Interval-based stochastic dominance: theoretical framework and application to portfolio choices

Jia Liu^{1,2} · Zhiping Chen^{1,2} · Giorgio Consigli^{3,4}

Accepted: 11 August 2021 / Published online: 30 August 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

We introduce a new stochastic dominance relationship, the *interval-based stochastic dominance* (ISD). By choosing different reference points, we show that ISD may span a continuum of preferences between k th and $(k + 1)$ th order stochastic dominance (SD). We distinguish accordingly between interval-based (or shortly just interval) SD of order 1 and of order 2: the former spanning from first- to second-order stochastic dominance, the latter from second- to third-order stochastic dominance. By examining the relationships between interval-based SD and SD, as well as between ISD and risk measures or utility functions, we frame the concept within decision theory and clarify its implications when applied to an optimal financial allocation problem. The formulation of ISD-constrained problems in the presence of discrete random variables is discussed in detail and applied to a portfolio selection problem.

Keywords Stochastic dominance · Bi-criteria decision making · Portfolio selection · Reference point

This research was supported by National Natural Science Foundation of China under Grant Numbers 11991023, 11901449 and 11735011.

✉ Giorgio Consigli
giorgio.consigli@ku.ac.ae

Jia Liu
jjialiu@xjtu.edu.cn

Zhiping Chen
zchen@xjtu.edu.cn

¹ School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, People's Republic of China

² Center for Optimization Technique and Quantitative Finance, Xi'an International Academy for Mathematics and Mathematical Technology, Xi'an 710049, People's Republic of China

³ Department of Mathematics, Khalifa University of Science and Technology, Abu Dhabi, United Arab Emirates

⁴ Department of Economics, University of Bergamo, Bergamo, Italy

1 Introduction

The concept of *stochastic dominance* (SD) originates in decision theory and economics to compare random outcomes (Fishburn 1964; Levy 2015). *First-order stochastic dominance* (FSD) defines a partial order between distribution functions of two random variables, say X and Y , defined on the same support and for which X gives at least as high a probability of receiving at least x as does Y , and for some x , X gives a higher probability than Y . *Second-order stochastic dominance* (SSD) defines an ordering relationship between integrated distribution functions. By recursively considering the integration of the associated distribution functions, we can extend the dominance relationship to higher integer orders. Since their origin, SD-based decision models have been studied in relationship with results from expected utility theory and risk measures (Bawa 1975; Bawa et al. 1985; Ogryczak and Ruszczyński 2001; Müller and Stoyan 2002; Levy 2015) with early applications to economics and finance in Quirk and Saposnik (1962), Hadar and Russell (1969) and Hanoch and Levy (1969).

The theory and computational approaches devoted to the formulation and solution of FSD- or SSD-constrained problems evolved consistently in recent years, without pretending to be exhaustive, thanks to Dentcheva and Ruszczyński (2003), Kuosmanen (2004), Noyan et al. (2006), Luedtke (2008), Gollmer et al. (2008), and Dupačová and Kopa (2014). Even if largely applied, SSD approaches may indeed have limited discriminating power as they do not rule out preferences for negatively skewed and fat-tailed distributions. On the other hand FSD conditions may be too hard to satisfy. Unlike FSD and SSD, *third-order stochastic dominance* (TSD) approaches can capture third-moment skewness preferences within the decision model. TSD and higher-order SD constraints were considered in Post and Kopa (2016), Chen and Jiang (2018), and Kallio and Hardoroudi (2019). More recently, optimization techniques for higher-order SD or decreasing absolute risk aversion have been proposed by Fang and Post (2017) and Post et al. (2015), respectively. Furthermore, outside parametric assumptions, distributionally robust SD constraints have also been studied by Dentcheva and Ruszczyński (2010a), Chen and Jiang (2018).

Optimization approaches with SD constraints became popular in finance after their application to several problem types: from financial planning problems (Post 2003; Kuosmanen 2004; Dentcheva and Ruszczyński 2006; Post and Kopa 2016; Kallio and Hardoroudi 2019) to enhanced portfolio indexation (Roman et al. 2013), asset-liability management (Yang et al. 2010; Consigli et al. 2019) and pension fund management (Kopa et al. 2018; Moriggia et al. 2019).

All the above developments rely on integer-based SD principles and already in the past an issue of hardly solvable FSD problems and rough spanning of agents' risk preferences was raised by several authors and motivated the search of more flexible approaches. Indeed, classical SD can not distinguish, for $k = 1, 2$, between k th- and $(k + 1)$ th-order SD-based preferences with respect to a benchmark distribution (Example 1 on page 4 shows that FSD and SSD can not distinguish two illustrative different risky assets). Furthermore, following the rationale proposed by Armbruster and Delage (2015) for SSD-constrained portfolio problems, an SSD constraint can be understood as a robust, worst-case, constraint imposed to an investor with arbitrary, even if concave, utility function, thus under an assumption of ambiguity over the decision maker risk preferences. We see below that, from a decision theoretic viewpoint, the SD criterion proposed in this article falls in this class.

Already in the 80s, Fishburn (1980) proposed fractional integration to define a continuum of stochastic dominance relations between integer SD partial orderings. The concept of

almost stochastic dominance (ASD) was proposed by Leshno and Levy (2002) with similar purposes and motivated by the evidence that FSD models, when solvable, often led to over-conservative policies. A similar drawback was later reported by Hu and Stepanyan (2017). Lizyayev and Ruszczyński (2012) studied an optimization problem with ASD constraints and found a tractable reformulation. Again to address the shortcomings of integer-based SD principles, Yang et al. (2010) proposed the *relaxed-interval* SSD between FSD and SSD with an application to asset-liability management. More recently, Tsetlin et al. (2015) further generalized ASD with improved computational properties and introduce the concept of *generalized almost SD* (GASD). Müller et al. (2016) also defined a continuum of SD between FSD and SSD, in this case, by relying on an application of utility theory. Finally, Hu and Stepanyan (2017) proposed a reference-based ASD, able to accommodate decision-makers' preferences and quantify their robustness with respect to alternative choices.

As in our approach, discussed below, the proposal of a reference-based ASD criterion leads to arguments similar to those put forward early on by prospect theorists Kahneman and Tversky (1979), who focused explicitly on reference points discriminating between lower and upper risk preferences. The associated concept of *Prospect Stochastic Dominance* (PSD) was indeed proposed by Levy and Levy (2002), and associated with all S-shaped utilities, i.e., convex on the negative part and concave on the positive part. Baucells and Heukamp (2006) further studied stochastic dominance induced by cumulative prospect theory and found generally supporting evidence for loss aversion.

The search for more general decision paradigms based on different types of relaxation of canonical SD criteria has thus been quite intense. Nevertheless, several issues remain untackled from either a decision theoretical or a specific financial perspective, and motivate this article. In particular:

- ASD or relaxed-interval SSD, in their original formulations, propose a partial order which is independent of the underlying loss magnitude and are thus unable to capture loss aversion, nor they may be adopted in the context of tail risk control as common in portfolio management. Indeed Generalized ASD (GASD) aimed at overcoming some limitations of ASD rankings, but still the focus is very much on the properties of utility functions and related agents' risk preferences, while in this article we cannot neglect the search for an effective risk control in a financial context. Unlike GASD, also, we do not need to assume any violation of canonical SD rules but rather FSD, SSD and TSD are presented as specific instances of ISD.
- The continuum of SD conditions proposed by Müller et al. (2016), specifically between first- and second-degree SD, concentrates on conditions to be satisfied by the marginal utility, setting a minimal rate of increase which is independent of the loss magnitude and constant over the loss domain, thus again unable to capture the variations of an investor's risk preferences and without any extension to higher degrees as in our context. The $1 + \gamma$ condition in Müller et al. (2016), however, is conceptually close to the minimal β derived through the bi-section method in Sect. 3 and it did actually inspire the definition of a maximum dominance level.
- PSD (Prospect SD) provides an essential decision criterion for investors with S-shape utilities. PSD however, when applied to a portfolio selection problem, undermines several key assumptions of a rational investor seeking a minimal risk exposure for given expected reward.

In this paper we follow this research line to extend SD principles to a theoretically continuous spanning of agents' risk preferences within an interval, from first to third-order SD

through a new partial ordering, based on the definition of first- and second-order interval-based stochastic dominance: ISD-1 and ISD-2, respectively, as rigorously defined below.

The article is organized as follows. In Sect. 2 we provide an example and convey the key motivations of the proposed generalization of SD conditions. In Sect. 3, we introduce formally the concept of Interval SD and study its' relationship with canonical SD, with risk measures and within utility theory. In Sect. 4, we derive reformulations of ISD-1 problems and efficient approximations of higher-order ISD problems for the case of discrete random variables. In Sect. 5, we introduce different portfolio selection models based on ISD principles, and in Sect. 6 an extended set of computational results is presented aimed at validating the introduced decision paradigm and analyzing its implications with an application to the US equity market. Conclusions follow.

2 Motivation and contribution

We intend to exemplify a market condition in which integer-based SD conditions are not sufficient to discriminate between two strategies and show how the ISD concept may help to choose between the two and bring about a specific issue of tail risk control below a reference point. Consider the following example.

Example 1 (Fig. 1) In a security market, there exists a market index Y and two portfolios, W and X , with the following return distributions:

- Y follows a uniform distribution on $[-1, 1]$;
- X follows a piecewise uniform distribution on $[-1, 1]$ with density

$$p(x) = \begin{cases} 1/8, & x \in [-1, -0.2], \\ 2, & x \in (-0.2, 0.1], \\ 1/3, & x \in (0.1, 1]; \end{cases}$$

Fig. 1 The distributions of W , X and Y in Example 1

- W follows a piecewise uniform distribution on $[-1, 1]$ with density

$$p(x) = \begin{cases} 4/11, & x \in [-1, 0.1], \\ 11/10, & x \in (0.1, 0.4], \\ 9/20, & x \in (0.4, 1]. \end{cases}$$

The distributions of X and W differ on the core portion of the support and the tails. Nevertheless, we can find that $X \succeq_{(2)} Y$, $W \succeq_{(2)} Y$; and $X \not\succeq_{(1)} Y$, $W \not\succeq_{(1)} Y$. It illustrates that FSD and SSD principles can not directly discriminate between W and X . Sometimes, the dominance constraint is the only factor to reflect the risk attitude in an optimization problem. Hence, a $(k + 1)$ th-order SD constraint would not be sufficient to describe agents' risk preferences, while a k th-order SD constraint may be too conservative. Hence, the need of a dominance relationship between k th- and $(k + 1)$ th-order SD constraints.

From the perspective of a portfolio manager/investor, she/he would surely prefer a strategy stochastically dominating the benchmark distribution Y to the first order. When neither W nor X are FSD-consistent with respect to Y as in the example, the portfolio manager will go for a strategy stochastically dominating the benchmark to the second order, where however both W and X would be acceptable, despite their different distribution shapes. Indeed W can dominate Y in the sense of FSD over a much wider portion of the support, $[-1, 0.345]$. Thus, in this example, the portfolio manager/investor does not need to go for SSD but insist on FSD criteria over a subset of the support. In practice, portfolio managers/investors are likely to be more sensitive to extreme losses on the left tail. In extreme scenarios, they would prefer to remain on the safe side, and an asymmetry between left and right tails would be natural. If we compare W and X directly, we may find that X dominates W in the FSD sense until point -0.1 and, thus, for a very risk-averse investor X could be preferred to W . Here, neither classical FSD nor SSD would have been able to clarify such evidence.

This way of reasoning can indeed be applied to comparisons not just over portfolio returns. Given a benchmark probability distribution, a decision-maker may find rational to discriminate between stochastic dominance conditions on different, contiguous, non-overlapping portions of the support and rely on a bi-criteria decision function. From which the idea of a decision paradigm based on a continuum of SD conditions in the $[1.0, 3.0]$ interval and a double parameterization of SD constraints: the *Interval-based Stochastic Dominance* of order k (ISD- k for short). This generalization relies on the definition of a reference point β and a different characterization of the decision criteria left and right of β , with k th-order SD conditions left and $(k + 1)$ th-order SD conditions right of such point. The reference point is given primarily a financial motivation in this article. By letting the reference point vary over the support, ISD-1 spans a continuum between FSD and SSD. Generalizing the definition to any k -th order, ISD- k provides a new SD criterion with any positive real number order. Differently from currently available SD-based decision frameworks, more generally, even considering recent non-integer and relaxed SD principles, which rely on a financial investors' unique risk characterization and rule out loss-dependent changes of risk preferences, we consider a bi-criteria decision-making paradigm. From a mathematical viewpoint, this change will lead to a double parametrization of SD constraints in an optimal decision problem. Such rationale will help discriminating among all those investors that may keep an overall risk-averse attitude and, at the same time, seek an effective tail risk control (see for instance the models in Roman et al. 2007; Gao et al. 2016).

Other examples may be proposed, outside portfolio selection, which rely on this bi-criteria decision approach in which an overall risk attitude is enriched with a specific threshold-based risk evaluation. These may reflect collective *public* decision making problems as in Atkinson (1987) and Killick (1995) with a welfare objective, or *private* entities allocation problems as

in Consigli et al. (2019). In what follows we stick to an individual optimal portfolio selection problem.

In relation to utility theory, the introduction of a reference point to span alternative SD conditions is delicate: specifically for ISD-1 problems, an equivalent *bi-criteria* reformulation involving jointly a class of utility functions and a class of downside risk functions can be determined. As shown in Sect. 5, ISD-1 implies indeed an investor who, given a, maybe bounded, $\mathbb{R}$ -valued portfolio domain, is risk-averse over such domain and particularly loss-averse below a pre-set reference point. Notice that, in general, as for any SD criterion, the specific utility function capturing a decision-maker overall risk aversion or the appropriate risk measure for her loss aversion with respect to extreme losses may both be assumed to be uncertain, thus not uniquely defined. In summary, the following can be regarded as main contributions of this research:

- The generalization of SD principles to a theoretically continuous spanning of agents' risk preferences within an interval, from first to third-order SD through a new partial ordering, based on the definition of first- and second-order interval stochastic dominance: ISD-1 and ISD-2, respectively.
- The definition of a viable approach to define the least feasible ISD-1 order in case of first-order SD problem infeasibility.
- The definition of a new risk measure, the *interval conditional value-at-risk* or Interval CVaR (ICVaR), whose relationship with other tail risk measures such as the VaR and the CVaR is clarified in the article.
- The proposed ISD-1 criterion provides a bi-criteria paradigm characterizing jointly risk aversion and, given a reference point, loss aversion to allow a stricter (relative to the entire loss domain) risk control over the extreme losses if compared with SSD. While, the ISD-2 criterion is linked to the bi-criteria paradigm through the concept of prudent utility (as for TSD) and ICVaR as associated risk measure for optimal control below the β .
- An extensive computational testing of the introduced ISD conditions of first and second orders on a range of portfolio models applied to the US market over the 2008–2021 period: specifically for financial planning problems we show that ISD- k , for $k = 1, 2$ leads to diversified optimal portfolios with distinctive features.

3 Interval-based stochastic dominance

We generalize canonical SD principles, whose mathematical characterization can be given first. Consider a probability space $(\Omega, \mathcal{F}, P)$ with its sample space Ω , σ -algebra $\mathcal{F}$ and associated probability measure P . A random variable W is defined on $L^m(\Omega, \mathcal{F}, P; \mathbb{R})$ with distribution function $F_1(W; \eta) := P(W \leq \eta)$, $\forall \eta \in \mathbb{R}$. $L^m(\Omega, \mathcal{F}, P; \mathbb{R})$ denotes the measurable functional space in which functionals mapping from the probability space to the value space $\mathbb{R}$ are integral to the m -th order. We define a series of non-decreasing functions recursively,

$$F_k(W; \eta) = \int_{-\infty}^{\eta} F_{k-1}(W; \xi) d\xi, \quad \forall \eta \in \mathbb{R}, \quad k = 2, 3, \dots, m+1.$$

By changing the order of integration and writing the probability function in functional expectation form, we have

$$F_k(W; \eta) = \frac{1}{(k-1)!} \mathbb{E}[(\eta - W)_+^{k-1}], \quad k \geq 1.$$

From which:

Definition 1 Given two random variables $W, Y \in L^{k-1}(\Omega, \mathcal{F}, P; \mathbb{R})$, we say that W dominates Y to the k th-order if

$$F_k(W; \eta) \leq F_k(Y; \eta), \quad \forall \eta \in \mathbb{R}. \quad (1)$$

We denote the k th-order SD relationship by $W \succeq_{(k)} Y$ for short.

Let Y be a benchmark random variable, we denote the set of W 's stochastically dominating Y as

$$A_k(Y) := \{W \in L^{k-1}(\Omega, \mathcal{F}, P; \mathbb{R}) : W \succeq_{(k)} Y\}.$$

An equivalent reformulation of k th-order SD constraints can be found in Dentcheva and Ruszczyński (2003).

We wish to extend the SD framework to allow a more accurate spanning of risk preferences. Consider, in particular, a case where k th-order SD constraints would lead to infeasibility of the associated optimization problem and let a reference point β to discriminate between k th and $(k+1)$ th SD.

Definition 2 Given two random variables $W, Y \in L^k(\Omega, \mathcal{F}, P; \mathbb{R})$, here $k \in \mathbb{N}$, we say that W stochastically dominates Y in the k th-based interval if, for given $\beta \in \mathbb{R}$, we have

$$\begin{cases} F_k(W; \eta) \leq F_k(Y; \eta), & \forall \eta \leq \beta, & (2-1) \\ F_{k+1}(W; \eta) \leq F_{k+1}(Y; \eta), & \forall \eta \geq \beta. & (2-2) \end{cases} \quad (2)$$

We denote this new dominance order by $W \succeq_{(\beta, k)} Y$. Moreover, we define the feasible set of W ISD- k dominating Y as

$$A_{(k, \beta)}(Y) := \{W \in L^k(\Omega, \mathcal{F}, P; \mathbb{R}) : W \succeq_{(\beta, k)} Y\}.$$

The $\beta \in \mathbb{R}$ in Definition 2 can be treated as a reference point for the dominance level. Below the β quantile, we adopt the stronger k th-order SD to describe the dominance relation; above the β quantile, we use the weaker $(k+1)$ th-order SD. β can be a reference point from a prospect theoretical viewpoint, or an exogenous parameter associated with a financial benchmark. In our formulation, the selection of β is very much related to a risk control problem. An investor may choose β according to behavioral or psychological considerations (which are endogenous to the decision problem formulation) or in relationship with the (exogenous) current market phase. For instance, when the decision is made from historical data, the investor may pre-set a probability level α , and let β to be the $(1 - \alpha)$ -quantile of the historical samples. Alternatively, investors and decision makers, may just be interested to that maximal β that will preserve ISD-1 feasibility. Notice that such β can be even updated dynamically according to previous decisions and investment outcomes (Baucells and Sarin 2010; Strub and Li 2019). We use the quantile method to derive a sample-based β in our numerical tests. Maybe what is sufficient to say on this specific point, is that, specifically in financial applications, we believe that such a reference point may very well depend on exogenous factors.

Remark 1 In (2), at the reference point β , both constraints are required to hold. The reason for this, similar to canonical SD results, is to ensure that the set of η for each constraint is closed, which would make it convenient for later discussion. Nevertheless, if we change the

interval of η in the k th-order SD constraint to $(-\infty, \beta)$, most of the conclusions that follow still hold.

With the help of the ISD concept, we can further notice from Example 1 that the dominance relationship between W and Y is stronger than that between X and Y on the left tail. By using the definition of ISD-1 with a fixed reference point, we have that, $X \succeq_{(1,0)} Y$, $W \succeq_{(1,0)} Y$; and $X \succeq_{(2)} Y$, $W \succeq_{(2)} Y$; but $X \not\succeq_{(1,0.2)} Y$, $W \succeq_{(1,0.2)} Y$. It means that, the ISD with $\beta = 0.2$ can distinguish the performances of X and W over Y , while neither FSD nor SSD can.

In the ISD constraint (2), we require the k th-order SD relationship to hold on $(-\infty, \beta]$, while on $[\beta, \infty)$ it is sufficient to satisfy the $(k+1)$ th-order SD relationship. Therefore, the ISD constraint can be regarded as a relaxation of k th-order SD constraints and a reinforcement of $(k+1)$ th-order SD constraints, in short, denoted by ISD- k in what follows. In practice, different β 's will correspond to different degrees of risk aversion and in financial applications be associated with a target return, as a reference value.

For constraints induced by the above ISD relationship, the selection of β is thus important, which depend largely investors' risk attitude. On one hand, we should choose β to preserve the feasibility of the resulting SD constrained problem. On the other hand, we should select the largest possible β so that the dominance constraints are tight and the resulting optimal solutions have better performance. To this end, we introduce the following concept.

Definition 3 Given two random variables $X, Y \in L^k(\Omega, \mathcal{F}, P; \mathbb{R})$ satisfying $X \succeq_{(k,\beta)} Y$, we define the maximum dominance level $\beta_k(X, Y)$ as the largest possible β such that $X \succeq_{(k,\beta)} Y$ holds. That is,

$$\beta_k(X, Y) = \sup \{ \beta \in \mathbb{R} \mid X \succeq_{(k,\beta)} Y \}.$$

Take ISD-1 for instance, $\beta_1(X, Y)$ is the first crossing point when the distribution function of X up-crosses the distribution function of Y .

Proposition 1 For any two normal distributions X and Y with the same mean value μ and different standard deviations σ_X, σ_Y ($\sigma_X < \sigma_Y$), $\beta_1(X, Y) = \mu$, say, X always dominates Y to the first-order interval with respect to any reference point $\eta \leq \mu$.

In Example 1, the maximum dominance level between X and Y is $\beta_1(X, Y) = 0$, while the maximum dominance level between W and Y is $\beta_1(W, Y) = 0.345$. By choosing a reference point between 0 and 0.345, the ISD constraint could distinguish between X and W over Y .

As for the dominance relationship between W and X , although X seems dominating W in the first-order-interval on the left tail, X does not dominate W in the sense of ISD-1 with any reference point β . The point is that X does not dominate W to the second order. While ISD-2 can further explain the dominance relationship between X and W . We draw $F_2(\cdot, \eta)$ and $F_3(\cdot, \eta)$ curves of the three distributions in Figs. 2 and 3, respectively. Figure 3 shows that X dominates W to the third-order. Thus, we know there exists a β such that X dominates W to the second-order-interval. Figure 2 tells us that the maximum $\beta_k(X, W)$ letting $X \succeq_{(2,\beta)} W$ is 0.468.

Remark 2 Instead of mixing two SD constraints with adjacent orders, one may consider just dropping the RHS constraints, above β , or (2-2). This can be viewed as a simplification of k th-order SD constraints within limited support set. In the final part of the numerical tests, we will examine the influence of dropping the constraints associated with the domain beyond

Fig. 2 The $F_2(\cdot, \eta)$ of W , X and Y

Fig. 3 The $F_3(\cdot, \eta)$ of W , X and Y

β . By using the new concept, we can find in Example 1 that X dominates W in the first-order-interval on the left tail until point -0.1 , thus for a very risk-averse investor X could be preferred to W over the interval $[-1, -0.1]$.

3.1 Relationship with stochastic dominance

Since k th-order SD implies $(k + 1)$ th-order SD and $\text{ISD-}k$ lies between the two, we can establish the following relationship:

Proposition 2 For any $\beta \in \mathbb{R}$ we have

$$W \succeq_{(k)} Y \Rightarrow W \succeq_{(k, \beta)} Y \Rightarrow W \succeq_{(k+1)} Y,$$

and

$$A_k(Y) \subseteq A_{(k,\beta)}(Y) \subseteq A_{k+1}(Y).$$

Proof : The key point for the second relationship is to show that (2-1) implies $F_{k+1}(W; \eta) \leq F_{k+1}(Y; \eta)$, $\forall \eta \leq \beta$. From (2-1), we know $F_k(W; \xi) \leq F_k(Y; \xi)$, $\forall \xi \leq \eta \leq \beta$. Hence,

$$\int_{-\infty}^{\eta} F_k(W; \xi) d\xi \leq \int_{-\infty}^{\eta} F_k(Y; \xi) d\xi, \quad \forall \eta \leq \beta,$$

which is the desired result. The second relationship follows similarly. $\square$

Thus β can span from the k th- to $(k + 1)$ th-order SD. Furthermore:

Proposition 3 When $\beta \leq \inf_{y \in \text{supp}(Y)} y$, ISD- k is equivalent to $(k + 1)$ th-order SD. When $\beta \rightarrow +\infty$, ISD- k is asymptotically equivalent to k th-order SD. For $k = 1$, when $\beta \geq \sup_{y \in \text{supp}(Y)} y$, ISD-1 is equivalent to FSD.

Proof : First, when $\beta \leq \inf_{y \in \text{supp}(Y)} y$, we have $F_k(Y; \eta) = F_{k+1}(Y; \eta) = 0$, $\forall \eta \leq \beta$. It means that, in this case, $W \geq \eta$, for all $\eta \leq \beta$. This implies that (2-1) is equivalent to $F_{k+1}(W; \eta) \leq F_{k+1}(Y; \eta)$, $\forall \eta \leq \beta$. Together with (2-2), we can obtain the equivalence between ISD- k and k th-order SD.

Second, equivalence in the limit when $\beta \rightarrow +\infty$ can be trivially recovered from (2).

Finally, when $\beta \geq \sup_{y \in \text{supp}(Y)} y$, $F_1(Y; \eta) = P(Y \leq \eta) = 1$ for any $\eta \geq \beta$. In this case (2-1) will imply the FSD constraint. $\square$

Remark 3 Following Propositions 2 and 3, for varying β with ISD- k ordering and increasing $k = 1, 2, 3$, we span a continuum of stochastic dominance degrees. In practice, $k = 1, 2$ and β will be given a discrete finite set of values. Thus when a decision is made from historical data, the investor may pre-set a probability level α , and let β be the $(1 - \alpha)$ -quantile of the historical samples. In this case, we may denote with ISD- $k.\alpha$ the ISD condition in which β is the $(1 - \alpha)$ -quantile. Relying on this notation, we see that we are approximating a continuous ordering scheme between traditional integer-order SD: FSD will correspond to ISD-1.0, SSD to ISD-2.0, TSD to ISD-3.0, and ISD- $k.\alpha$ for different α values will span the interval between the integer orders.

Whenever the probability distribution of Y is assumed with continuous support, as say the real line for a normal distribution, then if we let the reference point go to $-\infty$, the loss aversion will disappear, and ISD-1 will converge to SSD. By letting, instead, the reference point tend to ∞ , the loss aversion will span the entire support, and thus ISD-1 will converge to FSD. While, in case of finite support, as in Example 1 on Page 5, FSD corresponds to ISD-1 for $\beta = 1$ (denoted by ISD-1.0), while SSD is equivalent to ISD-1 for $\beta = -1$ as well as to ISD-2 for $\beta = 1$ (denoted by ISD-2.0).

3.2 ISD-1: relation to utility functions and risk measures

We clarify the relationship between ISD-1 principles and utility functions and risk measures. As FSD implies SSD, we have by Proposition 2 that ISD-1 is equivalent to

$$\begin{cases} F_1(W; \eta) \leq F_1(Y; \eta), & \forall \eta \leq \beta, & (3.1) \\ F_2(W; \eta) \leq F_2(Y; \eta), & \forall \eta \in \mathbb{R}. & (3.2) \end{cases} \quad (3)$$

We have:

Proposition 4 ISD-1 is equivalent to

$$\begin{cases} \mathbb{E}[u(W)] \geq \mathbb{E}[u(Y)], \quad \forall u \in \mathcal{U}_S, & (4.1) \\ \mathbb{E}[r(W)] \leq \mathbb{E}[r(Y)], \quad \forall r \in \mathcal{R}_F. & (4.2) \end{cases} \quad (4)$$

where

$$\mathcal{U}_S = \{u : u \text{ is monotone increasing and concave on } \mathbb{R}\},$$

and

$$\mathcal{R}_F = \{r : r \text{ is monotone decreasing on } (-\infty, \beta], \text{ and } r(x) = 0, \quad \forall x > \beta\}.$$

Proof The constraint (3.1) is equivalent to

$$\mathbb{E}[1_{(-\infty, \eta)}(W)] \leq \mathbb{E}[1_{(-\infty, \eta)}(Y)], \quad \forall \eta \leq \beta,$$

where $1_{(-\infty, \eta)}(\cdot)$ is a risk function in $\mathcal{R}_F$. Thus, (3.1) follows from (4.1).

On the other side, given a tolerance error ϵ , any function $r(\cdot)$ in $\mathcal{R}_F$ which is decreasing on $(-\infty, \beta]$ and zero-valued on (β, ∞) can be approximated by a step function, $r_n(\cdot) = \sum_{i=1}^n \alpha_i 1_{(-\infty, z_i]}(\cdot) + \alpha_0 1_{(-\infty, \beta]}(\cdot)$ with $\alpha_0 \in \mathbb{R}$, $\alpha_i > 0$ and $z_i \leq \beta$, $i = 1, \dots, n$, such that $|\mathbb{E}[r(W)] - \mathbb{E}[r_n(W)]| \leq \epsilon$ and $|\mathbb{E}[r(Y)] - \mathbb{E}[r_n(Y)]| \leq \epsilon$. Then, (3.1) implies

$$\sum_{i=1}^n \alpha_i \mathbb{E}[1_{(-\infty, z_i)}(W)] + \alpha_0 \mathbb{E}[1_{(-\infty, \beta]}(W)] \leq \sum_{i=1}^n \alpha_i \mathbb{E}[1_{(-\infty, z_i)}(Y)] + \alpha_0 \mathbb{E}[1_{(-\infty, \beta]}(Y)],$$

which is $\mathbb{E}[r_n(W)] \leq \mathbb{E}[r_n(Y)]$, and consequently $\mathbb{E}[r(W)] \leq \mathbb{E}[r(Y)]$ by setting a small enough ϵ . Thus, (4.1) follows from (3.1).

(3.2) is just SSD, which is equivalent to $\mathbb{E}[u(W)] \geq \mathbb{E}[u(Y)]$, $\forall u \in \mathcal{U}_S$, where $\mathcal{U}_S = \{u : u \text{ is monotone increasing and concave on } \mathbb{R}\}$. $\square$

Proposition 4 tells us that ISD-1 requires the order relationship to hold not only for all monotone increasing and concave utilities on the whole support but also for downside risk functions which are decreasing on the left part of the support, no matter if concave or convex.

Notice that, when β tends to $-\infty$, the constraints (4.1) would take no effect, and (4) reduces to SSD. When β tends to $+\infty$, (4.2) would imply (4.1) and (4) would reduce to FSD. Hence, by increasing β we would span the SSD–FSD interval.

We call risk functions in $\mathcal{R}_F$ downside risk functions. These include *lower partial moments* $\mathbb{E}[-(\beta - X)_+]^p$ of order p , which for $p = 1$ define the *expected regret* measure and for increasing p would increase the penalty on losses relative to the reference β . Downside risk functions include the exponentially weighted mean square risk (Satchell et al., 2000) as well: $\mathbb{E}[\omega(X)(X - \beta)^2]$ with $\omega(X) = -e^{-\theta(X - \beta)}$ for $X \leq \beta$ and $\omega(X) = 0$ for $x > \beta$.

A lower partial moment risk constraint with order p for all reference points $\beta \in \mathbb{R}$ corresponds to a $(p - 1)$ -th order SD constraint (Bawa 1975; Porter 1974). In contrast, the downside risk constraints in (4) provide risk control for any type of downside risk measure but only below β . Notice that, for $p \geq 2$ and a reference point below β , adopting the negative of lower partial moment as a utility function both (4.1) and (4.2) apply.

The set-up behind ISD-1-based decision making is thus consistent with the loss-aversion philosophy of behavioral finance. Given a benchmark, Kahneman and Tversky (2019) in their early contribution, pointed out that agents' degree of risk-aversion depends on loss-tolerance due to systematic biases of human behavior. In an ISD-1 setting, we would argue that the investor follows a bi-criteria approach: she/he is risk-averse over the portfolio domain

(equivalent to all risk aversion above the reference point), but remains with an undetermined risk attitude below the benchmark.

Looking back to Example 1 in Sect. 2, we can find that, $X \succeq_{(1,0)} Y$ implies first that X is always preferred to Y by any concave and increasing utility function on $[-1, 1]$ and second that X is less risky than Y on $[-1, 0]$, as measured by any downside risk function with negative reference point. Meanwhile, W surely SSD-dominates Y , but it is also less risky over $[-1, 0.345]$, consistently with a 0.345 reference point.

In summary, ISD-1 captures different risk attitudes at different loss levels: global risk-aversion over the full support (equivalent to risk aversion above the reference point) and stronger risk control for investment results below the reference point.

3.3 ISD-2 and risk measures

We can extend the previous reasoning to ISD-2 and discuss its relationship with risk measures. Again as SSD implies TSD, ISD-2 is equivalent to

$$\begin{cases} F_2(W; \eta) \leq F_2(Y; \eta), & \forall \eta \leq \beta, \\ F_3(W; \eta) \leq F_3(Y; \eta), & \forall \eta \in \mathbb{R}. \end{cases} \quad (5)$$

The second constraint in (5) is just TSD. Dentcheva and Ruszczyński (2006) showed that the SSD constraint is related to infinite conditional value-at-risk (CVaR) constraints. We have a similar property for the first constraint in (5).

Proposition 5 *The constraint*

$$F_2(W; \eta) \leq F_2(Y; \eta), \quad \forall \eta \leq \beta,$$

is equivalent to

$$\rho_{\alpha, \beta}(W) \geq \rho_{\alpha, \beta}(Y), \quad \forall \alpha \in [0, 1),$$

where

$$\rho_{\alpha, \beta}(W) = \sup_{\eta \leq \beta} \left\{ \eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - W]_+ \right\}, \quad \alpha \in [0, 1).$$

Proof First, as $F_2(W; \eta) = \mathbb{E}[\eta - W]_+$, $F_2(W; \eta) \leq F_2(Y; \eta), \forall \eta \leq \beta$, implies $\eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - W]_+ \geq \eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - Y]_+, \forall \eta \leq \beta, \alpha \in [0, 1)$, which further implies $\rho_{\alpha, \beta}(W) \geq \rho_{\alpha, \beta}(Y), \quad \forall \alpha \in [0, 1)$.

As for the converse, for any α , select $q_\alpha^* = \text{VaR}(W; q) = \arg \sup_{\eta \in \mathbb{R}} \{ \eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - W]_+ \}$. For all $q_\alpha^* \leq \beta$, we have $q_\alpha^* - \frac{1}{1-\alpha} \mathbb{E}[q_\alpha^* - W]_+ = \rho_{\alpha, \beta}(W) \geq \rho_{\alpha, \beta}(Y) \geq q_\alpha^* - \frac{1}{1-\alpha} \mathbb{E}[q_\alpha^* - Y]_+$, and thus $\mathbb{E}[q_\alpha^* - W]_+ \leq \mathbb{E}[q_\alpha^* - Y]_+$. By adjusting for every α such that $q_\alpha^* \leq \beta$, we get the assertion. $\square$

We find that ISD-2 is related to a new risk measure $\rho_{\alpha, \beta}(W)$: we call it *ISD-2 induced Conditional Value-at-Risk* (ICVaR). The only difference between ICVaR and CVaR is that the supreme is taken over $(-\infty, \beta]$ rather than over $\mathbb{R}$. We have:

Proposition 6 *For $\beta \geq \text{VaR}_\alpha(W)$,*

$$\rho_{\alpha, \beta}(W) = \text{CVaR}_\alpha(W);$$

while for $\beta \leq \text{VaR}_\alpha(W)$,

$$\rho_{\alpha, \beta}(W) = \beta - \frac{1}{1-\alpha} \mathbb{E}[\beta - W]_+.$$

Fig. 4 ICVaR and CVaR

Proof The global optimal value of $\sup_{\eta \in \mathbb{R}} \{\eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - W]_+\}$ is $\text{CVaR}_\alpha(W)$, while the optimal solution is $\eta^* = \text{VaR}_\alpha(W)$, which is essentially a stationary point. If the stationary point η^* is located in $(-\infty, \beta]$, the optimal η for $\sup_{\eta \leq \beta} \{\eta - \frac{1}{1-\alpha} \mathbb{E}[\eta - W]_+\}$ is η^* , as it is a concave maximization problem. If not, the optimal η should be on the bound, say, $\eta = \beta$. $\square$

It is worth noting that, CVaR can be defined on the loss or the return functions. Here, we use the definition from Dentcheva and Ruszczyński (2006), $\text{CVaR}_\alpha(W) = \sup_{\eta \in \mathbb{R}} \{\eta - \frac{1}{\alpha} \mathbb{E}[\eta - W]_+\}$, to remain consistent with the introduced notation.

Proposition 6 implies that $\rho_{\alpha, \beta}(W)$ is always smaller than or equal to $\text{CVaR}_\alpha(W)$. When $\text{VaR}_\alpha(W)$ is smaller than or equal to the preset benchmark β , the investor would just use CVaR to measure the risk. When $\text{VaR}_\alpha(W)$ is greater than β , the investor would just focus on the loss beyond the benchmark.

The new risk measure applies to those losses larger than both the benchmark β and the quantile estimation $\text{VaR}_\alpha(W)$. Figure 4 shows two cases: $\rho_{\alpha, \beta_1}(W) = \beta - \frac{1}{1-\alpha} \mathbb{E}[\beta_1 - W]_+$, where β_1 is smaller than $\text{VaR}_\alpha(W)$; and $\rho_{\alpha, \beta_2}(W) = \text{CVaR}_\alpha(W)$ where β_2 is larger than $\text{VaR}_\alpha(W)$. The established equivalence between ISD-2 and the ICVaR allows the formulation of a decision problem based on the canonical risk-return trade-off criterion, where the risk is in this case captured by the (ISD-2 consistent) ICVaR measure. We will further develop such parallelism in a separate contribution.

In summary, Proposition 5 clarifies that ISD-2 corresponds to investors who have a prudent utility as well as risk aversion below a reference point. The prudence is represented by the TSD constraint while risk-aversion is now characterized by a group of ICVaR constraints.

4 Sample reformulation of ISD constraints

From the definition of $F_k(\cdot)$, we can write $W \succeq_{(k, \beta)} Y$ equivalently in terms of infinite expectation value constraints

$$\begin{cases} \mathbb{E}[(\eta - W)_+^k] \leq \mathbb{E}[(\eta - Y)_+^k], & \forall \eta \leq \beta, \\ \mathbb{E}[(\eta - W)_+^{k+1}] \leq \mathbb{E}[(\eta - Y)_+^{k+1}], & \forall \eta \geq \beta. \end{cases} \quad (6)$$

Especially, when $k = 1$, $W \succeq_{(\beta,1)} Y$ is equivalent to

$$\mathbb{P}(W \leq \eta) \leq \mathbb{P}(Y \leq \eta), \quad \forall \eta \leq \beta, \quad (7)$$

$$\mathbb{E}[(\eta - W)_+] \leq \mathbb{E}[(\eta - Y)_+], \quad \forall \eta \geq \beta. \quad (8)$$

An optimization problem with the ISD constraints (6) is generally intractable due to the infinite constraints.

4.1 ISD-1 constraints

In case of Y with finite support and $k = 1$, FSD constraints (7) can be simplified into a group of finite constraints. We assume that the reference random variable Y has D possible realizations, y_i carrying probability p_i , $i = 1, 2, \dots, D$. Without loss of generality, we assume that $y_1 < y_2 < \dots < y_D$. From Proposition 3, we know that an ISD-1 constraint is equivalent to an FSD constraint when $\beta \geq y_D$ and to an SSD constraint when $\beta \leq y_1$. The reformulation of FSD and SSD constraints for discretely distributed Y can be found in Dentcheva and Ruszczyński (2003; 2006) and Luedtke (2008). Here, the only case we need to discuss is $y_1 < \beta < y_D$. As $y_1 < y_2 < \dots < y_D$, we can find a unique index $1 \leq l < D$ such that $y_l < \beta \leq y_{l+1}$.

Theorem 1 *When Y is discretely distributed with finite realizations, ISD-1 constraints (7–8) are equivalent to the following finite constraints:*

$$\mathbb{P}(W < y_i) \leq \mathbb{P}(Y \leq y_{i-1}), \quad \forall i = 2, \dots, l, \quad (9)$$

$$\mathbb{P}(W < \beta) \leq \mathbb{P}(Y \leq y_l), \quad (10)$$

$$\mathbb{P}(W \leq \beta) \leq \mathbb{P}(Y \leq \beta), \quad (11)$$

$$W \geq y_1, \text{ w.p.1}, \quad (12)$$

$$\mathbb{E}[(y_i - W)_+] \leq \mathbb{E}[(y_i - Y)_+], \quad \forall i = l + 1, \dots, D, \quad (13)$$

$$\mathbb{E}[(\beta - W)_+] \leq \mathbb{E}[(\beta - Y)_+]. \quad (14)$$

If only the SSD constraint is required to hold at β , then by changing the inequality $\eta \leq \beta$ in (7) into a strict inequality $\eta < \beta$, we can reformulate ISD-1 constraints as (9–14) without (11).

Proof To prove Theorem 1, we show the equivalence between (7) and (9–12), as well as the equivalence between (8) and (13–14), respectively.

First, we show that (7) implies (9–12). We have four cases:

- $\eta = y_1 - \delta$, where δ is a very small positive number. Then (7) is $\mathbb{P}(W \leq y_1 - \delta) \leq \mathbb{P}(Y \leq y_1 - \delta) = 0, \forall \delta > 0$, which is equivalent to $W \geq y_1$, w.p.1., that is, (12).
- $\eta = y_i - \delta$, where $0 < \delta \leq y_i - y_{i-1}$, $i = 2, \dots, l$. We have that $\mathbb{P}(W \leq y_i - \delta) \leq \mathbb{P}(Y \leq y_i - \delta) = \mathbb{P}(Y \leq y_{i-1}), \forall \delta \in (0, y_i - y_{i-1}]$, which implies (9).
- (7) with $\eta = \beta - \delta$, where $0 < \delta \leq \beta - y_l$, implies (10).
- (7) with $\eta = \beta$ is equivalent to (11).

Then, we show that (9–12) imply (7). We divide the support into four parts:

- (12) implies $\mathbb{P}(W < y_1) = 0$, i.e., (7) holds for any $\eta < y_1$.
- For any $y_1 \leq \eta \leq \beta$, either $y_i \leq \eta < y_{i+1}$, for some i , or $y_l \leq \eta \leq \beta$. For an η in $[y_i, y_{i+1})$, we have from (9) that

$$\mathbb{P}(W \leq \eta) \leq \mathbb{P}(W < y_{i+1}) \leq \mathbb{P}(Y < y_{i+1}) = \mathbb{P}(Y \leq y_i) \leq \mathbb{P}(Y \leq \eta).$$

- Similarly, for $\eta \in [y_l, \beta)$, we have from (10) that

$$\mathbb{P}(W \leq \eta) \leq \mathbb{P}(Y \leq \eta).$$

- (11) corresponds to the last case of (7) with $\eta = \beta$.

The above discussion shows the equivalence between (7) and (9–12).

As for the equivalence between (8) and (13–14), the necessity of (13) and (14) for (8) is obvious, while the sufficiency comes from the convexity of $F_2(W) = \mathbb{E}[(\eta - W)_+]$. The sufficiency proofs for $\eta \in [y_i, y_{i+1}]$, $i = l + 1, \dots, D - 1$, and $\eta > y_D$, are similar to Case 2 and Case 3 in the proof of Proposition 3.2 in Dentcheva and Ruszczyński (2003). The case for $\eta \in [\beta, y_{l+1}]$ is similar to the case for $\eta \in [y_i, y_{i+1}]$. $\square$

Furthermore, we assume that W also follows a discrete distribution with N realizations, w_i with probability q_i , $i = 1, \dots, N$. Then, by introducing new binary auxiliary parameters $c_{i,j}$, $j = 1, \dots, N$, $i = 1, \dots, l + 1$, and continuous auxiliary parameters $a_{i,j}$, $j = 1, \dots, N$, $i = l, \dots, D$, we can reformulate (9–14) as follows:

$$\sum_{j=1}^N q_j c_{i,j} \leq \sum_{s=1}^{i-1} p_s, \quad i = 2, \dots, l + 1, \quad (15)$$

$$w_j + M c_{i,j} \geq y_i, \quad j = 1, \dots, N, \quad i = 2, \dots, l, \quad (16)$$

$$w_j + M c_{l+1,j} \geq \beta, \quad j = 1, \dots, N, \quad (17)$$

$$\sum_{j=1}^N q_j c_{1,j} \leq P(Y \leq \beta), \quad (18)$$

$$w_j - M(1 - c_{1,j}) \leq \beta, \quad j = 1, \dots, N, \quad (19)$$

$$c_{i,j} \in \{0, 1\}, \quad j = 1, \dots, N, \quad i = 1, \dots, l + 1, \quad (20)$$

$$w_j \geq y_1, \quad j = 1, \dots, N, \quad (21)$$

$$\sum_{j=1}^N q_j a_{i,j} \leq \sum_{s=1}^D p_s (y_i - y_s)_+, \quad i = l + 1, \dots, D, \quad (22)$$

$$\sum_{j=1}^N q_j a_{l,j} \leq \sum_{s=1}^D p_s (\beta - y_s)_+, \quad (23)$$

$$y_i - w_j \leq a_{i,j}, \quad j = 1, \dots, N, \quad i = l + 1, \dots, D, \quad (24)$$

$$\beta - w_j \leq a_{l,j}, \quad j = 1, \dots, N, \quad (25)$$

$$a_{i,j} \geq 0, \quad j = 1, \dots, N, \quad i = l, \dots, D. \quad (26)$$

Here, the chance constraints, equivalent to the expectation of some indicator functions, are reformulated as linear constraints (15–20) by using the Big-M method, which provides a relatively straightforward procedure (Bonami et al. 2015) in this case. The idea is to switch on and off the constraint in the indicator function by multiplying the binary variable by a very large constant number M , so that (16) or (17) always holds when $c_{i,j} = 1$.

4.2 Sample approximations of higher-order ISD

Constraints (15–26) provide an equivalent reformulation of first-order ISD constraints when both W and Y are discretely distributed with finite bounded samples. However, for $(k \geq 2)$ -th-

order ISD constraints, even if both W and Y follow discrete distributions with N and D samples, the approach adopted above is not viable. We can nevertheless find inner and outer approximations of the feasible region and in practice adopt one of the two.

First, we directly choose a finite set of constraints from (6), which gives an outer approximation. Say, we generate H_l points between $\min\{y_1, w_1\}$ and β : $\eta_1^l, \eta_2^l, \dots, \eta_{H_l}^l$, and H_u points between β and $\max\{y_D, w_N\}$: $\eta_1^u, \eta_2^u, \dots, \eta_{H_u}^u$. Then we use $H_l + H_u$ constraints on $\{\eta_1^l, \eta_2^l, \dots, \eta_{H_l}^l, \eta_1^u, \eta_2^u, \dots, \eta_{H_u}^u\}$, to approximate (6):

$$\begin{cases} \sum_{j=1}^N q_j [(\eta_h^l - w_j)_+^k] \leq \sum_{i=1}^D p_i [(\eta_h^l - y_i)_+^k], & h = 1, \dots, H_l, \\ \sum_{j=1}^N q_j [(\eta_h^u - w_j)_+^{k+1}] \leq \sum_{i=1}^D p_i [(\eta_h^u - y_i)_+^{k+1}], & h = 1, \dots, H_u, \end{cases} \quad (27)$$

(27) can be further reformulated as the following group of polynomial constraints:

$$\sum_{j=1}^N q_j c_{h,j} \leq \sum_{i=1}^D p_i [(\eta_h^l - y_i)_+^k], \quad h = 1, \dots, H_l, \quad (28)$$

$$(\eta_h^l - w_j)^k \leq c_{h,j}, \quad j = 1, \dots, N, \quad h = 1, \dots, H_l, \quad (29)$$

$$0 \leq c_{h,j}, \quad j = 1, \dots, N, \quad h = 1, \dots, H_l, \quad (30)$$

$$\sum_{j=1}^N q_j a_{h,j} \leq \sum_{i=1}^D p_i [(\eta_h^u - y_i)_+^{k+1}], \quad h = 1, \dots, H_u, \quad (31)$$

$$(\eta_h^u - w_j)^{k+1} \leq a_{h,j}, \quad j = 1, \dots, N, \quad h = 1, \dots, H_u, \quad (32)$$

$$0 \leq a_{h,j}, \quad j = 1, \dots, N, \quad h = 1, \dots, H_u. \quad (33)$$

(28–33) provide an outer approximation to (6). To choose the η 's, one can select some equally spaced points between $\min\{y_1, w_1\}$ and $\max\{y_D, w_N\}$, or use a sample from the data history. (28–33) with η 's from an historical sample is the approximation method adopted in the case study.

In order to get a conservative (inner) approximation, we can follow the method for third-order SD from Bawa et al. (1985) and Post and Kopa (2016), and generalize it to higher-order ISD. This requires to first pre-set a tolerance error ϵ , and then choose a series of $\hat{\eta}_1^l, \hat{\eta}_2^l, \dots, \hat{\eta}_{H_l}^l$, such that, $\hat{\eta}_1^l = \min\{y_1, w_1\}$, $\hat{\eta}_{H_l}^l = \beta$. We then apply Post and Kopa (2016).

4.3 Determining the maximum dominance level

In many portfolio selection problems, such as an index tracking or enhanced index-tracking problems, SD based models often lead to infeasibility either because the index portfolio is not included in the feasible portfolio set or because the benchmark portfolio strategy is too ambitious.

In ISD modeling, the reference point β plays an important role in the infeasibility issue. The higher the β , the stronger the ISD constraint. However, when β is too high, the ISD constraint may lead to infeasibility, like FSD. Hence, choosing a proper β is a crucial issue. The decision-maker could set a fixed β according to his/her tolerance level of risk, or he/she could try to find the maximum β to keep the ISD constraint feasible: this is referred to as the maximum feasible dominance level.

$$\beta_k(X, Y) = \sup \{ \beta \in \mathbb{R} \mid X \succeq_{(k, \beta)} Y \}. \quad (34)$$

The derivation of such optimal β (or l equivalently in a finite sample case) is not straightforward even for discrete samples. From an algorithmic viewpoint, problem (34) may be tackled with a bisection method. Suppose X and Y follow discrete distributions with N samples. Then the outline of the bisection method to find the maximum β for ISD-1 is given in Algorithm 1. The algorithmic complexity of the bisection method is $O(\log_2 N)$, say, we have to solve ISD-1, FSD, or SSD for around $\log_2 N + 2$ times during the bisection method.

Algorithm 1 Bisection method to find maximum β for ISD-1

```

1: if FSD is feasible then
2:    $l^* \leftarrow N$ 
3: else if SSD is infeasible then
4:    $l^* \leftarrow 0$ 
5: else
6:    $a \leftarrow 0, b \leftarrow N$ 
7:   while  $b - a > 1$  do
8:      $l \leftarrow \text{floor}((b + a)/2)$ ;
9:     if ISD-1 with  $\beta = l/N$ -quantile is feasible then
10:       $a \leftarrow l$ 
11:     else
12:       $b \leftarrow l$ 
13:     end if
14:   end while
15:    $l^* \leftarrow a$ 
16: end if
17:  $\beta \leftarrow$  the  $l^*$ th smallest sample of  $Y$ 

```

Based on the bisection method, specifically for ISD-1 problems, we can identify the sup β to ensure the problem feasibility, the closest from a computational perspective to the FSD problem. We show in the next section that indeed such problem formulation has a strong rationale in the context of a portfolio allocation problem.

5 Portfolio selection with ISD constraints

We study the application of ISD constraints to a one-period portfolio selection problem. We consider a security market consisting of n risky assets and a market index. In general, such index or any other benchmark strategy is correlated with the investment universe, which will typically include a subset of the index-constituent assets. We denote the random returns of the n risky assets by $r = [r_1, r_2, \dots, r_n]^\top$. The return rate of the market index is denoted by y . A portfolio $u = [u_1, u_2, \dots, u_n]^\top$ is an allocation of the investor's initial wealth x_0 among the n risky assets. We can set $x_0 = 1$, hence u represents a vector of investment proportions. Then, we have $e^\top u = x_0$, where $e = [1, 1, \dots, 1]^\top$. The wealth of the investor at the end of the period is $x = r^\top u$. k th order ISD conditions are associated with the market benchmark and expressed by the constraints $r^\top u \succeq_{(k, \beta)} y$: y plays the role of a financial benchmark for the portfolio u . By choosing y as the market index, we get a pure *index-tracking* problem; while by letting y to be an index over the market index, we have the so-called α -strategies. In the case study below, we will consider an index-tracking problem only.

Given ISD- k feasibility, the investor wants to maximize her expected terminal wealth. Then, with β as the reference point and ruling out short selling, we can formulate the problem as follows:

$$\max_u \mathbb{E}[r^\top u] \quad (35)$$

$$\text{s.t. } r^\top u \succeq_{(k,\beta)} y, \quad (36)$$

$$e^\top u = x_0, \quad (37)$$

$$u \geq 0. \quad (38)$$

We denote problem (35) under (36–38) by ISD-1 for order $k = 1$, and ISD-2 for order $k = 2$.

As the market index is determined by the prices of all assets in the market, it is natural to believe that both r and y are driven by the same stochastic events. We can treat r and y as belonging to the same probability space. A popular way to describe the randomness of r and y is by using their historical data. We draw N pairs of historical samples of r and y . The samples of r are denoted by $r^1, \dots, r^N$, with equal probability $q_j = \frac{1}{N}$, $j = 1, \dots, N$. Then we merge the samples of y with the same realization values into D samples satisfying $y_1 < y_2 < \dots < y_D$, with probabilities $p_1, \dots, p_D$.

Consider first-order ISD. With the sample assumption, we can adopt the method proposed in Sect. 4 and reformulate ISD-1 as:

$$\max_{u,c,a} \sum_{j=1}^N q_j (u^\top r^j) \quad (39)$$

$$\text{s.t. } u^\top r^j + M c_{i,j} \geq y_i, \quad j = 1, \dots, N, \quad i = 2, \dots, l, \quad (40)$$

$$u^\top r^j + M c_{l+1,j} \geq \beta, \quad j = 1, \dots, N, \quad (41)$$

$$\sum_{j=1}^N q_j c_{1,j} \leq P(y \leq \beta), \quad (42)$$

$$u^\top r^j - M(1 - c_{1,j}) \leq \beta, \quad j = 1, \dots, N, \quad (43)$$

$$u^\top r^j \geq y_1, \quad j = 1, \dots, N, \quad (44)$$

$$\sum_{j=1}^N q_j a_{l,j} \leq \sum_{k=1}^D p_k (\beta - y_k)_+, \quad (45)$$

$$y_i - u^\top r^j \leq a_{i,j}, \quad j = 1, \dots, N, \quad i = l+1, \dots, D, \quad (46)$$

$$\beta - u^\top r^j \leq a_{l,j}, \quad j = 1, \dots, N, \quad (47)$$

$$(15), (20), (22), (26), (37-38). \quad (48)$$

The reformulation leads to a mixed-integer linear programming problem, which can be efficiently solved by commercial solvers as in Cplex, Xpress or Gurobi. From (20), we can see that the number of binary variables is $N \times (l+2)$. Hence, the tractability of the portfolio selection problem with ISD constraints depends on the sample size as well as the reference point. Alternatively to the mixed-integer programming approach, the cutting plane method also provides an efficient solution method for the FSD, SSD, and thus ISD constrained programming problems. One may refer to a series of works on the cutting plane approach, see, for example, Dentcheva and Ruszczyński (2010b), Sun et al. (2013) and Noyan and Rudolf (2018).

As for the portfolio selection model ISD-2 with the same asset universe and investment horizon, We assume that the distribution of r is discrete with samples $r^1, \dots, r^N$ and probabilities $q_j = \frac{1}{N}$, $j = 1, \dots, N$. We use the outer approximation method (28–33) proposed in

Sect. 4.2 to approximate the ISD constraint, which leads to a convex quadratic programming problem.

6 Computations

In this section, we develop an extensive set of computational experiments to a range of possible portfolio models with in-sample and out-of-sample analyses applied to the US equity market. After summarizing the adopted data-sets and experimental set-up we consider the results generated over the 2007–2021 period by those portfolio selection models in terms of model consistency and actual financial performance when assuming a 1-week investment horizon. In this computational section, we are primarily interested to analyze the implications of an ISD-based approach on the solution of a classical portfolio problem. Accordingly, the problem formulations (35) under (36–38) are kept very simple to focus and isolate the impact of ISD-1, 2 constraints on the optimal solutions.

The overall case study aims at clarifying the implications of adopting an ISD-based decision model relative to existing approaches and their specific properties.

6.1 Data sets, experimental set-up and portfolio selection models

We consider an investment universe with 9 *exchange-traded-funds* (ETF) in the US equity market corresponding to different industry sectors and 1 risk-free asset. The risky ETFs are Utilities Select Sector SPDR ETF (XLU), Energy Select Sector SPDR ETF (XLE), Financial Select Sector SPDR ETF (XLF), Technology Select Sector SPDR ETF (XLK), Health Care Select Sector SPDR ETF (XLV), Consumer Staples Select Sector SPDR ETF (XLP), Consumer Discretionary Select Sector SPDR ETF (XLY), Industrial Select Sector SPDR ETF (XLI), and Materials Select Sector SPDR ETF (XLB). We adopt the S&P500 as the market index. In (36) we set y as the weekly return rate of the S&P500 index. The risk-free asset is assumed to be a cash account carrying a null interest rate. We collect weekly data of S&P500 and the sector indices over the period 2007/1/8 - 2021/5/31. All data are downloaded from "<https://finance.yahoo.com/>".

We present in Table 1 a set of descriptive statistics for the S&P500 and the sector indices. In Sect. 6.2 we test the following models.

- Portfolio selection models under first (FSD), second (SSD) and third (TSD) order SD constraints;

Table 1 Statistics of return rates of S&P500 and sub-indices in 2007/1/8–2021/5/31, weekly data

	S&P500	XLU	XLE	XLF	XLK	XLV	XLP	XLY	XLI	XLB
Mean (%)	0.190	0.169	0.081	0.204	0.311	0.247	0.204	0.319	0.235	0.207
Max (%)	12.102	17.455	17.755	32.524	10.558	9.276	6.171	18.332	14.572	20.676
Min (%)	−18.196	−19.801	−25.185	−23.959	−15.168	−18.584	−13.331	−14.721	−18.335	−14.939
Std. (%)	2.628	2.623	4.027	4.275	2.785	2.428	1.954	3.087	3.184	3.356
Skewness	−0.716	−0.760	−0.675	0.630	−0.470	−1.070	−1.145	0.060	−0.373	−0.069
Kurtosis	10.438	16.066	9.386	15.819	6.834	10.682	9.413	9.165	7.984	7.762
Sharpe ratio	0.072	0.064	0.020	0.048	0.112	0.102	0.104	0.103	0.074	0.062

- ISD-1 and ISD-2 for $l = 1, 13, 26, 39, 51$ to span, over 52 weeks in a year, different reference points; and ISD-1 for l determined according to (34).

Table 2 summarizes the above problems' specification with an indication of the associated solvers. We test ISD-1 and ISD-2 for different values of the reference point β , where β is determined as the $l/52$ -th quantile of the historical sample.

With weekly data, we will have 52 observations every year. Thus, the l -th-smallest sample is equal to the $l/52$ -th quantile, which is the reference sample (reference point). By relying on the notation introduced at the end of Sect. 3.2, we denote with ISD- $k.\alpha$, the ISD- k case with $\beta = (1 - \alpha)$ quantile. Here, $\alpha = 1 - l/52$. Then, for $l = 1$ we consider the ISD- $k.9908$ problem (close to the k -th-order SD problem); for $l = 51$ the ISD- $k.0192$ problem (close to the $(k + 1)$ -th order SD problem) and for $l = 13, l = 26, l = 39$, we have the ISD- $k.75$, ISD- $k.5$ and ISD- $k.25$ problems. Under this convention, we see that we are approximating a continuous ordering scheme between traditional integer-order SD. Then FSD corresponds to ISD-1.0, SSD to ISD-2.0, TSD to ISD-3.0, and ISD- k with different l values will span the interval between the integer orders. For the case l determined according (34) we will find the strongest ISD-constrained problem. We denote by ISD-1.min, the ISD-1 problem with maximal l or equivalently minimal α . The relationship between l and ISD- $k.\alpha$ can be found in Table 2.

In the out-of-sample analysis, we investigate the performance of the models with ISD constraints against a set of benchmark models by examining their out-of-sample performance over the 2008–2021 period. To this aim we develop a back-testing analysis through a rolling window procedure based on 1-year (52 weeks) in-sample data and 1 week out-of-sample validation. Starting from a 2007/1/8–2007/12/31 history, we derive the optimal portfolios and analyze their performance over the following week using market data and then move the entire data window one week forward. Eventually we have 698 out-of-sample results, spanning from 2008/1/7 to 2021/5/31.

6.2 Model validation

We intend to validate the introduced ISD paradigm by analyzing the diversification properties of optimal portfolios generated by alternative ISD models relative to a set of current portfolio approaches and by comparing a range of ISD-1. α and ISD-2. α optimal portfolios for different orders (different values of l -th-quantile as β). The aim of this comparison is to discriminate between ISD-1. α or ISD-2. α optimal allocations and analyze their consistency against canonical optimal FSD or SSD portfolios.

The following evidences are presented in this section:

1. The optimization results associated with a single instance of several SD- and ISD-based problem formulations,
2. The resulting optimal portfolios diversification and financial properties,
3. A pairwise comparison of optimal portfolios generated by alternative optimization problems over the 2008–2021 period.

One problem instance

As evidence on portfolio compositions and ISD-constrained portfolios relative to existing approaches, we consider here the week 2018/6/26–2018/7/2, and show the resulting optimal portfolios.

We can find from Table 3 that, in this single case problem:

Table 2 Settings of the tested models

Model	Formulation	Equiv. model	Reference sample	Equivalent order	Solver
TSD	Bawa et al. (1975)*	Convex QP	–	ISD-3,0	SDPT3**
ISD-2	(39) s.t.(28-33)	Convex QP	$l = 1$ *****	ISD-2,9808*****	SDPT3**
			$l = 13$	ISD-2,75	
			$l = 26$	ISD-2,5	
			$l = 39$	ISD-2,25	
			$l = 51$	ISD-2,0192	
			–	ISD-2,0	SDPT3**
SSD	Luedtke (2008)*	LP	$l = 1$	ISD-1,9808	CPLEX***
ISD-1	(39) s.t.(40-48)	MILP	$l = 13$	ISD-1,75	
			$l = 26$	ISD-1,5	
			$l = 39$	ISD-1,25	
			$l = 51$	ISD-1,0192	
			Max- l determined by (34)	ISD-1.min	
FSD	Luedtke (2008)*	MILP	–	ISD-1,0	CPLEX***

*The model is with transaction constraints (37–38).

**SDPT3 solver is called by CVX v2.1 in MATLAB R2016a in a PC with 3.4GHz CPU and 16.0GB memory.

***CPLEX v12.8 is run in MATLAB R2016a in a PC with 3.4GHz CPU and 16.0GB memory.

****Corresponding β is the l th-smallest historical sample of S&P500.*****Equivalent order is ISD- k, α where $\alpha = 1 - l/52$

Table 3 Optimal portfolios for the week 2018/6/26–2018/7/2, in-sample data over 2017/7/3–2018/6/25

Model	XLU	XLE	XLF	XLK	XLV	XLP	XLX	XLX	XLX	Risk-free	Opt.val. (%)	HHI	Entropy
TSD (ISD-3.0)	0	0	0	0.252	0	0	0.584	0	0	0.164	0.375***	0.432	0.958
ISD-2.9808	0	0	0	0.252	0	0	0.584	0	0	0.164	0.375	0.432	0.958
ISD-2.9423	0	0	0	0.231	0	0	0.609	0	0	0.160	0.375	0.450	0.933
ISD-2.9038	0	0.003	0	0.295	0	0	0.522	0	0	0.180	0.372	0.392	1.025
ISD-2.8077	0	0.008	0	0.313	0	0	0.492	0	0	0.187	0.371	0.375	1.065
ISD-2.75 ~ ISD-2.0	0	0.009	0	0.313	0	0	0.492	0	0	0.186	0.371	0.375	1.068
SSD (ISD-2.0)	0	0.009	0	0.313	0	0	0.492	0	0	0.186	0.371	0.375	1.068
ISD-1.9808	0	0.009	0	0.313	0	0	0.492	0	0	0.186	0.371	0.375	1.068
ISD-1.9038	0	0	0	0.276	0	0	0.503	0	0	0.221	0.353	0.378	1.035
ISD-1.8077	0.004	0.289	0	0.175	0	0	0.297	0	0	0.234	0.331	0.258	1.387
ISD-1.75 ~ ISD-1.2885*	0.024	0.284	0	0.184	0	0	0.276	0	0	0.233	0.327	0.245	1.453
ISD-1.2693	0.030	0.284	0	0.175	0.009	0.014	0.264	0	0	0.225	0.318	0.233	1.555
ISD-1.25 (ISD-1.min)	0.031	0.285	0	0.165	0.018	0.029	0.252	0	0	0.220	0.309	0.223	1.618
≤ ISD-1.2308	Inf**												
FSD (ISD-1.0)	Inf**												

* A row with multi parameters means the optimal solutions are the same under these parameters;

** Inf means that the optimization problem is infeasible;

***Opt.val. (%) of TSD and ISD-2.α are lower bound of their optimal values

- SSD is feasible but the problem with FSD constraints cannot be solved to optimality and it is infeasible. As a compromise of SSD and FSD, ISD-1. α is feasible when the order 1. α is larger than or equal to 1.25 (equivalently β is less than or equal to 0.75-quantile) and infeasible with order smaller than or equal to 1.2308 (equivalently β larger than or equal to 0.7692-quantile). The optimal value of ISD-1. α is always smaller than or equal to SSD because FSD constraints are stronger than SSD constraints.
- When β is small enough, say, 0.0192-quantile, both the optimal portfolio and the optimal value of ISD-1.9018 are the same as those of SSD. The smaller (larger) the order (β), the smaller the optimal value of ISD-1. Thus the smaller the order, the stronger the ISD constraint.
- Similarly: the optimal ISD-2.9808 portfolio coincides with TSD portfolios and again that SSD is equivalent to ISD-2.0192 and ISD-1.9808. As the order decreases, ISD-2. α constraints have limited impact on portfolio composition.
- Based on the decreasing optimal value in Column 12, we can infer that $\text{TSD} \subset \text{ISD-2.}\alpha$ (larger order) $\subset \text{ISD-2.}\alpha$ (smaller order) $\subset \text{SSD} \subset \text{ISD-1.}\alpha$ (larger order) $\subset \text{ISD-1.}\alpha$ (smaller order) $\subset \text{FSD}$, confirming an effective spanning of risk profiles based on ISD models.
- When the order is close to 2, the optimal portfolio of ISD-1. α invests in only two or three risky assets, just like SSD. When the order is closer to 1, the optimal portfolio invests in four or six risky assets, almost half of the stock pool.
- The largest proportion of one risky asset is below 30%. It means that, when the order is smaller (benchmark is large), the optimal portfolio is further diversified to meet the stricter dominance constraint. While, for all SD based models, the largest proportion of one risky asset would not exceed a half. This evidence confirms the flexibility of the ISD approach.

The last two columns of Table 3 report the Herfindahl-Hirschman Index (HHI) and Shannon entropy (Entropy) of the associated portfolios. HHI was originally proposed as an indicator of firms' competition and then extended to finance theory and now a commonly accepted measure of portfolio concentration. It is calculated as $H = \sum_{i=1}^N s_i^2$, where s_i is the share of each asset in the portfolio. It ranges from 0 to 1.0, moving from a theoretically infinitely-diversified to a fully concentrated portfolio with 1 single asset in the portfolio. The Shannon entropy is $S = -\sum_i s_i \log s_i$. The more unequal the investment proportions of the portfolios, the larger the weighted geometric mean of the portfolio proportions and the smaller the corresponding Shannon entropy. When there is only one risky asset invested in the portfolio, Shannon entropy exactly equals zero. In Table 4 we extend the diversification analysis to the 2008–2021 test period.

Portfolio diversification

Table 4 shows the average values of the two concentration indices. Rows 5–6 in Table 4 show the number of weeks in which these portfolio problems are infeasible, as well as the percentage of infeasible weeks. When an ISD- $k.\alpha$ portfolio problem is feasible, the resulting portfolio dominates the benchmark in the ISD- $k.\alpha$ sense with respect to 52 in-sample data. When a portfolio problem is infeasible, the resulting portfolio will be concentrated in the risk-free asset during that week. As the infeasibility will affect the concentration, we compute two kinds of concentration indices. One reports the average HHI and Entropy of all portfolios over 2008/1/7–2021/5/31 (rows 3 and 4). The other shows average HHI and Entropy values only when the problems are solved to optimality (rows 7 and 8).

We summarize the following evidences from Table 4:

- The FSD-constrained problem is infeasible in 667 out of the 698 weeks. When feasible, the resulting optimal portfolio is highly diversified, but in practice, this is not a viable model. Similar evidence for the ISD-1.25 and ISD-1.5 models with infeasibility rates that make these models hardly practical. The model with minimal order (maximal- l) selection proves to overcome the problem and among all optimal portfolios also has the associated higher diversification on average.
- ISD-1.9808, SSD and ISD-2.0192 show very similar statistics and relatively good diversification properties: we show below that indeed the optimal portfolios associated with these problems are in general very close to each other, as expected.
- For decreasing order: ISD-2. α and ISD-1. α optimal portfolios do actually span respectively from TSD to SSD and from SSD to FSD optimal portfolios and provide evidence of the increasing difficulty to solve to optimality the associated portfolio problem.

Table 4 Average values of Herfindahl-Hirschman Index (HHI), Shannon entropy (Entropy), proportions invested in risk-free asset, numbers and percentages of infeasible weeks, of out-of-sample weekly portfolios over 2008/1/7–2021/5/31 (ISD-1.min is re-run in each week to find the maximal- l)

Order	TSD			ISD-2. α			SSD
	3.0	2.9808	2.75	2.5	2.25	2.0192	2.0
HHI*	0.748	0.747	0.729	0.717	0.716	0.728	0.731
Entropy*	0.402	0.404	0.436	0.460	0.461	0.438	0.433
Inf. Weeks***	5	5	5	5	5	19	20
Inf. Rate	0.72%	0.72%	0.72%	0.72%	0.72%	2.72%	2.87%
Feas. HHI**	0.746	0.745	0.727	0.715	0.714	0.720	0.723
Feas. Entropy**	0.404	0.407	0.439	0.464	0.464	0.450	0.446
RF	11.31%	11.31%	11.84%	12.10%	12.18%	13.57%	13.73%
Feas. RF**	10.67%	10.67%	11.20%	11.47%	11.55%	11.15%	11.19%
Order	ISD-1. α						FSD
	1.9808	1.75	1.5	1.25	1.0192	1.min	1.0
HHI*	0.737	0.621	0.654	0.776	0.958	0.543	0.9727
Entropy*	0.420	0.661	0.632	0.416	0.077	0.882	0.0512
Inf. Weeks***	20	41	179	403	640	20	667
Inf. Rate	2.87%	5.87%	25.64%	57.74%	91.69%	2.87%	95.56%
Feas. HHI**	0.730	0.597	0.535	0.469	0.496	0.530	0.386
Feas. Entropy**	0.432	0.702	0.850	0.985	0.927	0.908	1.153
RF	13.67%	18.57%	37.68%	62.96%	92.37%	31.02%	96.08%
Feas. RF**	11.12%	13.48%	16.18%	12.37%	8.14%	28.99%	11.79%

*HHI and Entropy give the average HHI and Entropy values of all optimal weekly portfolios over 2008/1/7–2021/5/31.

When a model is infeasible, the portfolio is all in risk-free asset and the HHI is 1 at that week.

**Feas.HHI, Feas.Entropy and Feas.RF account the average HHI, Entropy and RF values when the model is feasible.

***Inf. accounts when solver CPLEX returns '-2: No feasible point was found' or SDPT3 returns 'Failed' or 'Infeasible'

Pairwise optimal portfolios comparison

Each matrix value here next reflects the proportion of weeks in which the optimal portfolios of the two models in the row and in the column agree, in the sense that the norm of the difference of the two portfolios is less than $\epsilon = 0.001$.

We have adopted a graphical rather than a tabular representation to highlight in color the similarities across the many tested models. We see that indeed by decreasing order for ISD-1. α and ISD-2. α we span from FSD to SSD and TSD portfolios but also that, possibly due to the adopted outer approximation scheme for ISD-2. α the resulting optimal portfolio compositions do not differ significantly in the SSD-TSD range. For ISD-1.0192,1.9808,2.0192,2.9808, we see that indeed the associated optimal portfolios are within the same color sets.

From Fig. 5 we can confirm the following similarities, that will delimit the discussion below:

- Optimal TSD portfolios can be regarded as equivalent to ISD-2.9808 optimal portfolios. Same with respect to SSD-constrained portfolios considering ISD-2.0192 or ISD-1.9808.
- ISD-2.25,2.5,2.75 optimal portfolios show similar composition with negligible differences.
- ISD-1. α optimal portfolios provide for decreasing order on average different optimal portfolios when solved to optimality. The ISD-1. α optimal portfolios with minimal order have their own specific structure.
- For $k = 1, 2$, ISD- k . α formulations for increasing order and thus weaker feasibility conditions, lead to increasingly similar portfolio structures.

The key outcome of this model validation analysis is that ISD-based portfolio models do actually enrich currently available SD-based approaches. In both the ISD-1. α and ISD-2. α classes we can find sure alternative and distinctive investors' risk profiles. In the following performance analysis we will focus in particular on portfolio models surely leading to alternative investment strategies.

	ISD-3.0 (TSD)	ISD-2.9808	ISD-2.75	ISD-2.5	ISD-2.25	ISD-2.0192	ISD-2.0 (SSD)	ISD-1.9808	ISD-1.75	ISD-1.5	ISD-1.25	ISD-1.0192	ISD-1.min	ISD-1.0 (FSD)
ISD-3.0 (TSD)	1.00	0.96	0.59	0.55	0.55	0.56	0.56	0.60	0.30	0.16	0.06	0.06	0.03	0.07
ISD-2.9808	0.96	1.00	0.59	0.55	0.55	0.55	0.55	0.60	0.30	0.16	0.06	0.06	0.03	0.07
ISD-2.75	0.59	0.59	1.00	0.62	0.62	0.62	0.62	0.65	0.29	0.16	0.06	0.06	0.03	0.07
ISD-2.5	0.55	0.55	0.62	1.00	0.71	0.68	0.69	0.74	0.28	0.16	0.06	0.06	0.04	0.07
ISD-2.25	0.55	0.55	0.62	0.71	1.00	0.75	0.75	0.79	0.28	0.16	0.06	0.06	0.04	0.07
ISD-2.0192	0.56	0.55	0.62	0.68	0.75	1.00	0.78	0.82	0.31	0.18	0.08	0.08	0.06	0.09
ISD-2.0 (SSD)	0.56	0.55	0.62	0.69	0.75	0.78	1.00	0.84	0.31	0.18	0.08	0.08	0.06	0.09
ISD-1.9808	0.60	0.60	0.65	0.74	0.79	0.82	0.84	1.00	0.33	0.19	0.08	0.08	0.07	0.09
ISD-1.75	0.30	0.30	0.29	0.28	0.28	0.31	0.31	0.33	1.00	0.36	0.18	0.12	0.09	0.13
ISD-1.5	0.16	0.16	0.16	0.16	0.16	0.18	0.18	0.19	0.36	1.00	0.45	0.33	0.17	0.33
ISD-1.25	0.06	0.06	0.06	0.06	0.06	0.08	0.08	0.08	0.18	0.45	1.00	0.59	0.25	0.59
ISD-1.0192	0.06	0.06	0.06	0.06	0.06	0.08	0.08	0.08	0.12	0.33	0.59	1.00	0.24	0.95
ISD-1.min	0.03	0.03	0.03	0.04	0.04	0.06	0.06	0.07	0.09	0.17	0.25	0.24	1.00	0.21
ISD-1.0 (FSD)	0.07	0.07	0.07	0.07	0.07	0.09	0.09	0.09	0.13	0.33	0.59	0.95	0.21	1.00

Fig. 5 Correlations between different models (proportion of weeks in which the optimal portfolios of two models are the same)

6.3 Out-of-sample evidence

Following the analysis in Sect. 6.2 we limit the study to the benchmark and those portfolio models satisfying a sufficient diversification principle and that will surely generate different optimal portfolio allocations, namely: ISD-3.0(TSD)-, ISD-2.75-, ISD-2.0(SSD)-, ISD-1.75- and ISD-1.min- constrained portfolios. The following statistics are computed over 698 weeks and shown in Table 5:

- Mean, standard deviations and Sharpe ratios of weekly returns in percentage. We use the standard deviation of realized returns to analyse the effectiveness of volatility control.
- Expected shortfalls (CVaR) over the least 5% and 10% weekly returns. These are for assessing tail risk control.
- The number (proportion) of out-of-sample weeks in which the ex-post return exceeds the S&P500;
- The weekly average excess returns above and below the S&P500, denoted respectively with $E(ER)_+$ and $E(ER)_-$;

Table 5 provides a synthetic evidence of the optimal portfolios' average performance over the test period.

There are significant differences in terms of out-of-sample performances and we see that in general all optimal portfolios do provide an effective volatility control relative to the S&P500 leading to higher risk-adjusted returns. The best out-of-sample performance is associated with the ISD-1.min optimal portfolio when the reference point (order) is determined through the bisection method. The $E(ER)_-$ evidence is also interesting: it reflects the penalty in case of benchmark return underachievement and thus how good the index tracking model works. From Table 5 we can confirm that ISD-1.min performs best: the proposed stochastic dominance constraint is effective as to meeting the tracking requirement. The test-period includes the financial crisis of 2007-2008 as well as the COVID-19 recession in 2020, with substantial losses in specific weeks therein. The out-of-sample CVaR statistics illustrate that the stronger the SD constraint, the less extreme the out-of-sample losses faced by the portfolio manager. While the maximal β strategy, by adapting to a changing market environment, will be very effective specifically for tail risk control.

We present in Figs. 6 and 7 the cumulative distributions of out-of-sample weekly returns for the above models. The cumulative distributions provide further evidence of the different risk-return profiles generated by each model and the ex-post effectiveness of the imposed stochastic dominance constraints. These must be considered together with the postulated S&P500 index tracking problem. A rational investor will in general prefer jointly lower underperformance probabilities and higher overperformance probabilities. In Fig. 7 we see that ISD-1.min does dominate the other distributions but has a very high probability of ex-post null returns.

It is worth recalling that we are only taking into consideration static 1-week investment horizon problems based on an optimal index-tracking criterion and the Figures display only ex-post actual market performances. In Fig. 6 we see that the displayed ISD-1.min distribution dominates all other distributions on the left tail (strictly negative returns) but carries the highest probability of non-positive returns and it is dominated on the positive tail. In Fig. 7 the three models adopted do generate similar ex-post distributions.

Table 5 Out-of-sample weekly returns statistics over 2008/1/7–2021/5/31

Model	Mean (%)	Std. (%)	Sharpe ratio	CVaR _{95%} (%)	CVaR _{90%} (%)	Win. weeks (Win. rate %)	E(ER) ₊ (%)	E(ER) _− (%)
S&P500	0.201	2.619	0.077	− 6.533	− 4.860	–	0	0
ISD-3.0(TSD)	0.198	2.252	0.088	− 5.571	− 4.130	354 (50.72)	0.564	− 0.567
ISD-2.75	0.198	2.232	0.089	− 5.544	− 4.103	358 (51.29)	0.559	− 0.562
ISD-2.0(SSD)	0.197	2.194	0.090	− 5.415	− 3.989	360 (51.58)	0.571	− 0.575
ISD-1.75	0.145	2.031	0.071	− 5.220	− 3.821	361 (51.72)	0.523	− 0.579
ISD-1.min	0.189	1.963	0.096	− 4.798	− 3.515	360 (51.58)	0.497	− 0.509

Fig. 6 Cumulative distribution of out-of-sample portfolio return rates over 2008/1/7–2021/5/31 of ISD-2.0(SSD), ISD-1.75 and ISD-1.min models

Fig. 7 Cumulative distribution of out-of-sample portfolio return rates over 2008/1/7–2021/5/31 of ISD-2.0(SSD), ISD-2.75 and ISD-3.0(TSD) models

6.4 Dropping the constraints above β

Following Remark 2 just before Sect. 3.1, we analyse the implications of dropping the ISD-constraints above β . We take ISD-1 as an example and avoid including the SSD constraints (8) for $\eta \geq \beta$, while constraints (7) remain active. In Theorem 1, (7) and (9–12) have been shown to be equivalent. With N finite samples, furthermore (7) is equivalent to (15–21) with active constraints only from the l -th to the N -th sample. By dropping such constraint set, we generate the desired reformulation of the portfolio selection model (35–38). Notice that we also attain a further relaxation of FSD over ISD-1, since:

$$W \succeq_{(1)} Y \Rightarrow W \succeq_{(\beta,1)} Y \Rightarrow \text{MISD-1} \quad (7),$$

where (7) corresponds to the modified ISD-1 (MISD-1) model without constraints above β . We see that indeed an effective implementation of the ISD models requires both conditions, left and right of β , to hold and to be employed in the optimization problem.

Fig. 8 Cumulative distribution of out-of-sample portfolio return rates over 2008/1/7–2021/5/31 of MISD-1.9808, MISD-1.75, MISD-1.5, MISD-1.25 models and S&P500

We develop this comparative analysis, relying on the very same set-up and data history of Sect. 6.1, and determine the out-of-sample performance of ISD-1 and MISD-1 models for different β . Table 6 shows a set of statistics of realized returns for different β over the 698 out-of-sample weeks. The third column shows the number of weeks in which the tested model is infeasible, according to the adopted solution approach (CPLEX MILP, see Table 2). The fourth column shows the number (proportion) of weeks in which the optimal portfolios of ISD-1 and MISD-1 models agree, in the sense that the norm of the difference between the two portfolios is less than $\epsilon = 0.001$. Finally, we investigate to which extent the 698 ISD-1 out-of-sample portfolio returns exceed the corresponding S&P500 returns. Figure 8 plots the cumulative distributions of realized weekly returns for the MISD-1 models. We find from Figs. 6, 7 and 8 that, out-of-sample, ISD-1 (or MISD-1) portfolios do not ISD-1 dominate S&P500 return. However, when setting $\beta = 0$, any ISD/MISD portfolio, but MISD-1.9808, leads to a realized return greater than the S&P500 return. In the last column of Table 6 we show the maximal reference point β for which out-of-sample portfolio returns could dominate S&P500 return in the MISD-2 sense.

From Table 6 we find that:

- When dropping the constraints above β (α -quantile), as the reference point increases ISD-1 active constraints become increasingly binding.
- The two approaches generate ex-post very similar risk control evidences. In most cases, the ISD constraints are more effective on the left tail than on the right tail. Hence, dropping the right tail would keep a similar risk exposure.
- In terms of ex-post performance a full implementation of ISD-1 conditions has a limited but not negligible positive impact on portfolio performance.
- Except for MISD-1.9808, most out-of-sample portfolios could MISD-2 dominate S&P500 over a large portion of the support, confirming the robustness of ex-post performance.
- The MISD model fails to achieve an efficient volatility control with a small β , as it only keeps constraints below β but drops all others above.

Table 6 Out-of-sample weekly returns statistics over 2008/1/7–2021/5/31

Model	k, α	Inf. weeks (Prop.)	Same. weeks (Prop.)	Mean (%)	Std. (%)	Sharpe ratio	$E(ER)_+$ (%)	$E(ER)_-$ (%)	$\max \beta (P(X \leq \beta))$ s.t. MISD-2 S&P500
S&P500	–	–	–	0.201	2.619	0.077	0	0	–
ISD-1 (full support)	ISD-1.9808	20 (2.87%)	–	0.190	2.206	0.086	0.569	–0.580	0.0402 (95.70%)
	ISD-1.75	41 (5.87%)	–	0.145	2.031	0.071	0.523	–0.579	0.0360 (94.41%)
	ISD-1.5	179 (25.64%)	–	0.096	1.791	0.054	0.571	–0.676	0.0326 (93.41%)
	ISD-1.25	403 (57.74%)	–	0.011	1.636	0.007	0.593	–0.783	0.0322 (93.12%)
MISD-1 (dropping right)	MISD-1.9808	0 (0.00%)	225 (32.23%)	0.248	2.667	0.093	0.690	–0.644	–0.0940 (0.42%)
	MISD-1.75	0 (0.00%)	590 (84.53%)	0.174	2.062	0.085	0.514	–0.540	0.0441 (96.56%)
	MISD-1.5	43 (6.16%)	467 (66.91%)	0.126	1.839	0.069	0.554	–0.629	0.0362 (94.70%)
	MISD-1.25	335 (47.99%)	559 (80.09%)	0.029	1.715	0.017	0.555	–0.727	0.0310 (92.26%)

7 Conclusion

The main contribution of this article is related to the extension of the canonical partial order over probability measures induced by stochastic dominance (SD) to discriminate between random outcomes involving two subsequent SD relationships: to this aim the concept of *k*th-order interval-based stochastic dominance (ISD-*k*) has been introduced. By choosing different reference points, ISD can reflect different degrees of risk preferences as a continuum between *k*th- and (*k* + 1)th-order SD.

From a modeling and decision theoretical perspective, the introduced paradigm has been clarified with respect to utility theory and optimal tail risk control. For *k* = 1 we have shown in particular that ISD-1 allows the order relationship to hold for all monotone increasing utilities which on the left part of the support may be concave or convex and surely concave beyond the ISD reference point. For *k* = 2, ISD-2 was shown to lead to a new risk measure, the interval-based Conditional Value-at-Risk, whose relationship with the VaR and the CVaR was analyzed in the paper.

To clarify the practicality and operational relevance of the introduced optimization framework we developed for the case of discrete random variables, a computationally efficient reformulation of the ISD-*k* constraints leading to convex programs that can be solved efficiently for the cases of first- and second-order ISD. Of particular interest is the formulation of a surely feasible ISD-1.α optimization problem closest in terms of risk-preferences to an otherwise infeasible FSD problem.

As an application domain, we have considered a portfolio selection problem with ISD-*k* constraints, for *k* = 1, 2. The associated portfolio models do contribute to already established portfolio optimization models. An extended in- and out-of-sample validation has been conducted with an application to the US equity market. The collected evidences confirm the potentials of the introduced approach for decision-making under uncertainty and specifically in portfolio theory.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Armbruster, B., & Delage, E. (2015). Decision making under uncertainty when preference information is incomplete. *Management Science*, 61(1), 111–128.
- Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica*, 55(4), 749–764.
- Barro, D., Canestrelli, E., & Consigli, G. (2018). Volatility versus downside risk: Performance protection in dynamic portfolio strategies. *Computational Management Science*, 23, 1–47.
- Baucells, M., & Heukamp, F. (2006). Stochastic dominance and cumulative prospect theory. *Management Science*, 52(9), 1409–1423.
- Baucells, M., & Sarin, R. K. (2010). Predicting utility under satiation and habit formation. *Management Science*, 56(2), 286–301.
- Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2, 95–121.
- Bawa, V. S., Bodurtha, J. N., Jr., Rao, M. R., & Suri, H. L. (1985). On determination of stochastic dominance optimal sets. *Journal of Finance*, 40(2), 417–431.
- Bonami, P., Lodi, A., Tramontani, A., & Wiese, S. (2015). On mathematical programming with indicator constraints. *Mathematical Programming Series B*, 151(1), 191–223.
- Chen, Z., & Jiang, J. (2018). Stability analysis of optimization problems with *k*th order stochastic and distributionally robust dominance constraints induced by full random recourse. *SIAM Journal on Optimization*, 28(2), 1396–1419.

- Consigli, G., Moriggia, V., & Vitali, S. (2019). Long-term individual financial planning under stochastic dominance constraints. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-019-03253-8>.
- Dentcheva, D., & Ruszczyński, A. (2003). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization*, 14(2), 548–566.
- Dentcheva, D., & Ruszczyński, A. (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance*, 30(2), 433–451.
- Dentcheva, D., & Ruszczyński, A. (2010a). Robust stochastic dominance and its application to risk-averse optimization. *Mathematical Programming*, 123(1), 85–100.
- Dentcheva, D., & Ruszczyński, A. (2010b). Inverse cutting plane methods for optimization problems with second-order stochastic dominance constraints. *Optimization*, 59(3), 323–338.
- Dupačová, J., & Kopa, M. (2014). Robustness of optimal portfolios under risk and stochastic dominance constraints. *European Journal of Operational Research*, 234(2), 434–441.
- Fang, Y., & Post, T. (2017). Higher-degree stochastic dominance optimality and efficiency. *European Journal of Operational Research*, 261, 984–993.
- Fishburn, P. C. (1964). *Decision and Value Theory*. New York: Wiley.
- Fishburn, P. C. (1980). Continua of stochastic dominance relations for unbounded probability distributions. *Journal of Mathematical Economics*, 7(3), 271–285.
- Gao, J. J., Xiong, Y., & Li, D. (2016). Dynamic mean-risk portfolio selection with multiple risk measures in continuous-time. *European Journal of Operational Research*, 249, 647–656.
- Gollmer, R., Neise, F., & Schultz, R. (2008). Stochastic programs with first-order dominance constraints induced by mixed-integer linear recourse. *SIAM Journal on Optimization*, 19, 552–571.
- Hadar, J., & Russell, W. R. (1969). Rules for ordering uncertain prospects. *American Economic Review*, 59(1), 25–34.
- Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *Review of Economic Studies*, 36, 335–346.
- Hu, J., & Stepanyan, G. (2017). Optimization with reference-based robust preference constraints. *SIAM Journal on Optimization*, 27(4), 2230–2257.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263–291.
- Kallio, M., & Hardoroudi, N. D. (2019). Advancements in stochastic dominance efficiency tests. *European Journal of Operational Research*, 276(2), 790–794.
- Killick, T. (1995). Structural adjustment and poverty alleviation: An interpretative survey. *Development and Change*, 26(2), 305–330.
- Kopa, M., Moriggia, V., & Vitali, S. (2018). Individual optimal pension allocation under stochastic dominance constraints. *Annals of Operations Research*, 260(1), 255–291.
- Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science*, 50, 1390–1406.
- Leshno, M., & Levy, H. (2002). Preferred by all and preferred by most decision makers: Almost stochastic dominance. *Management Science*, 48(8), 1074–1085.
- Levy, H. (2015). *Stochastic dominance: Investment decision making under uncertainty*. Berlin: Springer.
- Levy, M., & Levy, H. (2002). Prospect theory: Much ado about nothing? *Management Science*, 48(10), 1334–1349.
- Lizyayev, A., & Ruszczyński, A. (2012). Tractable almost stochastic dominance. *European Journal of Operational Research*, 218(2), 448–455.
- Luedtke, J. (2008). New formulations for optimization under stochastic dominance constraints. *SIAM Journal on Optimization*, 19(3), 1433–1450.
- Moriggia, V., Kopa, M., & Vitali, S. (2019). Pension fund management with hedging derivatives, stochastic dominance and nodal contamination. *Omega*, 87, 127–141.
- Müller, A., Scarsini, M., Tsetlin, I., & Winkler, R. L. (2016). Between first- and second-order stochastic dominance. *Management Science*, 63(9), 2933–2947.
- Müller, A., & Stoyan, D. (2002). *Comparison methods for stochastic models and risks*. Chichester: Wiley.
- Noyan, N., & Rudolf, G. (2018). Optimization with stochastic preferences based on a general class of scalarization functions. *Operations Research*, 66, 463–486.
- Noyan, N., Rudolf, G., & Ruszczyński, A. (2006). Relaxations of linear programming problems with first order stochastic dominance constraints. *Operations Research Letters*, 34, 653–659.
- Ogryczak, W., & Ruszczyński, A. (2001). On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming*, 89, 217–232.
- Porter, R. B. (1974). Semivariance and stochastic dominance: A comparison. *American Economic Review*, 64, 200–204.
- Post, T. (2003). Empirical tests for stochastic dominance efficiency. *The Journal of Finance*, 58, 1905–1931.

- Post, T., Fang, Y., & Kopa, M. (2015). Linear tests for decreasing absolute risk aversion stochastic dominance. *Management Science*, 61, 1615–1629.
- Post, T., & Kopa, M. (2016). Portfolio choice based on third-degree stochastic dominance. *Management Science*, 63(10), 3381–3392.
- Quirk, J. P., & Saposnik, R. (1962). Admissibility and measurable utility functions. *The Review of Economic Studies*, 29, 140–146.
- Roman, D., Darby-Dowman, K., & Mitra, G. (2007). Mean-risk models using two risk measures: A multiobjective approach. *Quantitative Finance*, 7, 443–458.
- Roman, D., Mitra, G., & Zverovich, V. (2013). Enhanced indexation based on second-order stochastic dominance. *European Journal of Operational Research*, 228(1), 273–281.
- Satchell, S. E., Damant, D. C., & Hwang, S. (2000). Exponential risk measure with application to UK asset allocation. *Applied Mathematical Finance*, 7, 127–152.
- Strub, M. S., & Li, D. (2019). Failing to foresee the updating of the reference point leads to time-inconsistent investment. *Operations Research*. <https://doi.org/10.1287/opre.2019.1872>.
- Sun, H., Xu, H., Meskarian, R., & Wang, Y. (2013). Exact penalization, level function method, and modified cutting-plane method for stochastic programs with second order stochastic dominance constraints. *SIAM Journal on Optimization*, 23(1), 602–631.
- Tsetlin, I., Winkler, R. L., Huang, R. J., & Tzeng, L. Y. (2015). Generalized almost stochastic dominance. *Operations Research*, 63(2), 363–377.
- Yang, X., Gondzio, J., & Grothey, A. (2010). Asset liability management modelling with risk control by stochastic dominance. *Journal of Asset Management*, 11, 73–93.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.