigma_{Y+Z}|}\right),$$

which in turn depends on correlations between X, Z and between Y, Z. Note, however, that

$$|\sigma_{X+Z} - \sigma_{Y+Z}| \leq \sigma_X + \sigma_Y.$$

Therefore $Y+Z$ dominates $X+Z$ for any correlations with γ_1 given by Proposition 3.4. Thus, our sufficient bounds for dominance of Y over X are also sufficient for dominance of $Y+Z$ over $X+Z$ regardless of the dependence.

As mentioned earlier, our sufficient conditions are based on marginal distributions only, which makes them especially easy to implement. They are also useful in settings with background risk (Theorem 3.9), as we can establish dominance of $(\boldsymbol{Y}, \boldsymbol{Z})$ over $(\boldsymbol{X}, \boldsymbol{Z})$ by comparing only marginal distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$.

4 Almost stochastic dominance with substitution

In the univariate case, when $\gamma = 1$, we have $X \leq_\gamma Y \iff \mathbb{E}[X] \leq \mathbb{E}[Y]$. This means that 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. In the previous sections we considered a version of almost stochastic dominance where we got as an extreme case for $\gamma = (1, \ldots, 1)$ an ordering which holds if and only if $\mathbb{E}[X_i] \leq \mathbb{E}[Y_i]$ for all $i = 1, \ldots, N$. This means that we cannot compensate a loss in one prospect by a gain in another prospect, or in other words, there is no possibility of substitution. Given that substitution is often achievable by selling one good and buying another, we now consider a second version of almost stochastic dominance with a parameter $\gamma \in [0,1]$ and a parameter vector $\boldsymbol{\beta}$ such that we get a complete ordering based on the weighted expectations $\mathbb{E}\left[\sum_{i=1}^N \beta_i X_i\right]$ when $\gamma = 1$ and classical first order stochastic dominance when $\gamma = 0$.

4.1 Defining $(\gamma, \boldsymbol{\beta})$-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 4.1. 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

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

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 proportional to the other one.

If $\gamma = 1$, then we get a complete ordering by comparing $\mathbb{E}[\sum \beta_i X_i]$ and $\mathbb{E}[\sum \beta_i Y_i]$ and for $\gamma = 0$ we get classical FSD. Moreover, if $\gamma_i = \gamma$ for $i = 1, \ldots, N$, then $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{Y}$ implies $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$ for any $\boldsymbol{\beta}$.

4.2 Characterization via $(\gamma, \boldsymbol{\beta})$-transfers

We now consider the stochastic dominance rule of Definition 4.1 and define the corresponding transfers. As already mentioned, a loss in one attribute can be compensated by a gain in another attribute. We first discuss why it is sufficient to consider the case of $\beta_i = 1$ for all $i \in \{1, \ldots, N\}$ by changing the units of measurement.

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

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

then the function

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

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

$$\gamma u'_i(\boldsymbol{x}) \leq u'_j(\boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \text{ and for all } i, j. \tag{4.3}$$

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

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

and then

$$\gamma\beta \leq u'_j(\boldsymbol{y}) \leq \beta \quad \text{for all } \boldsymbol{y} \text{ and for all } j;$$

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

$$\gamma\beta_j u'_i(\boldsymbol{x}) \leq \beta_i u'_j(\boldsymbol{y}) \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \text{ and for all } i, j \in \{1, \ldots, N\}.$$

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

Definition 4.2. 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 $i, j \in \{1, \dots, N\}$,

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

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 2.3, we get the following result.

Theorem 4.3. *Let the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$ assume a finite number of values. Then $\boldsymbol{X} \leq_{\gamma, \mathbf{1}} \boldsymbol{Y}$ if and only if the distribution of $\boldsymbol{Y}$ can be obtained from the distribution of $\boldsymbol{X}$ by a finite number of 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

$$\mathbb{E}[\tilde{u}(\beta_1 X_1, \dots, \beta_N X_N)] \leq \mathbb{E}[\tilde{u}(\beta_1 Y_1, \dots, \beta_N Y_N)] \quad \text{for all } \tilde{u} \in \mathcal{U}_{(\gamma, \mathbf{1})}.$$

From this equivalence we get the general $(\gamma, \boldsymbol{\beta})$-transfers as follows. The expressions only become a bit more technical.

Definition 4.4. 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 $i, j \in \{1, \dots, N\}$ 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 $i, j \in \{1, \dots, N\}$,

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

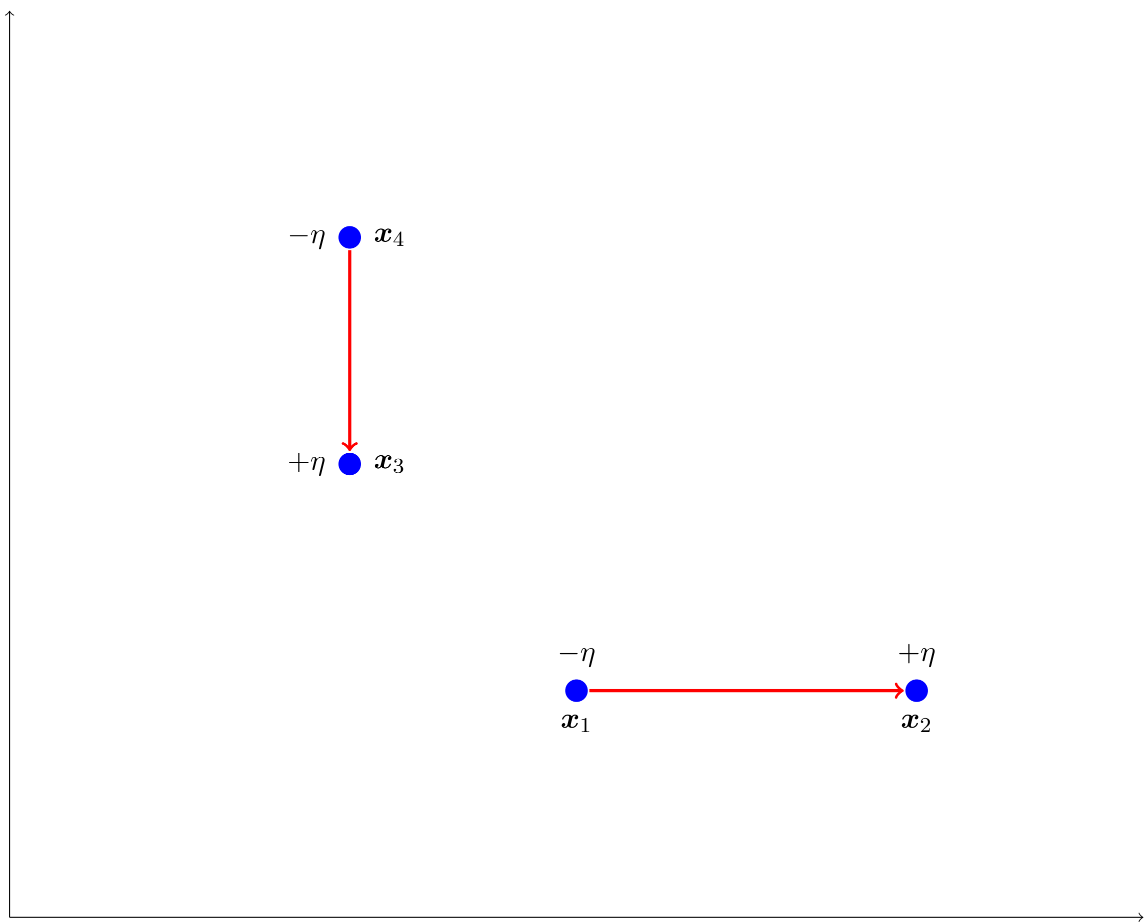

Figure 3: Example of $(\gamma, \mathbf{1})$-transfer with $\varepsilon_1 = 1.5$, $\varepsilon_2 = 1$, $\gamma = 2/3$, $\eta_1 = \eta_2 = \eta$.

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 4.5. *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

Sufficient conditions for this version of MASD are very similar to the conditions described in Section 3.

Theorem 4.6. *Assume that the marginal distributions of the components of* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *are known. Let* $\delta_i := \inf\{x\colon F_i(x) + G_i(x) \geq 1\}$ *and let*

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

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

5 A case study on investments in photovoltaic power systems

A company wants to compare the efficiency of a photovoltaic solar power system in two different locations. The productivity of a photovoltaic (PV) solar power system depends on solar irradiance, which varies through the day and depends on latitude and climate. We want to compare the two possible locations of Rome in Italy and Siegen in Germany, where two of the authors of this paper live. Data for solar irradiance are publicly available for all locations in Europe from the Copernicus Atmosphere Monitoring Service CAMS (2019) http://www.soda-pro.com/web-services/radiation/cams-radiation-service. From this source we downloaded the hourly data for the so-called global horizontal irradiation (GHI) for the year 2020. For each location we get a sample of 365 vectors of daily GHI data. These are displayed for the two cities of Rome and Siegen in Fig. 4 with the hourly means shown in red.

Positive values are only possible between 5 a.m. and 10 p.m. so that only 17 hours of the day are relevant. Therefore we can describe the possible productivity of the PV systems by a random vector $\boldsymbol{X} = (X_1, \ldots, X_{17})$ for Siegen and by a similar random vector $\boldsymbol{Y}$ for Rome, whose distributions we estimate by the empirical distributions of the data. It is not surprising that the values for Rome are typically larger than the ones for Siegen as Rome is more than 1000 km south of Siegen and less rainy. However, we do not have multivariate FSD between the two distributions as not even all hourly means are larger. This can easily be explained by the fact that in the summer days are longer in the north and therefore very early in the morning and late in the evening Siegen has a higher (though small) solar irradiance on average, whereas in the rest of the day Rome has much higher irradiance, as shown

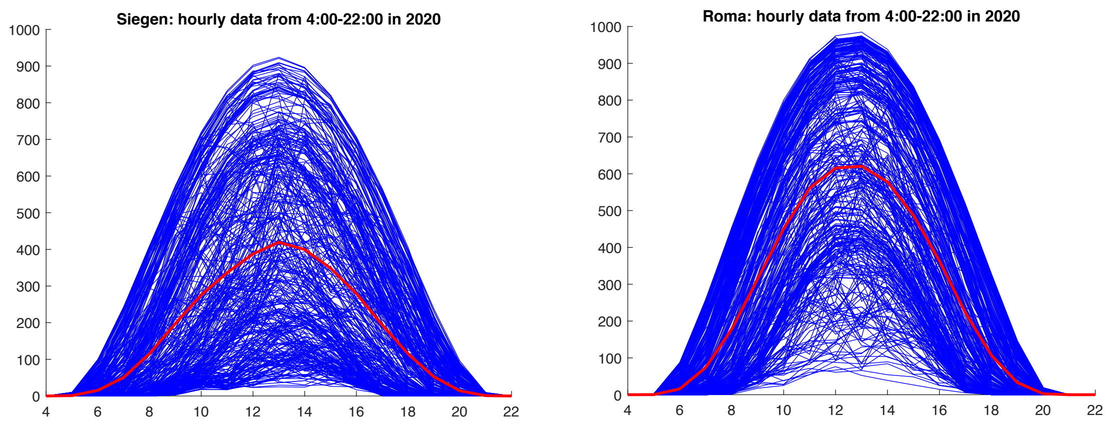

Figure 4: Global horizontal irradiation in Siegen and Rome.

in Fig. 5. It is quite clear, however, that an investment in Rome should be more profitable as there is some kind of almost FSD, as these minor effects in the morning and evening should be more than compensated by the much larger productivity for Rome compared to Siegen in the middle of the day.

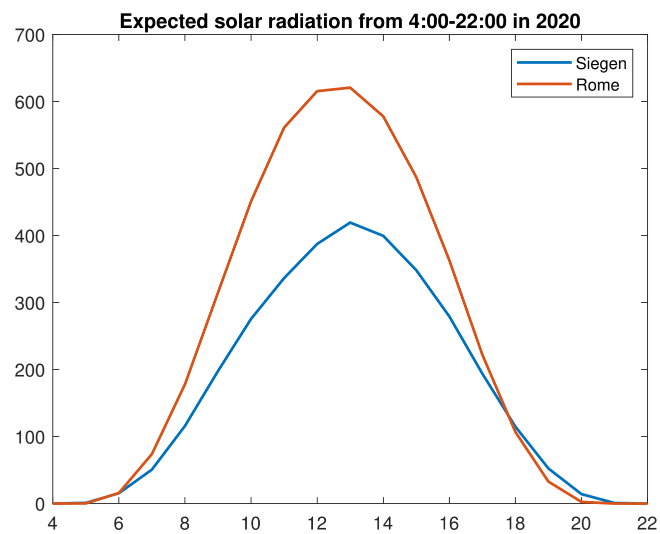

Figure 5: Expected hourly GHI for Siegen and Rome.

We now show that it is reasonable to assume that the expected reward of the decision maker can be written in form of an expected utility $\mathbb{E}[u(\boldsymbol{X})]$ that fulfills the assumptions of Definition 4.1 for appropriate parameters and that we can show with the results of Section 4 that $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$ holds for appropriate parameters. For this illustrative example we make a few simplifying assumptions as the real world problem is very complex. We assume that the output of the PV system is exactly proportional to the GHI. The decision maker is assumed to be a so called prosumer, i.e., simultaenously a producer as well as a consumer of electricity. For simplicity we assume that the consumption can be described by a random vector $\boldsymbol{Z}$ that is independent of the production of the PV system. Notice, however, that with the methods described in this paper, we can also handle situations with dependent background risk (see Theorem 3.9).

Unfortunately, the multivariate distribution of the consumption vector $\boldsymbol{Z}$ of a company is typically difficult to assess (see, e.g., Berk et al. (2018) for an attempt to describe electricity demand patterns of companies by a stochastic model). It is reasonable to assume that it is always possibile to buy electricity for a price of β per unit and sell it for a lower price of $\gamma\beta$ with $0 < \gamma < 1$. In practice, these prices may also vary with the hour of the day and therefore there would be a price vector $\boldsymbol{\beta}$, but we assume here for simplicity that β is constant. Prosumers have the incentive to consume the produced electricity themselves as much as possible to avoid the higher cost of β per unit for buying electricity. Electricity could be sold for the lower price $\gamma\beta$ if production exceeds consumption. If this simple strategy is applied, then, for a given output vector $\boldsymbol{x}$ of the PV system and a consumption vector $\boldsymbol{z}$, the payoff is

$$v(\boldsymbol{x}, \boldsymbol{z}) = \sum_{i=1}^{17} \beta \min\{x_i, z_i\} + \gamma\beta(x_i - z_i)_+. \tag{5.1}$$

For a random consumption vector $\boldsymbol{Z}$, the expected reward given output vector $\boldsymbol{x}$ is

$$u(\boldsymbol{x}) := \mathbb{E}[v(\boldsymbol{x}, \boldsymbol{Z})]. \tag{5.2}$$

In the case of this simple separable utility function the strong positive dependence between the production in different hours is irrelevant. However, strategic behavior of the prosumer may lead to a higher payoff. For instance, battery storage could be employed to store the produced electricity, so that the prosumer could adopt a policy π that allows electricity to be used later instead of being sold for a cheap price. Therefore, the real value that one gets as expected payoff is much more complicated and not separable anymore and the dependence structure of the multivariate distribution of $\boldsymbol{X}$ is relevant. When we have a random electricity consumption and in addition the possibility of using battery storage, we are not able to give a simple explicit expression for the value of the expected payoff of an operating policy π. However, it is still true that the marginal utilities are bounded by $\gamma\beta$ and β, so that we have $u_\pi \in \mathcal{U}_{\gamma,\boldsymbol{\beta}}$. Therefore the decision maker with operating policy π prefers the investment with a production vector $\boldsymbol{Y}$ to the one with production vector $\boldsymbol{X}$ if $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$. If the prosumer solves an optimization problem to find an optimal operating policy among all possible operating policies, the expected value of the PV system has the form $V(\boldsymbol{X}) = \sup_\pi \mathbb{E}[u_\pi(\boldsymbol{X})]$, which may not have the form of an expected utility anymore as the optimal policy may depend on the random vector $\boldsymbol{X}$. Nevertheless, it is still true that $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$ implies $V(\boldsymbol{X}) \leq V(\boldsymbol{Y})$ as the ordering property is preserved by taking a supremum.

It is also very difficult to determine the complicated dependence structure of the random vector of GHI data (see, e.g., Müller and Reuber (2022) for an attempt to describe the whole multivariate distribution of this time series by a stochastic model using time-dependent beta distributions and copula models). Therefore the results of Section 4 are useful to obtain bounds for the parameter γ that ensures $\boldsymbol{X} \leq_{\gamma,\boldsymbol{\beta}} \boldsymbol{Y}$. Approximating the marginal distributions with their empirical counterparts, and ignoring the dependence structure, from Theorem 4.6 we can derive the value $\gamma = 0.525$. Using only means and variances of the marginals, from Theorem 4.7 we get the value $\gamma = 0.576$. The difference between the two values is comparable to the difference for the values of γ_i that one obtains for normally distributed marginals in Section 3, as described in Fig. 5.

Table 1 shows the means, standard deviations, and the corresponding univariate γ for the different time slots. The numbers in boldface refer to the times when Siegen dominates Rome. One can see that the value of $\gamma = 0.576$ obtained for the multivariate version of the bounds discussed in Section 4 is about the same size as the bounds that one gets when considering the univariate problem for a single hour in the most relevant hours in the middle of the day.

Table 1: Expected values, standard deviations and γ's for different time slots

	Expected value		Standard deviation		γ
time	Siegen	Roma	Siegen	Roma	
04:00	0.000	0.000	0.000	0.000	-
05:00	**1.003**	**0.199**	2.393	0.542	**0.582**
06:00	15.401	15.744	26.398	25.754	0.987
07:00	50.707	73.894	70.451	87.724	0.747
08:00	115.811	177.978	128.355	157.311	0.649
09:00	197.805	314.928	185.353	204.742	0.553
10:00	275.557	450.991	224.584	232.883	0.473
11:00	336.451	561.013	249.567	248.803	0.418
12:00	387.412	615.468	262.666	260.380	0.429
13:00	419.324	620.697	266.839	263.684	0.476
14:00	399.560	577.937	257.419	252.466	0.504
15:00	348.290	487.155	235.694	234.401	0.558
16:00	279.321	363.207	209.142	212.732	0.674
17:00	193.808	222.166	171.812	176.166	0.850
18:00	**114.787**	**106.833**	122.830	113.958	**0.935**
19:00	**52.316**	**32.758**	68.294	47.429	**0.714**
20:00	**14.082**	**2.432**	24.152	5.123	**0.460**
21:00	**0.821**	**0.000**	2.020	0.000	**0.453**
22:00	0.000	0.000	0.000	0.000	-

In the context of this example the transfers corresponding to the stochastic dominance rule also have a simple and intuitive interpretation. If the marginal utility of the produced electricity is bounded by the buying price β and the selling price $\gamma\beta$, then we prefer a scenario where we produce more in hour i and less in hour j as long as the lower production in hour j is bounded by a fraction γ of the higher production in hour i.

6 Conclusions

Stochastic dominance (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 stochastic dominance is difficult for two reasons: First, often distributions cannot be ranked (e.g., by FSD); this can be overcome by using relaxations like $\boldsymbol{\gamma}$-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 multivariate distributions are very difficult in general and especially difficult when dependence assessment is concerned.

Another distinction of the multivariate case, compared to the univariate case, is that a real coordinate space is not completely ordered. To attain a version of SD that leads to a complete order in the boundary case of $\gamma = 1$, we need to constrain maximal marginal utilities for different attributes. Section 4 presents the corresponding definition of $(\gamma, \boldsymbol{\beta})$-MASD with substitution, its characterization via transfers, and sufficient conditions for comparing two risky multivariate alternatives.

Within the expected utility framework, $\boldsymbol{\gamma}$-MASD and $(\gamma, \boldsymbol{\beta})$-MASD translate into bounds on marginal utilities (Definitions 2.1 and 4.1). Alternatively, these preferences can be characterized via transfers (Definitions 2.2 and 4.4 and Theorems 2.3 and 4.5). Such transfers might be easier to explain to decision makers and use for elicitation of $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$.

Section 5 illustrates our approach by comparing two possible locations for photovoltaic power plants, where standard FSD comparison fails even though it is intuitively clear that FSD holds almost in some appropriate sense. In that context as well as in many other situations where we can buy and sell a product on a market, the parameter γ bounding the slope of marginal utilities has a very natural interpretation in terms of bid-ask spreads.

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 more and more towards more complex decision making processes 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 Proofs

Proofs of Section 2

The proof of Theorem 2.3 requires the following lemma.

Lemma A.1. *Let $u : \mathbb{R}^N \to \mathbb{R}$ be continuously differentiable. Then $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ if and only if*

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

for all $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ satisfying (2.5) for some i and γ_i.

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

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

so (A.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 (2.3) we get $u \in \mathcal{U}_{\boldsymbol{\gamma}}$.

Only if part: Now assume that $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ is continuously differentiable. Let $\boldsymbol{h} \coloneqq \boldsymbol{x}_2 - \boldsymbol{x}_1$. For $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4$ satisfying (2.5) 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 (A.1) we can deduce

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

□

Proof of Theorem 2.3. The proof is based on the duality theory for transfers. Lemma A.1 shows that $\mathcal{U}_{\boldsymbol{\gamma}}$ can be described by a set of inequalities, as in Müller (2013, Definition 2.2.1). Therefore it is induced by the corresponding set of transfers. The proof thus follows from Müller (2013, Theorem 2.4.1). □

Proofs of Section 3

The following lemma is the building block in the proofs of most of the subsequent results in our paper. The basic idea is that increments of functions $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ can be bounded above and below by separable piecewise linear utility functions that depend on $\boldsymbol{\gamma}$. This fact allows us to find sufficient conditions for $\boldsymbol{\gamma}$-dominance that do not depend on the joint distributions of the random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$, but only on the marginal distributions of their components.

Lemma A.2. *Let*

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

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

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 $\boldsymbol{c} \in \mathbb{R}^N$. Then, for any $\boldsymbol{x} \in \mathbb{R}^N$, we have

$$\sum_{i=1}^{N} b_i v_L(x_i - c_i; \gamma_i) \leq u(\boldsymbol{x}) - u(\boldsymbol{c}) \leq \sum_{i=1}^{N} b_i v_U(x_i - c_i; \gamma_i). \tag{A.2}$$

An instance of functions v_L and v_U is shown in Fig. 6.

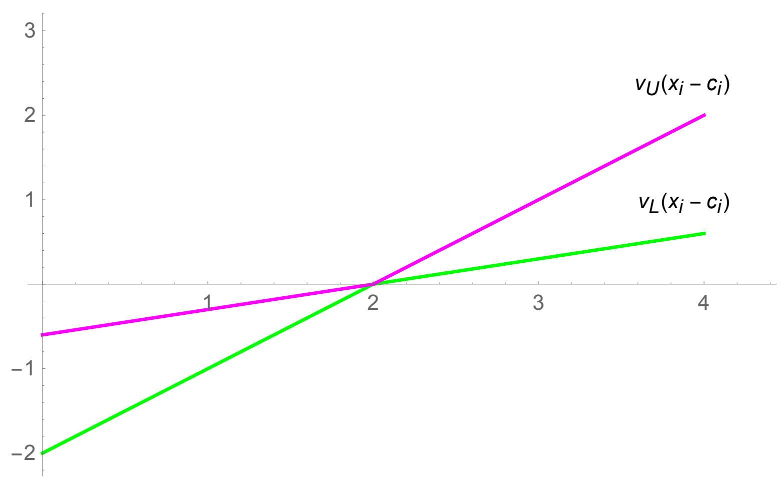

Figure 6: Functions v_L and v_U.

Proof of Lemma A.2. Note that $u_i'(\boldsymbol{x}) \leq \sup(u_i'(\boldsymbol{x})) = b_i$ and that by inequality (2.4) we have $u_i'(\boldsymbol{x}) \geq \gamma_i b_i$. By a multivariate first-order Taylor expansion, $u(\boldsymbol{x}) - u(\boldsymbol{c}) = \sum_{i=1}^{N} u_i'(\boldsymbol{y})(x_i - c_i)$, where y_i is between x_i and c_i. Then, using $u_i'(\boldsymbol{y}) \leq b_i$ if $x_i > c_i$ and $u_i'(\boldsymbol{y}) \geq \gamma_i b_i$ if $x_i < c_i$ provides an upper bound, whereas using $u_i'(\boldsymbol{y}) \geq \gamma_i b_i$ if $x_i > c_i$ and $u_i'(\boldsymbol{y}) \leq b_i$ if $x_i < c_i$ provides a lower bound. $\square$

Proof of Proposition 3.1. We prove (a). The proof of (b) is similar. Let $u \in \mathcal{U}_{\boldsymbol{\gamma}}$ and let

$$b_i \coloneqq \sup_{\boldsymbol{x} \in \mathbb{R}^N} u_i'(\boldsymbol{x}). \tag{A.3}$$

Without any loss of generality, assume $u(\boldsymbol{c}) = 0$. By Lemma A.2 we have

$$u(\boldsymbol{x}) \leq \sum_{i=1}^{N} b_i v_U(x_i - c_i; \gamma_i), \tag{A.4}$$

where $v_U(x_i - c_i; \gamma_i) = -\gamma_i (c_i - x_i)_+ + (x_i - c_i)_+$. This implies

$$\mathbb{E}[u(\boldsymbol{X})] \leq \sum_{i=1}^{N} b_i (-\gamma_i \, \mathbb{E}[(c_i - X_i)_+] + \mathbb{E}[(X_i - c_i)_+]) \tag{A.5}$$

Therefore $\mathbb{E}[u(\boldsymbol{X})] \leq 0$ if $-\gamma_i \, \mathbb{E}[(c_i - X_i)_+] + \mathbb{E}[(X_i - c_i)_+] \leq 0$ for all $i = 1, \ldots, N$.

Notice that $-\gamma_i \, \mathbb{E}[(c_i - X_i)_+] + \mathbb{E}[(X_i - c_i)_+] \leq 0$ is equivalent to $X_i \leq_{\gamma_i} c_i$. This proves the if part.

Now we prove the only if part. Consider a sequence of utility functions

$$u_n(\boldsymbol{x}) = \sum_{i=1}^{N} b_{i,n} v_U(x_i - c_i; \gamma_i)_+ \in \mathcal{U}_{\boldsymbol{\gamma}} \tag{A.6}$$

such that $\lim_{n \to \infty} b_{j,n} = 0$ for $j \neq i$ and $b_{i,n} \equiv 1$ for all n.

If $\boldsymbol{X} \leq_{\boldsymbol{\gamma}} \boldsymbol{c}$, then $\mathbb{E}[u_n(\boldsymbol{X})] \leq u_n(\boldsymbol{c}) = 0$. This implies $-\gamma_i \, \mathbb{E}[(c_i - X_i)_+] + \mathbb{E}[(X_i - c_i)_+] \leq 0$ for all $i = 1, \ldots, N$, i.e., $X_i \leq_{\gamma_i} c_i$, for all $i = 1, \ldots, N$. □

Proof of Theorem 3.2. 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 A.2 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. (3.4). This follows from

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

and the definition of γ_i.

Therefore, from inequality (A.2) 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

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

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

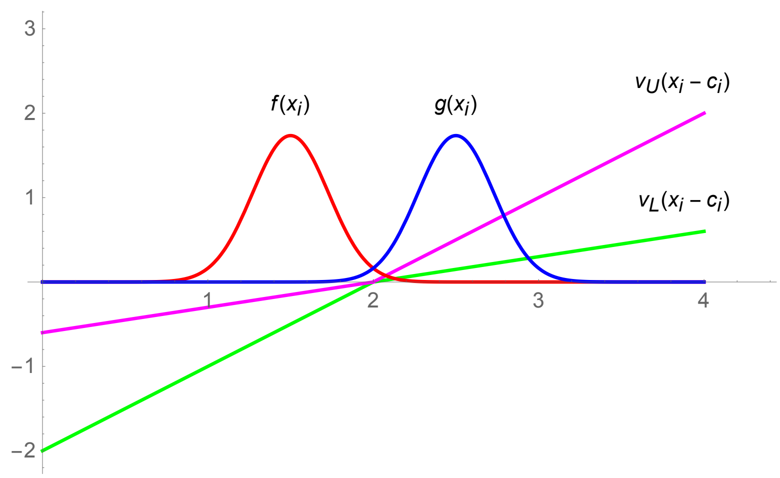

Figure 7: The variable Y_i $\boldsymbol{\gamma}$-dominates the constant c_i, which in turns dominates the variable X_i.

In Fig. 7, for some $\boldsymbol{\gamma}$, the variable Y_i dominates c_i and c_i dominates X_i.

Proof of Proposition 3.4. In this case we can solve for δ_i from Theorem 3.2:

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

Hence

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

The proof of Proposition 3.6 is along the lines of Müller et al. (2017, example 2.11).

Proof of Proposition 3.6. The following condition for γ_i^{M}-dominance in location-scale models can be found in Müller et al. (2017, bottom of page 2940):

$$\gamma_i^{\mathrm{M}} = \frac{\displaystyle\int_{-\infty}^{\infty}(G_i(x)-F_i(x))_+\,\mathrm{d}x}{\displaystyle\int_{-\infty}^{\infty}(F_i(x)-G_i(x))_+\,\mathrm{d}x} = \frac{\displaystyle\int_{-\infty}^{\infty}\left(H\left(\frac{x-\mu_{Y_i}}{\sigma_{Y_i}}\right)-H\left(\frac{x-\mu_{X_i}}{\sigma_{X_i}}\right)\right)_+\mathrm{d}x}{\displaystyle\int_{-\infty}^{\infty}\left(H\left(\frac{x-\mu_{X_i}}{\sigma_{X_i}}\right)-H\left(\frac{x-\mu_{Y_i}}{\sigma_{Y_i}}\right)\right)_+\mathrm{d}x}. \tag{A.7}$$

The two distribution functions F_i and G_i single-cross at a point δ_i such that

$$\frac{\delta_i-\mu_{X_i}}{\sigma_{X_i}} = \frac{\delta_i-\mu_{Y_i}}{\sigma_{Y_i}}, \tag{A.8}$$

which implies

$$\delta_i = \frac{\mu_{Y_i}\sigma_{X_i}-\mu_{X_i}\sigma_{Y_i}}{\sigma_{X_i}-\sigma_{Y_i}}. \tag{A.9}$$

Notice that, for $x<\delta_i$, the distribution with a larger variance takes larger values than the other one. Moreover, integrating by parts, we get the well-known equalities:

$$\int_{\infty}^{\delta_i} F_i(x)\,\mathrm{d}x = \mathbb{E}\big[(\delta_i-X_i)_+\big], \qquad \int_{\delta_i}^{\infty} F_i(x)\,\mathrm{d}x = \mathbb{E}\big[(X_i-\delta_i)_+\big]. \tag{A.10}$$

Therefore, when $\sigma_{Y_i}>\sigma_{X_i}$, Eq. (A.7) becomes

$$\gamma_i^{\mathrm{M}} = \frac{\displaystyle\int_{-\infty}^{\delta_i}\left(H\left(\frac{x-\mu_{Y_i}}{\sigma_{Y_i}}\right)-H\left(\frac{x-\mu_{X_i}}{\sigma_{X_i}}\right)\right)\mathrm{d}x}{\displaystyle\int_{\delta_i}^{\infty}\left(H\left(\frac{x-\mu_{X_i}}{\sigma_{X_i}}\right)-H\left(\frac{x-\mu_{Y_i}}{\sigma_{Y_i}}\right)\right)\mathrm{d}x} = \frac{\mathbb{E}\big[(\delta_i-Y_i)_+\big]-\mathbb{E}\big[(\delta_i-X_i)_+\big]}{\mathbb{E}\big[(Y_i-\delta_i)_+\big]-\mathbb{E}\big[(X_i-\delta_i)_+\big]}. \tag{A.11}$$

Since

$$\mathbb{E}\big[(\delta_i-Y_i)_+\big] = \mathbb{E}\big[(\delta_i-\mu_{Y_i}-\sigma_{Y_i}Z)_+\big] = \sigma_{Y_i}\,\mathbb{E}\left[\left(\frac{\delta_i-\mu_{Y_i}}{\sigma_{Y_i}}-Z\right)_+\right], \tag{A.12}$$

we have

$$
\begin{aligned}
\frac{\delta_i - \mu_{Y_i}}{\sigma_{Y_i}} &= \frac{1}{\sigma_{Y_i}}\left(\frac{\mu_{X_i}\sigma_{Y_i} - \mu_{Y_i}\sigma_{X_i}}{\sigma_{Y_i} - \sigma_{X_i}} - \mu_{Y_i}\right) \\
&= \frac{1}{\sigma_{Y_i}}\left(\frac{\mu_{X_i}\sigma_{Y_i} - \mu_{Y_i}\sigma_{X_i} - \mu_{Y_i}\sigma_{Y_i} + \mu_{Y_i}\sigma_{X_i}}{\sigma_{Y_i} - \sigma_{X_i}}\right) \\
&= \frac{1}{\sigma_{Y_i}}\left(\frac{\mu_{X_i}\sigma_{Y_i} - \mu_{Y_i}\sigma_{Y_i}}{\sigma_{Y_i} - \sigma_{X_i}}\right) \\
&= \frac{\mu_{X_i} - \mu_{Y_i}}{\sigma_{Y_i} - \sigma_{X_i}}.
\end{aligned} \tag{A.13}
$$

This implies that

$$
\mathbb{E}\big[(\delta_i - Y_i)_+\big] = \sigma_{Y_i}\,\mathbb{E}\left[\left(\frac{\mu_{X_i} - \mu_{Y_i}}{\sigma_{Y_i} - \sigma_{X_i}} - Z\right)_+\right]. \tag{A.14}
$$

Applying a similar argument to the other components in Eq. (A.11), we obtain

$$
\gamma_i^{\mathrm{M}} = \frac{\mathbb{E}\left[\left(\frac{\mu_{X_i}-\mu_{Y_i}}{\sigma_{Y_i}-\sigma_{X_i}} - Z\right)_+\right]}{\mathbb{E}\left[Z - \left(\frac{\mu_{X_i}-\mu_{Y_i}}{\sigma_{Y_i}-\sigma_{X_i}}\right)_+\right]}. \tag{A.15}
$$

A similar derivation holds for $\sigma_{Y_i} > \sigma_{X_i}$. □

Proof of Theorem 3.7. The proof uses ideas that are similar to the ones in the proof of Theorem 3 in Müller et al. (2021). Fix arbitrary $\boldsymbol{\delta}$, consider $u \in \mathcal{U}_{\boldsymbol{\gamma}}$, and let $b_i = \sup\left(u_i'(\boldsymbol{x})\right)$. Without loss of generality assume $u\left(\boldsymbol{\delta}\right) = 0$. By Lemma A.2,

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

We need to show that, for some appropriate δ_i and γ_i, $\mathbb{E}[v_L(Y_i - \delta_i; \gamma_i)] \geq \mathbb{E}[v_U(X_i - \delta_i; \gamma_i)]$ for $i = 1, \ldots, N$. With the same tedious but straightforward calculation as in the proof of Theorem 3 in Müller et al. (2021), we can establish that the smallest possible choice for γ_i is obtained by choosing

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

and

$$
\gamma_i = \frac{1}{1 + 2t\left(t + \sqrt{t^2 + 1}\right)}
$$

for

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

□

Proof of Theorem 3.9. The proof is similar to the proof of Theorem 3.2. We get

$$\sum_{i=1}^{N} b_i v_L(x_i - \delta_i; \gamma_i) \le u\left(\boldsymbol{x}, \boldsymbol{z}\right) - u\left(\boldsymbol{\delta}, \boldsymbol{z}\right) \le \sum_{i=1}^{N} b_i v_U(x_i - \delta_i; \gamma_i),$$

and thus

$$\begin{aligned}
\mathbb{E}\left[u\left(\boldsymbol{Y}, \boldsymbol{Z}\right)\right] &\ge \sum_{i=1}^{N} b_i \mathbb{E}\left[v_L(Y_i - \delta_i; \gamma_i)\right] + \mathbb{E}\left[u\left(\boldsymbol{\delta}, \boldsymbol{Z}\right)\right] \\
&= \sum_{i=1}^{N} b_i \mathbb{E}\left[v_U(X_i - \delta_i; \gamma_i)\right] + \mathbb{E}\left[u\left(\boldsymbol{\delta}, \boldsymbol{Z}\right)\right] \\
&\ge \mathbb{E}\left[u\left(\boldsymbol{X}, \boldsymbol{Z}\right)\right].
\end{aligned}$$

□

Proofs of Section 4

Proof of Theorem 4.6. As in Lemma A.2, we get for $\mathcal{U}_{\gamma,\boldsymbol{\beta}}$

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

Therefore, as in Theorem 3.2, a sufficient condition for $\mathbb{E}[u\left(\boldsymbol{Y}\right)] \ge \mathbb{E}[u\left(\boldsymbol{X}\right)]$ is

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

which is equivalent to

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

□

Proof of Theorem 4.7. Assume that (4.4) holds. Fix arbitrary $\boldsymbol{\delta}$, consider $u \in \mathcal{U}_{\gamma,\boldsymbol{\beta}}$, and without loss of generality set $u\left(\boldsymbol{\delta}\right) = 0$. As in Lemma A.2, it follows that

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

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

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

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

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

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

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

which is equivalent to

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

Define

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

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

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

Acknowledgments

We thank Franziska Wollny for help with the tables and figures in Section 5. Marco Scarsini is a member of GNAMPA-INdAM. His work was partially supported by the Italian MIUR PRIN 2017 Project ALGADIMAR "Algorithms, Games, and Digital Markets."

References

Arvanitis, S., Post, T., and Topaloglou, N. (2021). Stochastic bounds for reference sets in portfolio analysis. *Management Sci.*, 67(12):7737–7754.

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.

Bartl, D., Kupper, M., Lux, T., Papapantoleon, A., and Eckstein, S. (2022). Marginal and dependence uncertainty: bounds, optimal transport, and sharpness. *SIAM J. Control Optim.*, 60(1):410–434.

Berk, K., Hoffmann, A., and Müller, A. (2018). Probabilistic forecasting of industrial electricity load with regime switching behavior. *Int. J. Forecast.*, 34(2):147–162.

Bernard, C., Denuit, M., and Vanduffel, S. (2018). Measuring portfolio risk under partial dependence information. *J. Risk Insurance*, 85(3):843–863.

Bernard, C., Jiang, X., and Wang, R. (2014). Risk aggregation with dependence uncertainty. *Insurance Math. Econom.*, 54:93–108.

Bernard, C. and Müller, A. (2020). Dependence uncertainty bounds for the energy score and the multivariate Gini mean difference. *Depend. Model.*, 8(1):239–253.

Bernard, C., Rüschendorf, L., and Vanduffel, S. (2017). Value-at-risk bounds with variance constraints. *J. Risk Insurance*, 84(3):923–959.

Cambanis, S., Huang, S., and Simons, G. (1981). On the theory of elliptically contoured distributions. *J. Multivariate Anal.*, 11(3):368–385.

Chen, L., Ma, W., Natarajan, K., Simchi-Levi, D., and Yan, Z. (2022). Distributionally robust linear and discrete optimization with marginals. *Oper. Res.*, 70(3):1822–1834.

Delage, E. and Ye, Y. (2010). Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Oper. Res.*, 58(3):595–612.

Doan, X. V., Li, X., and Natarajan, K. (2015). Robustness to dependency in portfolio optimization using overlapping marginals. *Oper. Res.*, 63(6):1468–1488.

Doan, X. V. and Natarajan, K. (2012). On the complexity of nonoverlapping multivariate marginal bounds for probabilistic combinatorial optimization problems. *Oper. Res.*, 60(1):138–149.

Elton, J. and Hill, T. P. (1992). Fusions of a probability distribution. *Ann. Probab.*, 20(1):421–454.

Embrechts, P., Wang, B., and Wang, R. (2015). Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance Stoch.*, 19(4):763–790.

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

Ghossoub, M., Hall, J., and Saunders, D. (2022). Maximum spectral measures of risk with given risk factor marginal distributions. *Math. Oper. Res.*, forthcoming.

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

Kamihigashi, T. and Stachurski, J. (2020). Partial stochastic dominance via optimal transport. *Oper. Res. Lett.*, 48(5):584–586.

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.

Li, L., Shao, H., Wang, R., and Yang, J. (2018). Worst-case range value-at-risk with partial information. *SIAM J. Financial Math.*, 9(1):190–218.

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

Meilijson, I. and Nádas, A. (1979). Convex majorization with an application to the length of critical paths. *J. Appl. Probab.*, 16(3):671–677.

Mishra, V. K., Natarajan, K., Padmanabhan, D., Teo, C.-P., and Li, X. (2014). On theoretical and empirical aspects of marginal distribution choice models. *Management Sci.*, 60(6):1511–1531.

Mosler, K. C. (1984). Stochastic dominance decision rules when the attributes are utility independent.

Management Sci., 30(11):1311–1322.

Müller, A. (2001). Stochastic ordering of multivariate normal distributions. *Ann. Inst. Statist. Math.*, 53(3):567–575.

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 Reuber, M. (2022). A copula-based time series model for global horizontal irradiation. *Int. J. Forecast.*, forthcoming.

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.

Natarajan, K. (2022). *Optimization with Marginals and Moments*. Dynamic Ideas LLC, Belmont, Massachusetts.

Natarajan, K., Song, M., and Teo, C.-P. (2009). Persistency model and its applications in choice modeling. *Management Sci.*, 55(3):453–469.

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.

Pratt, J. W. (1988). Aversion to one risk in the presence of others. *J. Risk Uncertain.*, 1(4):395–413.

Puccetti, G. and Rüschendorf, L. (2012). Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.*, 236(7):1833–1840.

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. (2005). Risky choices and correlated background risk. *Management Sci.*, 51(9):1336—345.

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

Wang, R., Xu, Z. Q., and Zhou, X. Y. (2019). Dual utilities on risk aggregation under dependence uncertainty. *Finance Stoch.*, 23(4):1025–1048.

Wiesemann, W., Kuhn, D., and Sim, M. (2014). Distributionally robust convex optimization. *Oper. Res.*, 62(6):1358–1376.

Zhu, S. and Fukushima, M. (2009). Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.*, 57(5):1155–1168.