Portfolio Optimization Based on Almost Second-degree Stochastic Dominance

Chunling Luo
Department of Data Science and Management, Alibaba Business School, Hangzhou Normal University, Hangzhou , China, 310030, iseluoc@gmail.com

Piao Chen
ZJUI Institute, Zhejiang University, Haining, China, 314400, piaochen@intl.zju.edu.cn

Patrick Jaillet
Department of Electrical Engineering and Computer Science, Laboratory for Information and Decision Systems, and Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, jaillet@mit.edu

Accepted, *Management Science*, April 2024

In portfolio optimization, the computational complexity of implementing almost stochastic dominance has limited its practical applications. In this study, we introduce an optimization framework aimed at identifying the optimal portfolio that outperforms a specified benchmark under almost second-degree stochastic dominance (ASSD). Our approach involves discretizing the return range and establishing both sufficient and necessary conditions for ASSD. We then propose a three-step iterative procedure: first, identifying a candidate portfolio; second, assessing its optimality; and third, refining the discretization scheme. Theoretical analysis guarantees that the portfolio identified through this iterative process improves with each iteration, ultimately converging to the optimal solution. Our empirical study, utilizing industry portfolios, demonstrates the efficacy of our approach by consistently identifying an optimal portfolio within a few iterations. Furthermore, comparative analysis against other decision criteria, such as mean-variance, second-degree stochastic dominance, and third-degree stochastic dominance, reveals that ASSD generally leads to portfolios with higher out-of-sample average excess returns but also entails increased variations and risks.

Key words: portfolio optimization; almost stochastic dominance; stochastic dominance constraints; quadratically constrained programming; cutting-plane algorithm

1. Introduction

Portfolio optimization based on stochastic dominance (SD) has been studied with growing interest in recent literature (e.g., Dentcheva and Ruszczyński (2006), Roman et al. (2006), Kallio and

Hardoroudi (2018), Post et al. (2018)). Compared with mainstream classical mean-variance (MV) models (Markowitz 1952), SD approaches are theoretically more appealing because of its sound basis for decision making under risk and uncertainty. To be specific, SD is consistent with expected utility theory and can guarantee the preference of one random variable over another across a set of permissible utility functions. For example, a portfolio which dominates another portfolio by second-degree SD (SSD) will be preferred by all non-satiable and risk averse decision makers, that is, individuals whose utility function is non-decreasing and concave ($u'(x) \geq 0, u''(x) \leq 0$). Another advantage of SD is that it can provide a partial ranking of random variables based on partial utility information.

However, it has been recognized by researchers that conventional SD conditions might appear excessively stringent, encompassing certain 'pathological' utility functions that correspond to very few, if any, observed preferences in practical scenarios (Leshno and Levy 2002). To address this limitation, Leshno and Levy (2002) first proposed almost SD (ASD) as a relaxation of conventional SD conditions. In particular, they proposed almost first-degree SD (AFSD) that considers nondecreasing utility functions whose deviation in marginal utilities $u'(x)$ is bounded, and almost second-degree SD (ASSD) that considers risk-averse utility functions whose ratio of second derivative $u''(x)$ is bounded. The condition of ASSD was refined by Tzeng et al. (2013) later.

AFSD and ASSD mark the inception of the ASD field, motivating researchers to refine the set of utility functions associated with traditional SD rules by excluding extreme ones. This modification aims to enable ASD to more effectively discern clear preferences between random variables and explain certain common investment practices that conventional SD conditions cannot capture (Bali et al. 2009, 2013, Levy 2009). ASSD endeavors to achieve this objective by limiting the ratio of $u''(x)$, in line with AFSD. However, some researchers (Luo and Tan 2020, Arvanitis et al. 2021) highlight that this approach still accommodates certain pathological utility functions (e.g., all quadratic utility functions) while potentially excluding realistic utility functions (e.g., logarithmic utility with a wide wealth range). To address these issues, different forms of ASD have been proposed in the literature, considering different features of utility functions. For example, Tsetlin et al. (2015) generalized ASD by simultaneously restricting Nth derivatives of utility functions. Müller et al. (2017) and Huang et al. (2020) relaxed FSD and developed a continuum of SD by bounding the ratio of marginal utilities and the degree of absolute risk aversion, respectively. Luo and Tan (2020) focused on SSD and proposed another specification of ASSD by imposing restrictions on the degree of absolute risk aversion. Liu and Meyer (2021) further provided an alternative form of ASSD by considering utility functions whose absolute or relative risk aversion measure is constrained by both an upper bound and a lower bound. In addition to these works, generalizations of ASD designated for other settings have also emerged in the literature (e.g., Denuit et al. (2014), Lizyayev and Ruszczyński (2012), Tsetlin and Winkler (2018)).

Despite the theoretically enticing characteristics of SD and ASD, their practical implementation in portfolio optimization encounters significant computational complexity. SD encompasses the entire probability distribution of investment returns, leading to computationally demanding approaches. This complexity becomes apparent, especially when compared to the prevalent two-moment MV models that have long dominated the field of portfolio analysis. MV models excel in constructing efficient portfolios through parametric optimization problems. However, the literature lacks effective methods for SD, leaving a gap in the optimization methodologies available for portfolio analysis (Kuosmanen 2004).

In addressing this complexity, various strategies have been proposed in literature. Some researchers have focused on approaches to assess the efficiency or optimality of a specified portfolio. For instance, Bawa et al. (1985) introduced linear programming (LP) algorithms for identifying FSD and SSD optimal sets concerning discrete distributions. Furthermore, several scholars (Post 2003, Kuosmanen 2004, Roman et al. 2006, Kopa and Chovanec 2008, Kopa and Post 2015) have developed LP methodologies to evaluate the stochastic dominance efficiency of given portfolios. The work by Fang and Post (2017) expanded this research by characterizing optimality and efficiency in higher-degree stochastic dominance using LP or convex quadratic programming. Numerous other studies have also delved into similar methodologies, including those by Kopa and Post (2009), Post and Kopa (2013), Post et al. (2015), Longarela (2016), Post and Potì (2017). For a comprehensive exploration of this subject, interested readers can refer to Post (2017). While these works offer methods to assess individual portfolios based on stochastic dominance, constructing an optimal portfolio remains a challenging task, particularly in investment scenarios with an infinite array of feasible portfolios.

Another main stream of research, which addresses the problem of constructing portfolios, is to identify an enhanced portfolio that dominates a given benchmark, that is, a desirable 'reference' portfolio such as the return rate of an index. Early publications in this direction include Dentcheva and Ruszczyński (2003, 2004a,b), which introduced stochastic optimization models involving FSD, SSD and higher-degree SD constraints to guarantee that solutions dominate a benchmark random variable. Later, Dentcheva et al. (2007) and Dentcheva and Römisch (2013) further investigated the stability and sensitivity of SD-enhanced portfolio optimization models. These models typically contain a large number of constraints (Kallio and Hardoroudi 2018) and are difficult to solve. To deal with this issue, a number of approaches have been proposed, such as cutting-plane method (Rudolf and Ruszczyński 2008, Dentcheva and Ruszczyński 2010, Fábián et al. 2011, Sun et al. 2013), sample average approximation method (Hu et al. 2012), primal-dual algorithm (Haskell et al. 2017). Portfolio optimization enhanced by SD has also been specifically studied in the literature. For example, Dentcheva and Ruszczyński (2006) constructed LP models for SSD constrained portfolio optimization problem and analyzed optimality and duality of these models. Post and Kopa (2017) developed an optimization method to identify a portfolio that dominates a benchmark in terms of third-degree SD (TSD). Additional studies within this domain include works by Post et al. (2018), Kallio and Hardoroudi (2018), Liesiö et al. (2020), among others. In this paper, we follow a similar trajectory to explore the portfolio optimization problem based on ASD.

Compared to conventional SD conditions, portfolio optimization based on ASD is even more challenging. ASD has more complex formulation, which requires not only to compare cumulative probability distribution (CDF) but also to quantitatively calculate integrals of the distances between CDFs. Due to these complexities in implementation (Bruni et al. 2017), this area has been relatively underexplored. This paper addresses this gap by introducing a novel approach and devising an optimization method for portfolio selection based on ASSD proposed by Leshno and Levy (2002) and Tzeng et al. (2013). Our contribution lies in the development of mathematical programming models and algorithms aimed at constructing an optimal portfolio that invests in a set of base assets and dominates a benchmark portfolio in terms of ASSD. The primary contributions are summarized below:

- We propose an approximate ASSD (AASSD) condition via approximating the probability distribution of the portfolio return by a mean-preserving spread of the exact one. AASSD has the advantage that it could be determined based on mean, variance and values of first-order lower partial moment (LPM) at a series of discrete points. We show that AASSD is a sufficient condition of ASSD. The gap between AASSD and ASSD could be measured by a closeness index, which is defined to be the half variance difference between the approximate and the true distribution. As we refine the discretization of return range, the closeness index becomes smaller and AASSD gradually approaches ASSD.
- We further develop a three-step iterative method for identifying an optimal portfolio that satisfies the ASSD constraints. To the best of our knowledge, this is the first attempt to do so. Our proposed procedure includes three iterative steps: obtain a candidate portfolio by finding an optimal AASSD-enhanced portfolio; check the optimality of the candidate portfolio based on a necessary condition of ASSD by means of quadratic programming (QP) and quadratically constrained programming (QCP); refine discretization of return range to update and improve the obtained candidate portfolio. We theoretically prove that the identified portfolio improves over iterations and converges to the optimal one. Empirical applications show that our iterative procedure converges in a few iterations.
- An important component in the iterative procedure is a tailored cutting-plane algorithm. This algorithm mainly involves a QCP model that only contains one quadratic constraint and a series

of linear constraints, which could be solved by off-the-shelf algorithms and tools. It is used to find a candidate portfolio which dominates the benchmark in terms of AASSD and also ASSD in the first step, and to find an upper bound of the optimal objective value of the ASSD-enhanced portfolio problem in the second step so that optimality gap could be explicitly calculated.

This paper is structured as follows. Section 2 describes the portfolio optimization problem under study. Section 3 introduces an approximation condition of ASSD and discusses its properties. Section 4 presents optimization models and algorithms to identify ASSD-enhanced portfolios. Section 5 applies our optimization methods to the Fama and French 49 industry portfolios. Further discussions are expounded in Section 6, followed by the conclusion in Section 7.

2. Preliminaries

We consider here the setting of a one-period portfolio optimization problem, which is also the general setting commonly considered in portfolio optimization literature (e.g., Dentcheva and Ruszczyński (2006), Post and Potì (2017), Post and Kopa (2017), Post et al. (2018), Kallio and Hardoroudi (2018), Liesiö et al. (2020)). Consider n distinct base assets $\{1, 2, ..., n\}$ with random investment returns $(X_1, X_2, ..., X_n) \in \mathcal{X}^n$, where $\mathcal{X} := [a, b], -\infty < a < b < \infty$. Without loss of generality, we assume that the investment return is finite and bounded between a and b. Let u_i denote the expected value of X_i and $Cov(x_i, x_j)$ denote the covariance between X_i and X_j. Suppose the joint CDF of the returns of the base assets $F(x_1, ..., x_n)$ are discrete with T mutually exclusive and collectively exhaustive scenarios. Let $x_{i,t}$ denote the realization of X_i under scenario t. The probability of scenario t denoted by p_t satisfies $\sum_{t=1}^{T} p_t = 1$. A portfolio is a convex combination of these assets characterized by a vector of asset weights $\lambda = (\lambda_1, ..., \lambda_n) \in \Lambda$ representing the share of initial capital allocated for each asset, where $\Lambda := \{\lambda \in \mathbb{R}^n : \sum_{i=1}^{n} \lambda_i = 1, \lambda_i \geq 0\}$ is the set of all possible portfolios. The return of a portfolio λ, denoted by $X = \sum_{i=1}^{n} X_i \lambda_i$, is a random variable. We use Y to denote the investment return of a given benchmark portfolio, which can be a mixture of the base assets or any desired target return distributions of the decision maker. The objective of our research is to construct an enhanced portfolio that dominates the benchmark portfolio in terms of ASSD. Without loss of generality, we assume that Y is discrete with scenarios $t' = 1, ..., T_b$ with probability $p_{t'}$ and its return is ranked in ascending order $b_1 \leq b_2 ... \leq b_{T_b}$. The scenarios of Y can be either the same or different from the scenarios of the base assets. The CDF of the investment return X of portfolio λ is given by:

$$F_\lambda(x) = \sum_{t=1}^{T} p_t I_{\lambda,t}(x),$$

where $I_{\lambda,t}(x)$ is an indicator function that takes value 1 if $\sum_{i=1}^{n} x_{i,t}\lambda_i \leq x$ and zero otherwise. The first-order LPM for X is given by:

$$F_\lambda^{(2)}(z) = \int_a^z F_\lambda(x)dx = \mathbb{E}_{F_\lambda}[(z-x)I_{\lambda,t}(z)] = \sum_{t=1}^{T} p_t(z - \sum_{i=1}^{n} x_{i,t}\lambda_i)I_{\lambda,t}(z).$$

In general, $F_\lambda^{(2)}(z)$ is a nonnegative, nondecreasing and convex function. We could obtain the value of $F_\lambda^{(2)}(z)$ with the following LP problem (Rockafellar et al. (2000), Section 3):

$$\begin{aligned} F_\lambda^{(2)}(z) = &\min_\theta \sum_{t=1}^{T} p_t\theta_t \\ &s.t.\ \theta_t \geq z - \sum_{i=1}^{n} x_{i,t}\lambda_i, t = 1, ..., T \\ &\theta_t \geq 0, t = 1, ..., T. \end{aligned} \tag{1}$$

Similarly, the CDF of the benchmark portfolio Y is given by:

$$G(x) = \sum_{t'=1}^{T_b} p_{t'} I(b_{t'} \leq x)$$

and the first-order LPM for Y is given by:

$$G^{(2)}(z) = \sum_{t'=1}^{T_b} p_{t'}(z - b_{t'})I(b_{t'} \leq z),$$

where $I(b_{t'} \leq z)$ is an indicator function that takes value of 1 if $b_{t'} \leq z$ and zero otherwise. Under our problem setting, $G^{(2)}(z)$ is a piecewise-linear, nondecreasing and convex function.

3. Approximation of ASSD

In this paper, we aim at identifying the best portfolio that dominates the given benchmark portfolio by ASSD. The key challenge in this problem is to identify the set of feasible portfolios that satisfy the ASSD constraint when there exist an infinite number of candidate portfolios. To address this challenge, we propose an approximation condition of ASSD in this section, which is the foundation of our optimization models and algorithms.

We first begin with an introduction of ASSD. In this study, we will focus on ASSD proposed by Leshno and Levy (2002) and Tzeng et al. (2013) and introduce it in detail in the following section. Other versions of ASD are omitted here for concise presentation. Interested readers are referred to Levy (2016) and Luo and Tan (2020) for a detailed review.

3.1. Almost second-order stochastic dominance

Consider two random variables X and Y. Let F and G denote the CDF of X and Y respectively. Without loss of generality, we assume $\mathbb{E}_F(X) \geq \mathbb{E}_G(Y)$ throughout this paper, where $\mathbb{E}_F(X)$ and $\mathbb{E}_G(Y)$ denote the expected value of X and Y respectively. Suppose X and Y are bounded in the range of $[a,b]$. Define the set of outcomes S_2 as $S_2 = \{t \in [a,b] : F^{(2)}(t) > G^{(2)}(t)\}$, where $F^{(2)}(t) = \int_{-\infty}^{t} F(x)dx$ and $G^{(2)}(t) = \int_{-\infty}^{t} G(x)dx$. In particular, S_2 is the range of outcomes where the condition for F dominating G by SSD is violated. Let $\bar{S}_2$ denote the complement of S_2.

Definition 1 (ASSD). For $\tau \in (1,\infty)$, F dominates G by τ-ASSD if and only if:

$$\tau \leq \frac{\int_{\bar{S}_2} \left(G^{(2)}(t) - F^{(2)}(t)\right) dt}{\int_{S_2} \left(F^{(2)}(t) - G^{(2)}(t)\right) dt},$$

and $\mathbb{E}_F(X) \geq \mathbb{E}_G(Y)$.

The parameter ε in the definition of ASSD in Leshno and Levy (2002) and Tzeng et al. (2013) is replaced by τ (i.e., $\tau = \frac{1}{\varepsilon} - 1$) in our definition above, which is consistent with Tan (2015), Tan and Luo (2017), Luo and Tan (2020). Tzeng et al. (2013) showed that τ-ASSD is sufficient and necessary for the preferences of decision makers whose utility function is in $U_2^*(\tau)$:

$$U_2^*(\tau) = \{u \in U_2 : -u''(t) \leq \inf\{-u''(t)\}\tau, \forall t\},$$

where U_2 denotes the set of utility functions with $u'(t) \geq 0$ and $u''(t) \leq 0$. In particular, ASSD is necessary and sufficient for ensuring the preferences of all risk-averse decision makers whose second derivatives of the utility function deviate by no more than τ.

3.2. Approximate ASSD

Let $\mathcal{D} := \{d_1, d_2, ..., d_{T_{\mathcal{D}}}\}$ denote a discretization scheme of the investment return range, where $d_s < d_{s+1}, s = 1, ..., T_{\mathcal{D}} - 1$. In particular, $\mathcal{D}$ divides the investment return range $[a, b]$ into $T_D - 1$ return sub-intervals: $[d_s, d_{s+1}], s = 1, ..., T_{\mathcal{D}} - 1$. To guarantee the first-order LPM of the benchmark portfolio $G^{(2)}(t)$ to be linear in each return sub-interval $[d_s, d_{s+1}], s = 1, ..., T_{\mathcal{D}} - 1$, $\mathcal{D}$ is set to include all possible realizations of the benchmark portfolio, that is, $b_{t'} \in \mathcal{D}, \forall t' = 1, 2, ..., T_b$. Without loss of generality, we further set that $d_1 = a$ and $d_{T_{\mathcal{D}}} = b$.

Definition 2 (Approximate SSD Violation). Given a discretization scheme $\mathcal{D} = \{d_1, d_2, ..., d_{T_{\mathcal{D}}}\}$, the approximate SSD violation of X dominating Y in return sub-interval $[d_s, d_{s+1}], s = 1, ..., T_{\mathcal{D}} - 1$, denoted by $A_{\mathcal{D},s,\lambda}$, is defined by the following way:

- Case 1: if $F_\lambda^{(2)}(d_s) \leq G^{(2)}(d_s)$ and $F_\lambda^{(2)}(d_{s+1}) \leq G^{(2)}(d_{s+1})$, then $A_{\mathcal{D},s,\lambda} = 0$.
- Case 2: if $F_\lambda^{(2)}(d_s) \leq G^{(2)}(d_s)$ and $F_\lambda^{(2)}(d_{s+1}) \geq G^{(2)}(d_{s+1})$ with at least one inequality to be strict, then:
$$A_{\mathcal{D},s,\lambda} = \frac{(d_{s+1} - d_s)\left(F_\lambda^{(2)}(d_{s+1}) - G^{(2)}(d_{s+1})\right)^2}{2\left(F_\lambda^{(2)}(d_{s+1}) - G^{(2)}(d_{s+1}) + G^{(2)}(d_s) - F_\lambda^{(2)}(d_s)\right)}.$$
- Case 3: if $F_\lambda^{(2)}(d_s) \geq G^{(2)}(d_s)$ and $F_\lambda^{(2)}(d_{s+1}) \leq G^{(2)}(d_{s+1})$ with at least one inequality to be strict, then:
$$A_{\mathcal{D},s,\lambda} = \frac{(d_{s+1} - d_s)\left(F_\lambda^{(2)}(d_s) - G^{(2)}(d_s)\right)^2}{2\left(G^{(2)}(d_{s+1}) - F_\lambda^{(2)}(d_{s+1}) + F_\lambda^{(2)}(d_s) - G^{(2)}(d_s)\right)}.$$
- Case 4: if $F_\lambda^{(2)}(d_s) \geq G^{(2)}(d_s)$ and $F_\lambda^{(2)}(d_{s+1}) \geq G^{(2)}(d_{s+1})$ with at least one inequality to be strict, then:
$$A_{\mathcal{D},s,\lambda} = \left(F_\lambda^{(2)}(d_{s+1}) + F_\lambda^{(2)}(d_s)\right)\frac{d_{s+1} - d_s}{2} - \left(G^{(2)}(d_{s+1}) + G^{(2)}(d_s)\right)\frac{d_{s+1} - d_s}{2}.$$

The essence of the approximate SSD violation is that we approximate F_λ (the true probability distribution of X) by an approximate probability distribution $\hat{F}_{\mathcal{D},\lambda}$ such that:

$$\hat{F}_{\mathcal{D},\lambda}^{(2)}(z) = F_\lambda^{(2)}(d_s) + \frac{F_\lambda^{(2)}(d_{s+1}) - F_\lambda^{(2)}(d_s)}{d_{s+1} - d_s} * (z - d_s) \qquad if\ z \in [d_s, d_{s+1}], s = 1, ...T_{\mathcal{D}} - 1. \tag{2}$$

Specifically, $\hat{F}^{(2)}_{\mathcal{D},\lambda}(z)$ is a piecewise linear function such that $\hat{F}^{(2)}_{\mathcal{D},\lambda}(d_s) = F^{(2)}_{\lambda}(d_s), s = 1, 2, ..., T_{\mathcal{D}}$. By simple derivation, it can be proved that $\hat{F}_{\mathcal{D},\lambda}$ is a valid discrete probability distribution whose support is a subset of the discretization scheme $\mathcal{D}$.

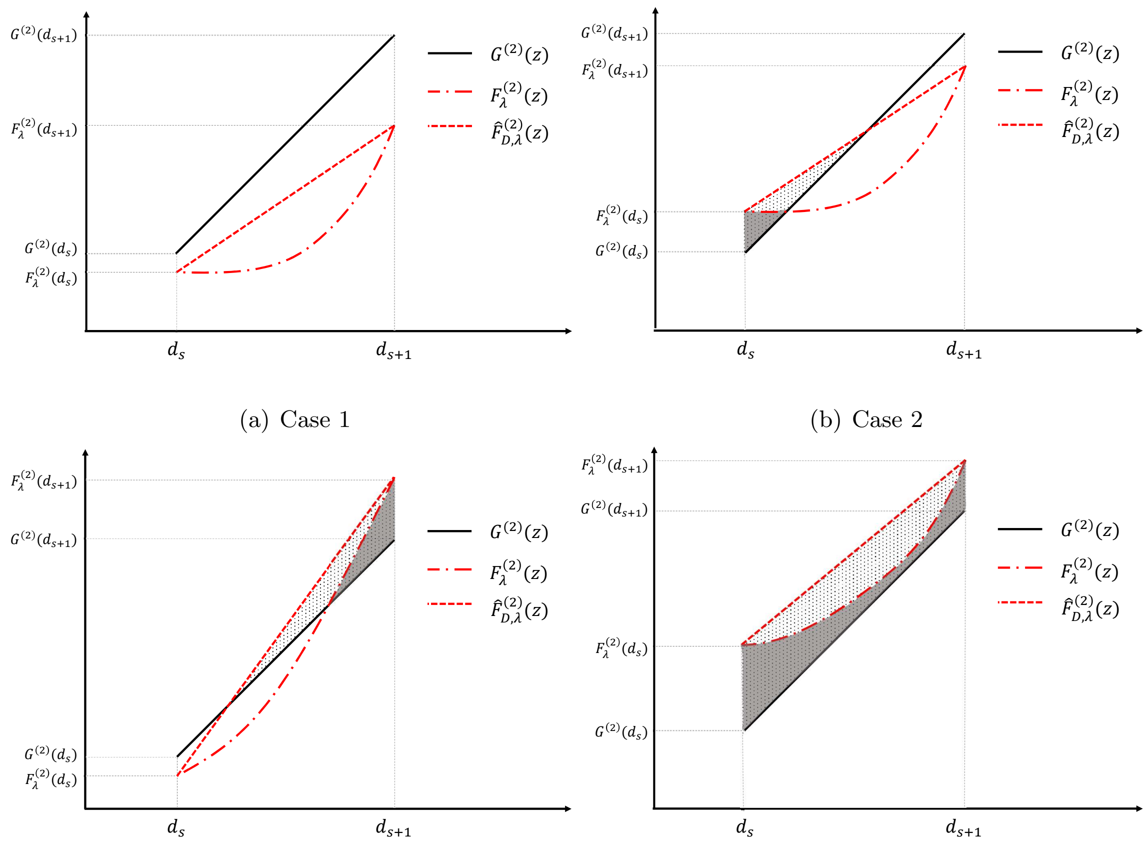
The approximate SSD violation corresponding to the four cases in Definition 2 is demonstrated in Figure 1. Note that $G^{(2)}(z)$ is linear in sub-interval $[d_s, d_{s+1}]$ under our setting and $F^{(2)}_{\lambda}(z)$ is non-decreasing and convex. In Case 1, the condition $F^{(2)}_{\lambda}(z)$ is no greater than $G^{(2)}(z)$ at the endpoints implies that $F^{(2)}_{\lambda}(z) \leq \hat{F}^{(2)}_{\mathcal{D},\lambda}(z) \leq G^{(2)}(z), \forall z \in [d_s, d_{s+1}]$ (see Figure 1(a)). Hence, both true and approximate SSD violation equal zero in this case. In Case 2, 3, and 4, the true SSD violation is the area where $F^{(2)}_{\lambda}(z) > G^{(2)}(z)$ (the gray area in Figure 1) while the approximate SSD violation $A_{\mathcal{D},s,\lambda}$ is the area where $\hat{F}_{\mathcal{D},\lambda} > G^{(2)}(z)$ (the dotted area Figure 1). Due to the convexity property of $F^{(2)}_{\lambda}(z)$, $\hat{F}^{(2)}_{\mathcal{D},\lambda}(x) \geq F^{(2)}_{\lambda}(x)$ always holds (shown in the following Proposition 1). As a result, the approximate SSD violation is always no less than the true SSD violation. Next, we present some important properties of $\hat{F}_{\mathcal{D},\lambda}$.

Proposition 1. *$\hat{F}^{(2)}_{\mathcal{D},\lambda}(x) \geq F^{(2)}_{\lambda}(x), \forall x \in \mathbb{R}$.*

Proposition 1 states that the first-order LPM of the approximate probability distribution $\hat{F}_{\mathcal{D},\lambda}$ is always no smaller than that of the true probability distribution F_{λ} under any discretization scheme of the return range. Proposition 1 also implies that $\hat{F}_{\mathcal{D},\lambda}$ is dominated by F_{λ} in terms of SSD. In the following, we use $\mathbb{E}_{\hat{F}_{\mathcal{D},\lambda}}$ and $Var_{\hat{F}_{\mathcal{D},\lambda}}$ to denote the expected value and variance of $\hat{F}_{\mathcal{D},\lambda}$, respectively. Similar notations are used for F_{λ}.

Proposition 2. *$\mathbb{E}_{\hat{F}_{\mathcal{D},\lambda}} = \mathbb{E}_{F_{\lambda}}$ and $Var_{\hat{F}_{\mathcal{D},\lambda}} \geq Var_{F_{\lambda}}$.*

Proposition 2 shows that the approximate distribution $\hat{F}_{\mathcal{D},\lambda}$ is a mean-preserving spread of true distribution F_{λ}, that is, $\hat{F}_{\mathcal{D},\lambda}$ and F_{λ} has identical expected value. Moreover, $\hat{F}_{\mathcal{D},\lambda}$ has higher variance than F_{λ}. Hence, $\hat{F}_{\mathcal{D},\lambda}$ is also dominated by F_{λ} in terms of MV criterion. Proposition 1 and 2 show that the approximate distribution $\hat{F}_{\mathcal{D},\lambda}$ has higher risk than F_{λ} in terms of both SSD and MV criterion.



(a) Case 1

(b) Case 2

(c) Case 3

(d) Case 4

Figure 1 Approximate SSD violations. The gray area is the true SSD violation while the dotted area is the approximate SSD violation.

Proposition 3. $\int_{\hat{S}_2} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)dt = \sum_{s=1}^{T_\mathcal{D}-1} A_{\mathcal{D},s,\lambda}$, *where* $\hat{S}_2 = \{t \in [a,b] : \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) > G^{(2)}(t)\}$.

Proposition 3 indicates that the approximate SSD violation between F_λ and G equals the SSD violation between $\hat{F}_{\mathcal{D},\lambda}$ and G. Next, we define approximate ASSD between portfolio λ with investment return X and the benchmark portfolio Y based on approximate SSD violation. For convenience of notation, in the rest of this paper we will use X (or Y) to represent portfolio λ (or the benchmark) as well as its return when no ambiguity is present. In addition, we will use $\mathbb{E}_{F_\lambda}(X)$ and $Var(X)$ to denote the expected value and variance of X, and similar notations for Y.

Definition 3 (Approximate ASSD). Given a discretization scheme $\mathcal{D}$, for $\tau \in (1,\infty)$, portfolio X dominates portfolio Y by τ-AASSD if and only if:

$$\tau \leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_\mathcal{D}-1} A_{\mathcal{D},s,\lambda}} + 1,$$

and $\mathbb{E}_{F_\lambda}(X) \geq \mathbb{E}_G(Y)$.

AASSD is intended as an approximation of ASSD rather than as an alternative of ASSD. Compared with ASSD, AASSD is easier to identify and use in practice. To be specific, ASSD needs to know the exact CDF formulation of the compared portfolios. However, AASSD only needs mean, variance and values of LPM at a series of discrete points in $\mathcal{D}$. Next, we show that AASSD is a sufficient condition of ASSD.

THEOREM 1. *If there exist a discretization scheme $\mathcal{D}$ such that portfolio X dominates portfolio Y by τ-AASSD, then X dominates Y by τ-ASSD, that is, (τ-AASSD)$\Rightarrow$(τ-ASSD).*

Theorem 1 shows that if portfolio X dominates Y by τ-AASSD under a discretization scheme, then X dominates Y by τ-ASSD. However, the other direction does not necessarily hold. This implies that AASSD is sufficient but not necessary for ASSD. Another thing of interest about AASSD is whether it converges to ASSD in some way. In particular, whether the 'gap' between AASSD and ASSD could be reduced or eliminated as the discretization scheme is refined? In the following section, we show this is the case.

3.3. Convergence of AASSD

In this section, we will show that AASSD approaches ASSD as we refine the discretization scheme by dividing some or all return sub-intervals into multiple smaller sub-intervals. In essence, AASSD approximates ASSD by approximating F_λ by $\hat{F}_{\mathcal{D},\lambda}$. Hence, the closeness between AASSD and ASSD is determined by the difference between F_λ and $\hat{F}_{\mathcal{D},\lambda}$. Next, we introduce a closeness index $\delta_{\mathcal{D},\lambda}$, which is defined to be the half variance difference between $\hat{F}_{\mathcal{D},\lambda}$ and F_λ. We will show how $\delta_{\mathcal{D},\lambda}$ measures the closeness between $\hat{F}_{\mathcal{D},\lambda}$ and F_λ as well as closeness between AASSD and ASSD.

DEFINITION 4 (CLOSENESS INDEX BETWEEN $\hat{F}_{\mathcal{D},\lambda}$ AND F_λ). Given a discretization scheme $\mathcal{D}$, $\hat{F}_{\mathcal{D},\lambda}$ and F_λ, the closeness index $\delta_{\mathcal{D},\lambda} = \frac{Var_{\hat{F}_{\mathcal{D},\lambda}} - Var_{F_\lambda}}{2}$.

PROPOSITION 4. $\delta_{\mathcal{D},\lambda} = \int_a^b \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - F^{(2)}_\lambda(t) \right) dt$.

Proposition 4 shows that $\delta_{\mathcal{D},\lambda}$ equals the integral of the first-order LPM difference between $\hat{F}_{\mathcal{D},\lambda}$ and F_λ. As $\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) \geq F^{(2)}_\lambda(t), \forall t \in \mathbb{R}$ (see Proposition 1), a smaller $\delta_{\mathcal{D},\lambda}$ implies a smaller difference between $\hat{F}_{\mathcal{D},\lambda}$ and F_λ and $\delta_{\mathcal{D},\lambda} = 0$ implies that $\hat{F}_{\mathcal{D},\lambda} = F_\lambda$.

THEOREM 2. *Given a discretization scheme $\mathcal{D}$, $0 \leq \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \int_{S_2} \left(F_\lambda^{(2)}(t) - G^{(2)}(t)\right) dt \leq \delta_{\mathcal{D},\lambda}$.*

Note that in Definition 3, AASSD replaces true SSD violation $\int_{S_2} \left(F_\lambda^{(2)}(t) - G^{(2)}(t)\right) dt$ in the definition of ASSD by approximate SSD violation $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}$. Hence, the difference between $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}$ and $\int_{S_2} \left(F_\lambda^{(2)}(t) - G^{(2)}(t)\right) dt$ is the cause of the difference between ASSD and AASSD. Theorem 2 shows that the difference between $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}$ and $\int_{S_2} \left(F_\lambda^{(2)}(t) - G^{(2)}(t)\right) dt$ is nonnegative and bounded from above by $\delta_{\mathcal{D},\lambda}$. The nonnegative property is consistent with the property that AASSD is sufficient for ASSD. The upper bound property implies that the smaller $\delta_{\mathcal{D},\lambda}$ is, the smaller the difference between ASSD and AASSD is.

COROLLARY 1. *When $\delta_{\mathcal{D},\lambda} \to 0, \hat{F}_{\mathcal{D},\lambda} \to F_\lambda$ and τ-AASSD $\to$ τ-ASSD.*

Corollary 1 states that when the closeness index $\delta_{\mathcal{D},\lambda}$ approaches zero, the approximate distribution $\hat{F}_{\mathcal{D},\lambda}$ will approach the true distribution F_λ and τ-AASSD will approach τ-ASSD. Corollary 1 shows the convergence of AASSD to ASSD at its limiting case. Next, we will show that refining discretization scheme will make AASSD closer to ASSD.

Given a discretization scheme $\mathcal{D} = \{d_1, d_2, ..., d_{T_{\mathcal{D}}}\}$, we define $\mathcal{D}^r$ as a refined discretization scheme of $\mathcal{D}$ if $\mathcal{D} \subset \mathcal{D}^r$, that is, $\mathcal{D}^r$ contains all discrete points $d_1, d_2, ..., d_{T_{\mathcal{D}}}$ in $\mathcal{D}$ and at least one extra distinct discrete point. More specifically, $\mathcal{D}^r$ refines $\mathcal{D}$ by dividing some or all return sub-intervals into multiple smaller sub-intervals.

PROPOSITION 5. *Given a discretization scheme $\mathcal{D}$, if $\mathcal{D}^r$ is a refined discretization scheme of $\mathcal{D}$, then the following two statements are true:*

(1) $\delta_{\mathcal{D}^r,\lambda} \leq \delta_{\mathcal{D},\lambda}$;

(2) $\sum A_{\mathcal{D}^r,s,\lambda} \leq \sum A_{\mathcal{D},s,\lambda}$.

The first statement of Proposition 5 indicates that refining the discretization scheme by dividing some or all return sub-intervals into smaller ones will lead to a smaller closeness index, and therefore make the approximate distribution closer to the true distribution. The second statement of Proposition 5 further shows that refining the discretization scheme will make approximate SSD violation smaller and closer to the true SSD violation, which therefore will make AASSD closer to ASSD.

PROPOSITION 6. *If X dominates Y by τ-AASSD under discretization scheme $\mathcal{D}$, then X dominates Y by τ-AASSD under any refined discretization scheme of $\mathcal{D}$ (i.e., $\mathcal{D}^r$).*

Proposition 6 states that existing AASSD will be preserved when refining the discretization scheme. On the other hand, it follows from Proposition 5 that AASSD relationship that does not exist under $\mathcal{D}$ may be identified under the refined discretization scheme $\mathcal{D}^r$. Hence, refining discretization scheme shall identify more portfolios dominating the benchmark in terms of AASSD as well as ASSD.

EXAMPLE 1. (Numerical Example of AASSD). We consider a simple numerical example here to demonstrate the relationship between AASSD and ASSD. Suppose X and Y are two random portfolios with random equal-probability investment returns $\{-0.1, 0.1, 0.3, 0.5\}$ and $\{0.02, 0.04\}$, respectively. Let $F(t)$ and $G(t)$ denote the probability distribution of X and Y respectively. Their CDFs and first-order LPMs are shown in Figure 2(a) and 2(b), respectively. The SSD violation area of X dominating Y is the orange area filled with circles in Figure 2(b) and equals: $\int_{S_2} \left(F^{(2)}(t) - G^{(2)}(t)\right) dt = 0.002767$. The light blue area filled with slash "/" in Figure 2(b) represents: $\int_{\bar{S}_2} \left(G^{(2)}(t) - F^{(2)}(t)\right) dt = 0.04327$. The existence of SSD violation indicates that X does not dominates Y by SSD and vice-versa. According to Definition 1, portfolio X dominates portfolio Y by τ-ASSD if and only if:

$$\tau \leq \frac{\int_{\bar{S}_2} \left(G^{(2)}(t) - F^{(2)}(t)\right) dt}{\int_{S_2} \left(F^{(2)}(t) - G^{(2)}(t)\right) dt} = \frac{0.04327}{0.00277} = 15.64.$$

Next, we show how AASSD works. Consider a discretization scheme of return range $\mathcal{D}_1 = \{-0.1, 0.02, 0.04, 0.5\}$, which includes all possible realizations of the benchmark portfolio Y as required by our setting. It can be observed from Figure 2(b) that the approximate first-order LPM $\hat{F}^{(2)}_{\mathcal{D}_1}(z)$ (the red dash-dotted line) is always greater than the true function $F^{(2)}(z)$ (the solid blue line), which is consistent with Proposition 1. In addition, $\hat{F}^{(2)}_{\mathcal{D}_1}(z)$ and $F^{(2)}(z)$ takes the same value when $z=0.5$ (upper bound for investment return of X as well as Y), indicating $\mathbb{E}_{\hat{F}_{\mathcal{D},\lambda}} = \mathbb{E}_{F_\lambda}$ and $Var_{\hat{F}_{\mathcal{D},\lambda}} \geq Var_{F_\lambda}$, as shown in Proposition 2. By calculation, the approximate ASSD violation under $\mathcal{D}_1$ equals:

$$\sum_{s=1}^{T_{\mathcal{D}_1}-1} A_{\mathcal{D}_1,s,\lambda} = 0.003087 \geq \int_{S_2} \left(F^{(2)}(t) - G^{(2)}(t)\right) dt = 0.002767,$$

which is shown in Theorem 2. According to the definition of AASSD (Definition 3), X dominates Y by τ-AASSD if and only if:

$$\tau \leq \frac{Var(Y) + (b-\mathbb{E}_G(Y))^2 - Var(X) - (b-\mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_\mathcal{D}-1} A_{\mathcal{D},s,\lambda}} + 1 = \frac{0.081}{2*0.003087} + 1 = 14.12.$$

Note that X dominates Y by AASSD for all $\tau \geq 14.12$ while ASSD holds for all $\tau \geq 15.64$. Hence, for any given value of τ, if AASSD holds, ASSD also holds. But the other direction does not necessarily hold. This is consistent with Theorem 1.

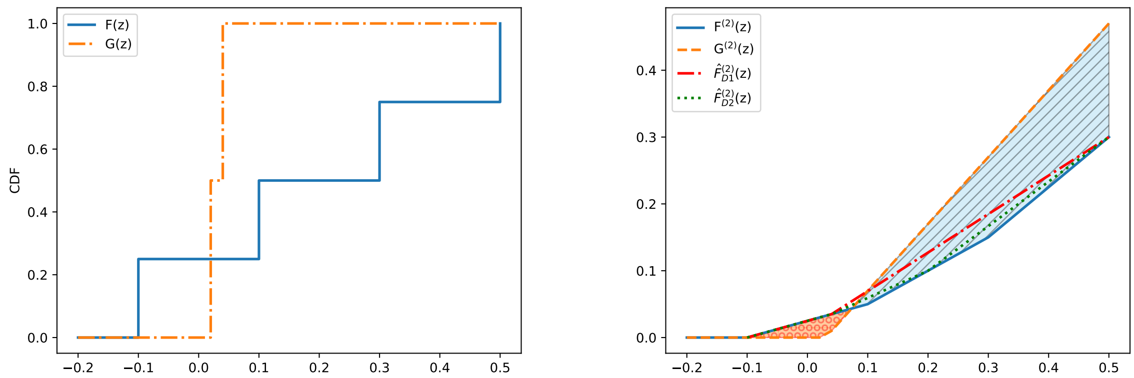
Next, we show the convergence of AASSD to ASSD. We consider a refined discretization scheme $\mathcal{D}_2 = \{-0.1, 0.02, 0.04, 0.2, 0.5\}$, as shown by the green dotted line in Figure 2(b). Under $\mathcal{D}_2$, approximate SSD violation $\sum_{s=1}^{T_{\mathcal{D}_2}-1} A_{\mathcal{D}_2,s,\lambda} = 0.0028763$ and X dominates Y by τ-AASSD for all $\tau \leq 15.08$, both of which become closer to the values under ASSD. When the discretization scheme is further refined as $\mathcal{D}_3 = \{-0.1, 0.02, 0.04, 0.1, 0.2, 0.3, 0.5\}$, $\sum_{s=1}^{T_{\mathcal{D}_3}-1} A_{\mathcal{D}_3,s,\lambda} = 0.002767$ and X dominates Y by τ-AASSD for all $\tau \leq 15.64$, indicating convergence to ASSD.

4. Optimization model and algorithm

In this section, we develop an optimization method for constructing portfolios that dominate a benchmark portfolio in terms of ASSD based on AASSD. In particular, we focus on the following problem:

$$\begin{aligned} &\max_{\lambda \in \Lambda} \ f(\lambda) \\ &\textit{subject to} \ \ \lambda \succeq_{\tau-ASSD} Y \end{aligned} \tag{3}$$

where $\Lambda := \{\lambda \in \mathbb{R}^n : \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0\}$ and $f(\lambda)$ is the user-defined objective function.



(a) CDF (b) First-order LPM

Figure 2 CDF and first-order LPM of X and Y in Example 1.

To achieve this goal, we first develop optimization models and algorithm to identify the optimal portfolio that dominates the benchmark portfolio in terms of AASSD. An optimal AASSD-enhanced portfolio is a candidate solution of Problem (3), which dominates the benchmark solution by ASSD but may not have optimal $f(\lambda)$. Then we develop a method to calculate the optimality gap between the obtained candidate solution and the optimal solution of Problem (3). If the obtained optimality gap does not satisfy optimality tolerance level, then the discretization scheme will be refined. This three-step procedure will be repeated (i.e., obtain a solution, check if it is optimal, refine discretization scheme) until an optimal ASSD-enhanced solution is found.

4.1. AASSD-enhanced portfolio optimization

In this section, we develop an approach to solve the AASSD-enhanced portfolio optimization problem:

$$\begin{aligned} &\max_{\lambda\in\Lambda} f(\lambda) \\ &\textit{subject to} \;\; \lambda \succeq_{\tau-AASSD} Y \end{aligned} \tag{4}$$

According to the definition of AASSD (Definition 3), Problem (4) can be transformed into the following equivalent problem:

$$\max_{\lambda} f(\lambda) \tag{5a}$$

$$s.t. \sum_{i=1}^{n} \lambda_i = 1 \tag{5b}$$

$$\sum_{i=1}^{n} \lambda_i \mu_i \geq \mu_y \tag{5c}$$

$$2(\tau - 1) \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} \leq Var(Y) + (b - \mu_y)^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j Cov(x_i, x_j) - (b - \sum_{i=1}^{n} \lambda_i \mu_i)^2 \tag{5d}$$

$$\lambda \geq 0 \tag{5e}$$

where μ_i and μ_y is the mean of X_i and Y respectively. Constraints (5b) and (5c) are linear. However, Constraint (5d) is very complex. The right-hand-side of Constraint (5d) is quadratic in terms of λ. The left-right-hand-side of Constraint (5d) is even more challenging and intractable because $A_{\mathcal{D},s,\lambda}$ depends on λ with a complex relationship given by Definition 2, which is neither linear nor quadratic. Hence, it is impossible to solve this AASSD-enhanced portfolio optimization problem with off-the-shelf algorithms or tools. To address this issue, we develop a cutting-plane algorithm.

Let $S_{2,s,\lambda}$ denote the SSD violation area between $\hat{F}_{\mathcal{D},\lambda}(t)$ and $G(t)$ in the area of $[d_s, d_{s+1}]$, that is, $S_{2,s,\lambda} = \{t \in [d_s, d_{s+1}] : \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) > G^{(2)}(t)\}$ and let $\bar{S}_{2,s,\lambda} = \{t \in [d_s, d_{s+1}] : \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) \leq G^{(2)}(t)\}$ denote the complementary set of $S_{2,s,\lambda}$. Next, we present a theorem which is the foundation of the cutting-plane algorithm.

THEOREM 3. *Given a known portfolio λ', then for all portfolio $\lambda \in \Lambda$ and $s \in [1, 2, ..., T_{\mathcal{D}} - 1]$:*

$$A_{\mathcal{D},s,\lambda} \geq \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{(2)}_{\mathcal{D},\lambda'}(t) \right) dt \right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t) \right) dt \right)^+$$

Theorem 3 shows that given a known solution λ', then a lower bound of approximate SSD violation $A_{\mathcal{D},s,\lambda}$ of any portfolio λ can be obtained.

COROLLARY 2. *If $S_{2,s,\lambda} = S_{2,s,\lambda'}$, then*

$$A_{\mathcal{D},s,\lambda} = \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{(2)}_{\mathcal{D},\lambda'}(t) \right) dt \right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t) \right) dt \right)^+ .$$

Next, we show that the lower bound of $A_{\mathcal{D},s,\lambda}$ given by Theorem 3 can be represented by a series of linear constraints. As $\hat{F}^{(2)}_{\mathcal{D},\lambda'}(t)$ and $G^{(2)}(t)$ are both linear in return interval $[d_s, d_{s+1}]$, they can intersect at most once in $[d_s, d_{s+1}]$. Without loss of generality we let $S_{2,s,\lambda'} = (\underline{v}_s, \bar{v}_s)$ and $\bar{S}_{2,s,\lambda'} = (\underline{c}_s, \bar{c}_s)$.

THEOREM 4. *Given* $A_{\mathcal{D},s,\lambda'}$, $S_{2,s,\lambda'} = (\underline{v}_s, \bar{v}_s)$ *and* $\bar{S}_{2,s,\lambda'} = (\underline{c}_s, \bar{c}_s)$, *inequality:*

$$A_{\mathcal{D},s,\lambda} \geq \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{(2)}_{\mathcal{D},\lambda'}(t)\right) dt\right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt\right)^+$$

is equivalent to the following system of linear constraints:

$$\theta_{s,t} \geq d_s - \sum_{i=1}^{n} x_{i,t}\lambda_i, t = 1, ..., T \tag{6a}$$

$$\theta_{s+1,t} \geq d_{s+1} - \sum_{i=1}^{n} x_{i,t}\lambda_i, t = 1, ..., T \tag{6b}$$

$$\begin{aligned} h_{1,s} \geq A_{\mathcal{D},s,\lambda'} &+ \left(\sum_{t=1}^{T} p_t\theta_{s,t}\left(2d_{s+1} - \bar{v}_s - \underline{v}_s\right) + \sum_{t=1}^{T} p_t\theta_{s+1,t}\left(\bar{v}_s + \underline{v}_s - 2d_s\right)\right)\frac{\bar{v}_s - \underline{v}_s}{2(d_{s+1} - d_s)} \\ &- \left(\hat{F}^{(2)}_{\mathcal{D},\lambda'}(\bar{v}_s) + \hat{F}^{(2)}_{\mathcal{D},\lambda'}(\underline{v}_s)\right)\frac{\bar{v}_s - \underline{v}_s}{2} \end{aligned} \tag{6c}$$

$$\begin{aligned} h_{2,s} \geq &\left(\sum_{t=1}^{T} p_t\theta_{s,t}\left(2d_{s+1} - \bar{c}_s - \underline{c}_s\right) + \sum_{t=1}^{T} p_t\theta_{s+1,t}\left(\bar{c}_s + \underline{c}_s - 2d_s\right)\right)\frac{\bar{c}_s - \underline{c}_s}{2(d_{s+1} - d_s)} \\ &- \left(G^{(2)}(\bar{c}_s) + G^{(2)}(\underline{c}_s)\right)\frac{\bar{c}_s - \underline{c}_s}{2} \end{aligned} \tag{6d}$$

$$A_{\mathcal{D},s,\lambda} \geq h_{1,s} + h_{2,s} \tag{6e}$$

$$h_{1,s}, h_{2,s}, A_{\mathcal{D},s,\lambda}, \lambda, \theta_{s,t}, \theta_{s+1,t} \geq 0, t = 1, ..., T \tag{6f}$$

Note $A_{\mathcal{D},s,\lambda'}, \hat{F}^{(2)}_{\mathcal{D},\lambda'}$ and $G^{(2)}$ in (6) are all known constant given that λ' is known. Theorem 3 and 4 indicate that $A_{\mathcal{D},s,\lambda}$ (defined by Definition 2) in Constraint (5d) could be relaxed by a system of linear constraints. Based on this property, we develop a cutting-plane algorithm to solve the AASSD-enhanced portfolio optimization Problem (5). In this algorithm, we first relax the problem based on a given portfolio and check whether the optimal solution of the relaxed problem dominates the benchmark portfolio Y. If yes, then the obtained solution is an optimal solution of Problem (5). Otherwise, new cuts will be iteratively introduced into the relaxed problem until an optimal solution is found.

Let $\lambda^k, k = 0, ..., M$, denote an optimal solution of the relaxed problem (also called the master problem) in the kth iteration, which are all known in $(M+1)$th iteration. Let $(\underline{v}^k_s, \bar{v}^k_s)$ and $(\underline{c}^k_s, \bar{c}^k_s), k = 0, ..., M$, denote the area of S_{2,s,λ^k} and $\bar{S}_{2,s,\lambda^k}$ respectively. The relaxed AASSD-enhanced portfolio optimization problem in $(M+1)$th iteration is as follows:

$$\max\ f(\lambda)$$

$$s.t. \sum_{i=1}^{n} \lambda_i = 1 \tag{7a}$$

$$\sum_{i=1}^{n} \lambda_i \mu_i \geq \mu_y \tag{7b}$$

$$2(\tau - 1) \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} \leq Var(Y) + (b - \mu_y)^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j Cov(x_i, x_j) - (b - \sum_{i=1}^{n} \lambda_i \mu_i)^2 \tag{7c}$$

$$\theta_{s,t} \geq d_s - \sum_{i=1}^{n} x_{i,t} \lambda_i, \ t = 1, ..., T, \ \ and \ s = 1, ..., T_{\mathcal{D}} \tag{7d}$$

$$\begin{aligned} h_{1,s}^k \geq A_{\mathcal{D},s,\lambda^k} &+ \left(\sum_{t=1}^{T} p_t \theta_{s,t} \left(2d_{s+1} - \bar{v}_s^k - \underline{v}_s^k \right) + \sum_{t=1}^{T} p_t \theta_{s+1,t} \left(\bar{v}_s^k + \underline{v}_s^k - 2d_s \right) \right) \frac{\bar{v}_s^k - \underline{v}_s^k}{2(d_{s+1} - d_s)} \\ &- \left(\hat{F}_{\mathcal{D},\lambda'}^{(2)}(\bar{v}_s^k) + \hat{F}_{\mathcal{D},\lambda'}^{(2)}(\underline{v}_s^k) \right) \frac{\bar{v}_s^k - \underline{v}_s^k}{2}, \ s = 1, ..., T_{\mathcal{D}} - 1, \ \ and \ k = 0, ..., M \end{aligned} \tag{7e}$$

$$\begin{aligned} h_{2,s}^k \geq &\left(\sum_{t=1}^{T} p_t \theta_{s,t} \left(2d_{s+1} - \bar{c}_s^k - \underline{c}_s^k \right) + \sum_{t=1}^{T} p_t \theta_{s+1,t} \left(\bar{c}_s^k + \underline{c}_s^k - 2d_s \right) \right) \frac{\bar{c}_s^k - \underline{c}_s^k}{2(d_{s+1} - d_s)} \\ &- \left(G^{(2)}(\bar{c}_s^k) + G^{(2)}(\underline{c}_s^k) \right) \frac{\bar{c}_s^k - \underline{c}_s^k}{2}, \ s = 1, ..., T_{\mathcal{D}} - 1, \ \ and \ k = 0, ..., M \end{aligned} \tag{7f}$$

$$A_{\mathcal{D},s,\lambda} \geq h_{1,s}^k + h_{2,s}^k, \ s = 1, ..., T_{\mathcal{D}} - 1, \ \ and \ k = 0, ..., M \tag{7g}$$

$$h_{1,s}^k, h_{2,s}^k, A_{\mathcal{D},s,\lambda}, \theta_{s,t}, \lambda \geq 0, s = 1, ..., T_{\mathcal{D}} - 1, \ t = 1, ..., T, \ \ and \ k = 0, ..., M \tag{7h}$$

Constraint (7a) ensures that $\lambda \in \Lambda$. Constraints (7b) and (7c) follow from the definition of AASSD. Constraints (7d) to (7g) follow from Theorem 4 and relax Problem 5 by relaxing $A_{\mathcal{D},s,\lambda}$ with a series of lower bounds corresponding to $\lambda^k, k = 1, ..., M$. Constraints (7h) defines the domain of decision variables. Note that Problem (7) is a QCP problem with one single convex quadratic constraint (7c). All other constraints in Problem (7) are linear. This problem is a relaxed problem of Problem (5). An optimal solution λ^{M+1} of Problem (7) is optimal to Problem (5) if it satisfies the AASSD constraint. Otherwise, new constraints (6a)-(6f) corresponding to λ^{M+1} for $s = 1, ..., T_{\mathcal{D}} - 1$ shall be introduced to Problem (7) to further refine the feasible set. The cutting-plane algorithm is presented in Algorithm 1.

Two key concerns on Algorithm 1 are: (1) whether this algorithm will be trapped in an endless loop where subsequent iterations obtain the same optimal solution as the previous iteration, i.e., the situation where $\lambda^M = \lambda^{M+1}$; (2) whether updating the master problem in Step 2 could efficiently reduce the feasible set. We address these concerns in the following. Let λ^k denote an optimal

Algorithm 1 The cutting-plane algorithm for AASSD-enhanced portfolio optimization problem

Step 1: Set $M = 0$ and obtain the initial portfolio λ^0 by solving the following relaxed portfolio optimization problem:

$$\begin{aligned} &\max_{\lambda \in \Lambda} \ f(\lambda) \\ &s.t. \ \sum_{i=1}^{n} \lambda_i \mu_i \geq \mu_y \end{aligned} \tag{8}$$

Step 2: Check whether λ^M dominates Y by τ-AASSD.

- If λ^M is not dominated, update master Problem (7) by adding Constraints (6a)-(6f) for $s = 1, ..., T_{\mathcal{D}} - 1$ that correspond to λ^M.
- Otherwise, terminate and return λ^M.

Step 3: Set $M = M + 1$. Solve updated master Problem (7).

- If the model is feasible, let λ^M denote an optimal solution to this problem. Go back to Step 2.
- Otherwise, terminate and return the result that the model is infeasible.

Step 4: Repeat Step 2 and Step 3 until the obtained solution λ^M dominates Y by τ-AASSD or the master Problem (7) becomes infeasible.

Remark:
(1) In Setp 1, we assume Problem (8) is feasible without loss of generality.
(2) In Step 2, we set a tolerance value $\delta \geq 0$ for dominance condition and check whether λ^M dominates Y by τ-AASSD by the following condition:

$$2(\tau - 1) \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} \leq (1+\delta) \left(Var(Y) + (b - \mu_y)^2 - \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_i \lambda_j Cov(x_i, x_j) - (b - \sum_{i=1}^{n} \lambda_i \mu_i)^2 \right)$$

solution of the master problem in the kth iteration of Algorithm 1. Let $\Omega(\lambda, \lambda^k) = \{\lambda : S_{2,s,\lambda} = S_{2,s,\lambda^k}, s = 1, 2, ..., T_{\mathcal{D}} - 1\}$. To be specific, $\Omega(\lambda, \lambda^k)$ is the set of portfolios whose approximate SSD violation area is the same as that of λ^k.

Proposition 7. *The following statements are true for Algorithm 1:*

(1) If $\lambda^{k+1} \in \Omega(\lambda, \lambda^k)$, then λ^{k+1} is an optimal solution of AASSD-enhanced portfolio optimization Problem (5). On the other hand, if λ^{k+1} is not an optimal solution of Problem (5), then $\lambda^{k+1} \notin \Omega(\lambda, \lambda^k)$.

(2) In $(k+1)$th iteration, all portfolios $\lambda \in \Omega(\lambda, \lambda^k)$ which do not dominate Y by τ-AASSD are excluded from the feasible set of the master Problem (7)*.*

The first statement of Proposition 7 demonstrates that the obtained solution in $(k+1)$th iteration will lie outside $\Omega(\lambda, \lambda^k)$ if it is not an optimal solution of AASSD-enhanced portfolio optimization problem. As λ^k is in $\Omega(\lambda, \lambda^k)$ and does not satisfy dominance condition if the algorithm does not terminate at kth iteration, λ^{k+1} will be different from λ^k as well as $\lambda^n, \forall n < k$. This implies that the solution found in each iteration will be different from all solutions found in preceding iterations. Hence, additional cuts introduced to the master problem in each iteration will also be different from all cuts introduced before. The second statement of Proposition 7 indicates that in each iteration, new introduced cuts shall efficiently refine the feasible region of Problem (7) and make it closer to the feasible region of Problem (5) by excluding at least an extra set of portfolios. These facts show that Algorithm 1 does not fall into an infinite loop and guarantee that the solution obtained via this algorithm gradually converges to the desired optimal portfolio.

AASSD-enhanced portfolio optimization model identifies an optimal solution that dominates the benchmark solution by AASSD. As AASSD is a sufficient condition of ASSD (Theorem 1), this solution is a candidate solution of the ASSD-enhanced portfolio optimization problem (3). In particular, it dominates the benchmark by ASSD but may not have the best objective among all ASSD-enhanced solutions.

Example 2. (Example of AASSD-enhanced portfolio optimization.) We consider a small portfolio optimization problem with AASSD constraint and an objective function that maximizes expected return. The distributions of three base assets with investment return (X_1, X_2, X_3) and a benchmark asset Y are shown in Table 1. To be specific, base assets X_1 and X_2 are random variables and X_3 are risk-free with deterministic return rate. It can be easily find that $\mathbb{E}(X_1) > \mathbb{E}(X_2) > \mathbb{E}(X_3)$ while $Var(X_1) > Var(X_2) > Var(X_3)$, indicating declining level of return as well as risk. The benchmark asset Y has two possible returns 0.02 and 0.1 with equal probability in Scenario 1 and 2.

Table 1 Distributions of investment return of the base assets and benchmark.

Scenario	X_1	X_2	X_3	Probability	Scenario	Y	Probability
1	-0.2	-0.1	0.06	0.25	1	0.02	0.5
2	-0.2	0.3	0.06	0.25	2	0.1	0.5
3	0.5	-0.1	0.06	0.25			
4	0.5	0.3	0.06	0.25			

We consider the case where $\tau = 6$ and the discretization scheme is $\mathcal{D} = \{-0.2, -0.1, 0.02, 0.06, 0.1, 0.3, 0.5\}$, which is simply set to include all possible realizations of the base assets and benchmark. Next, we show how to apply Algorithm 1 to find an optimal AASSD-enhanced portfolio.

- Step 1: Solve Problem (8) with $f(\lambda) = \sum_{i=1}^{3} \lambda_i \mu_i$ and we obtain a candidate solution $\lambda^0 = [1.0, 0.0, 0.0]$.
- Step 2: Check whether λ^0 dominates Y by τ-AASSD. By calculation, we find that $R^* = \frac{Var(Y)+(b-\mathbb{E}_G(Y))^2 - Var(X)-(b-\mathbb{E}_{F_{\lambda^0}}(X))^2}{2\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda^0}} + 1 = 0.25 < 6$, indicating that λ^0 does not satisfy the AASSD Constraint (5d) and does not dominate the benchmark Y[1]. Hence, we update master Problem (7) by adding Constraints (6a)-(6f) corresponding to λ^0 for all $s = 1, 2, ..., 6$.
- Step 3: Solve the updated master Problem (7) and obtain a new candidate solution $\lambda^1 = (0.2341, 0.3466, 0.4193)$.
- Step 4: Repeat step 2 and 3 until termination. The obtained solution and its performance in each iteration are reported in Table 2. In this problem Algorithm 1 terminates at the second iteration and the obtained optimal AASSD-enhanced solution is $(0.2231, 0.3636, 0.4133)$.

4.2. Iterative Method for ASSD-enhanced portfolio optimization

In this section, we first introduce how to quantify the optimality gap between an obtained candidate solution (i.e., an obtained optimal AASSD-enhanced portfolio) and an optimal ASSD-enhanced portfolio. Next, we develop a method for identifying an optimal ASSD-enhanced portfolio based on these results.

[1] In this example, we define $R^* = \frac{Var(Y)+(b-\mathbb{E}_G(Y))^2 - Var(X)-(b-\mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}} + 1$, which is used to determine whether the obtained solution dominates the benchmark by τ-AASSD.

Table 2 Solution update of Algorithm 1 in Example 2

Iteration	Solution	Objective Value	R^*
Initialized	(1.0, 0.0, 0.0)	0.1500	0.25
1st	(0.2341, 0.3466, 0.4193)	0.0949	5.92
2ed	(0.2231, 0.3636, 0.4133)	0.0946	6.05

Let $\delta_{\mathcal{D}} := \max\{\delta_{\mathcal{D},\lambda}, \forall\lambda \in \Lambda\}$ denote the maximal closeness index among all valid portfolios. A necessary condition of ASSD could be established based on $\delta_{\mathcal{D}}$, which is presented as follows.

THEOREM 5. *Given $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}} > 0$, if X dominates Y by τ-ASSD, then:*

$$\tau \leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2(\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}})} + 1. \tag{9}$$

Theorem 5 present a necessary condition of ASSD based on $A_{\mathcal{D},s,\lambda}$ and $\delta_{\mathcal{D}}$. Incorporating this necessary condition into the portfolio optimization problem will lead to a feasible set containing all portfolios that dominate the benchmark in terms of ASSD. Hence, an optimal portfolio that satisfies this necessary condition has better or equivalent performance than all ASSD-enhanced portfolios and provides an upper bound for the optimal objective value of the ASSD-enhanced portfolio optimization problem. Based on this result, optimality gap between the candidate and optimal ASSD-enhanced solution could be calculated. To achieve this goal, we first develop a QP model to identify the value of $\delta_{\mathcal{D}}$.

PROPOSITION 8. *Given a discretization scheme $\mathcal{D}$,*

$$\delta_{\mathcal{D}} = \max_{\lambda,\theta} \sum_{s=1}^{T_{\mathcal{D}}-1} \left(\sum_{t=1}^{T} p_t\theta_{s,t} + \sum_{t=1}^{T} p_t\theta_{s+1,t} \right) \frac{d_{s+1} - d_s}{2} - \frac{1}{2} \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j Cov(x_i, x_j) + (b - \sum_{i=1}^{n} \lambda_i\mu_i)^2 \right) \tag{10a}$$

$$s.t. \quad \sum_{i=1}^{n} \lambda_i = 1 \tag{10b}$$

$$\sum_{i=1}^{n} \lambda_i\mu_i \geq \mu_y \tag{10c}$$

$$\theta_{s,t} = \{d_s - \sum_{i=1}^{n} x_{i,t}\lambda_i\}^+, \quad s = 1, ..., T_{\mathcal{D}} \text{ and } t = 1, ..., T \tag{10d}$$

$$\lambda \geq 0 \tag{10e}$$

where $\{z\}^+ = \max\{z, 0\}$.

Proposition 8 shows that $\delta_{\mathcal{D}}$ equals optimal value of objective function (10a). With $\delta_{\mathcal{D}}$ obtained, an upper bound of ASSD-enhanced portfolio could be identified based on Theorem 5 via the following optimization problem[2].

$$\max_{\lambda} \ f(\lambda) \tag{11a}$$

$$s.t. \ \sum_{i=1}^{n} \lambda_i = 1 \tag{11b}$$

$$\sum_{i=1}^{n} \lambda_i \mu_i \geq \mu_y \tag{11c}$$

$$2(\tau - 1)\left(\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_d\right) \leq Var(Y) + (b - \mu_y)^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j Cov(x_i, x_j) - (b - \sum_{i=1}^{n} \lambda_i \mu_i)^2 \tag{11d}$$

$$\lambda \geq 0 \tag{11e}$$

Problem (11) identifies an optimal portfolio that satisfies the necessary condition of ASSD presented in Theorem 5. Constraint (11d) is a simple equivalent transformation of Condition (9). Note that Problem (11) is almost the same as Problem (5) except that Constraint (11d) contains an extra constant $\delta_{\mathcal{D}}$, whose value could be calculated by solving Problem (10). Hence, Problem (11) could also be solved by the cutting-plane Algorithm 1.

Let λ^U denote an optimal solution of Problem (11). According to the above analysis, this solution has better or equivalent performance than all portfolios that dominates the benchmark portfolio by ASSD and could be used as an upper bound for the optimal objective value of ASSD-enhanced optimization problem (3). Hence, given a candidate solution λ for problem (3), its optimality gap could be estimated by:

$$gap(\lambda) = f(\lambda^U) - f(\lambda).$$

[2] Without loss of generality, we assume $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}} > 0$. This assumption could always be achieved by refining the discretization scheme finely enough.

Based on all of the above results, we develop an iterative method for solving Problem (3) and identifying an optimal portfolio that dominates the benchmark in terms of ASSD. Given a user-defined value of τ for ASSD, the steps of this method are as follows:

Start. Initialization: define an initialized discretization scheme $\mathcal{D}$ of the return range.

Step 1. Solve the τ-AASSD enhanced portfolio optimization problem (5) under discretization scheme $\mathcal{D}$ with Algorithm 1 to obtain a candidate solution λ^* for τ-ASSD enhanced optimization problem.

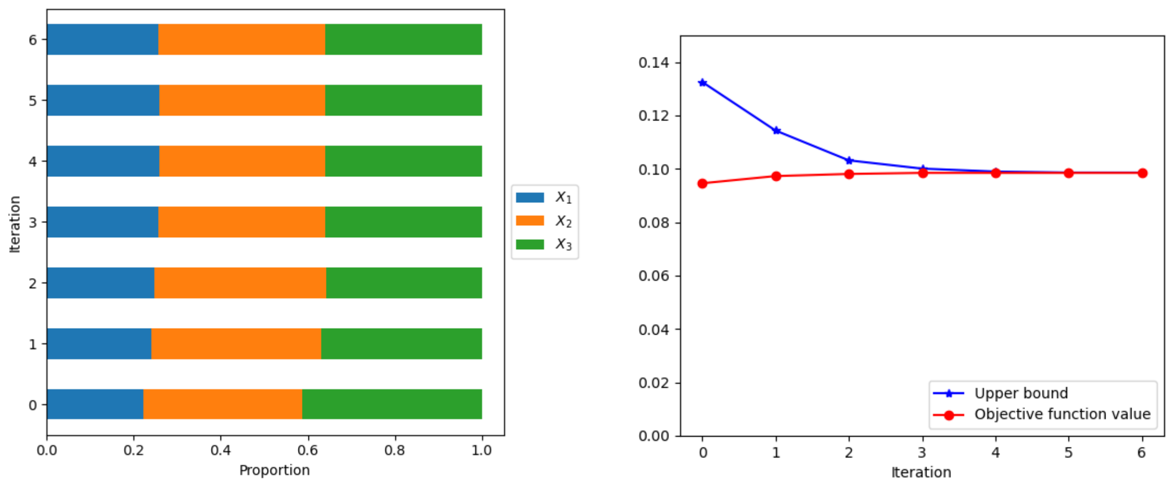
Step 2. Check whether the obtained candidate solution is optimal: solve Problem (10) and (11) and calculate the optimality gap $gap(\lambda^*)$.

- If the optimality gap is smaller than the optimality tolerance, then terminate and return the obtained candidate solution λ^* as the optimal solution.
- Otherwise, go to the next step.

Step 3. Refine the discretization scheme $\mathcal{D}$ and go back to the first step to update the candidate solution under the refined discretization scheme. Repeat the process until an optimal ASSD-enhanced portfolio is found[3].

The main contribution of this method is that it provides for the first time a feasible and tractable way to solve a portfolio optimization problem with ASSD constraints. Due to the complex formulation of ASSD, the analytical challenge of optimization with ASSD has been recognized (Arvanitis et al. 2021), and has prevented the application of ASSD in portfolio optimization as well as other areas of optimizations ever since its introduction in 2002 (Leshno and Levy 2002). Our method overcomes this challenge and transforms the problem into multiple iterative steps mainly involving QP and QCP, both of which could be solved by off-the-shelf solvers. Based on results about convergence of AASSD we proved in Section 3, it is theoretically guaranteed that a solution generated by our method improves and converges to an optimal ASSD-enhanced portfolio. Although this paper

[3] More details about how to apply this method is shown in Algorithm 2 in Appendix B. We adopt a naive rule in our work but other different refinement rules may be used. Specific study may be done in future on finding more efficient refinement rules.



(a) Asset allocation of solutions

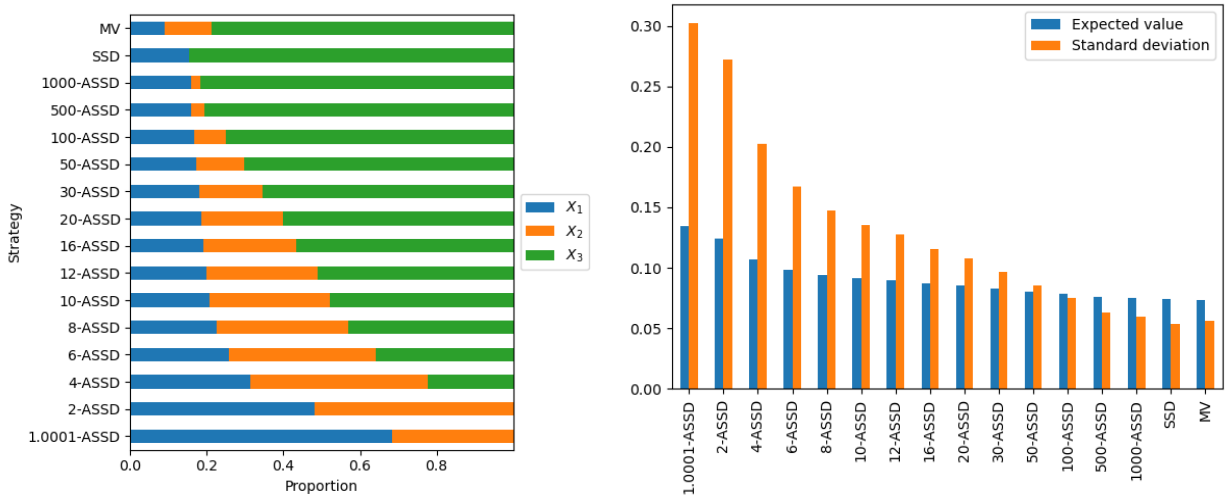
(b) Solution performance

Figure 3 Solution update of the iterative method in Example 3

focuses on portfolio optimization, to the best of our knowledge, this is the first attempt to tackle an optimization problem with ASSD constraints, and the proposed approach could be tailored and extended to other problems.

Example 3. (Example 2 continued: ASSD-enhanced portfolio optimization.) We continue to study the problem considered in Example 2 and show how to apply our iterative method to identify an optimal ASSD-enhanced portfolio. First, we consider the case where $\tau = 6$ and set the initialized discretization scheme $\mathcal{D}^0 = \{-0.2, -0.1, 0.02, 0.06, 0.1, 0.3, 0.5\}$. Detailed steps of our iterative method are presented as follows.

- Identify candidate solution. As shown in Example 2, we apply Algorithm 1 to identify an optimal AASSD-enhanced portfolio $\lambda^{*0} = (0.2231, 0.3636, 0.4133)$ with expected return 0.0946. This solution is a candidate optimal solution that dominates Y by ASSD with $\tau = 6$ but it may not be optimal.
- Check optimality. Next, we check whether λ^{*0} is optimal by calculating the optimality gap of λ^{*0}. Solve Problem (11) with $\delta_{\mathcal{D}^0}$ obtained by solving Problem (10). The obtained optimal objective value of Problem (11) is $f(\lambda^U) = 0.1326$ and the optimality gap equals $gap(\lambda^{*0}) = 0.1326 - 0.0946 = 0.0379$. As the gap is greater than our optimality tolerance (set to be 0.001 here), λ^{*0} is not optimal.



(a) Asset allocation (b) Expected value and standard deviation of return

Figure 4 Solutions of ASSD-enhanced portfolio optimization problems with different values of τ

- Refine discretization scheme and update candidate solution. Here we simply refine $\mathcal{D}^0$ by adding extra points $\frac{d_s+d_{s+1}}{2}$ for all $d_s \in \mathcal{D}^0$ and the refined scheme $\mathcal{D}^1 = \{-0.2, -0.15, -0.1, -0.04, 0.02, 0.04, 0.06, 0.08, 0.1, 0.2, 0.3, 0.4, 0.5\}$. Again we apply Algorithm 1 under $\mathcal{D}^1$ to find an updated candidate solution $\lambda^{*1} = (0.2403, 0.3909, 0.3688)$.
- Go back to the step of checking optimality step. Repeat the processes until the gap of the obtained solution is smaller than the tolerance level. The solution and its performance in each iteration are shown in Figure 3. It can be found that the optimality gap of the obtained solution declined quickly as the number of iterations increases and our method identifies an optimal ASSD-enhanced portfolio $(0.2581, 0.3816, 0.3604)$ with expected return 0.0985 in the 4th iteration.

Next, we investigate the effect of value of τ. Figure 4 show that as τ increase, τ-ASSD enhanced strategy invests smaller proportion in asset with high risk and high return (i.e., X_1) and assigns higher proportion in risk-free asset (i.e., X_3), leading to lower expected return and variance. This observation is consistent with the implication of ASSD that a larger τ implies a larger set of utility functions to be considered and higher level of conservatism (or risk-aversion). When τ becomes extremely large, solution and performance of ASSD-enhanced strategy converge to those of SSD-

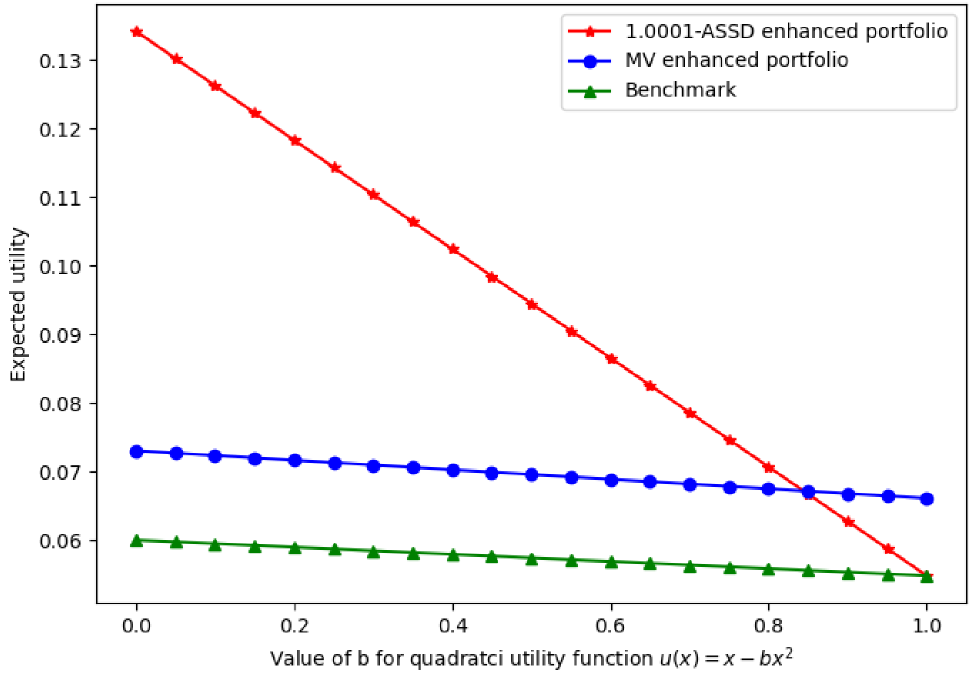


Figure 5 **Expected utility of ASSD-enhanced portfolio with τ approaches 1 and MV-enhanced portfolio.**

enhanced strategy. This is because when τ approaches infinity, $U_2^*(\tau)$ contains utility functions of all non-satiable and risk averse decision makers and therefore ASSD approaches SSD.

We further study another special case when τ approaches one. In this case, $U_2^*(\tau)$ only contains quadratic utility function, which is known to be an important sufficient condition for MV model. However, as it is shown in Figure 4, the optimal portfolios based on MV and ASSD strategy with τ approaches one[4] are quite different from each other. Specifically, the optimal solution of MV model is $(0.0902, 0.1227, 0.787)$ while it is $(0.6848, 0.3152, 0.0)$ for ASSD strategy with $\tau = 1.0001$. The former has smaller expected return and also smaller variance (0.073 and 0.0016) than the latter (0.134 and 0.0614 respectively). The difference between these two solutions could be explained by the fact that quadratic utility is sufficient but not necessary for the use of MV analysis in practice and "the necessary and sufficient condition for the practical use of mean–variance analysis is that a careful choice from a mean–variance efficient frontier will approximately maximize expected utility for a wide variety of concave (risk-averse) utility functions", as stated by Markowitz (2014).

[4] As τ is defined to be in the range of $(1, \infty)$, we take $\tau = 1.0001$ to represent τ approaches 1 in this paper.

Compared with the benchmark asset, portfolio optimization models based on MV and ASSD with τ approaches 1 generate two different solutions belonging to MV efficient frontier. Choice between these two solutions depends on decision maker's preference. We compare the expected utility of these two solutions by considering a quadratic utility function $u(x) = x - bx^2$ with $b \in [0, 1]$, which guarantees that the utility function is non-decreasing and concave in the range of return outcomes. Figure 5 shows that both portfolios have higher expected utility than the benchmark across all quadratic utility functions. However, portfolio generated by ASSD with τ approaches one has larger expected utility than MV portfolio for b smaller than 0.85 and the opposite for b greater than 0.85. This indicates that quadratic-utility decision makers with low or moderate degree of risk-aversion shall prefer ASSD-enhanced portfolio but these with high degree of risk-aversion will prefer MV-enhanced portfolio.

5. Application

We implement our optimization approach on the combination of the Fama and French 49 industry portfolios, which has been extensively studied in the literature (e.g., Hodder et al. (2015), Post and Kopa (2017), Liesiö et al. (2020)). We use average value weighted daily return data retrieved from the data library of Kenneth French[5] and our analysis commences from January 1927. The benchmark reflects the stock market return, based on the value-weighted return of all CRSP firms incorporated in the US and listed on the NYSE, AMEX, or NASDAQ, also available through Kenneth French's online library. Our study involves a comparison of portfolio performances, particularly portfolios enhanced by τ-ASSD against those enhanced by other criteria, including MV, SSD, superconvex TSD (SCTSD) proposed by Post and Kopa (2017), and stochastic bounding (SBK) methods using three distinct reference sets (K_2, K_3, and K_4) as detailed in Arvanitis et al. (2021). Additionally, we explore heuristic rules that involve purchasing an equal-weighted combination of 5, 10, and 15 industries with the highest average returns among the 49 industries.

Motivated by existing SD studies on momentum strategies (Post and Kopa 2017, Liesiö et al. 2020), a standard buy-hold trading strategy is employed with a 12-month formation period and

[5] `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

a 3-month holding period. No short positions are allowed. Under this strategy, the portfolio is constructed through optimization with 12-month historical return data and rebalanced after 3 months. The objective is to maximize the expected return with constraints that guarantee the enhanced portfolio dominates the benchmark by ASSD and other criteria (MV, SSD, SCTSD and SBK), respectively. We consider 377 overlapping formation periods from January 1927 to the end of 2021 and invest in base assets with the above investment strategies from January 1928 to March 2022 (for more than 94 years). We estimate the empirical joint return distribution using daily returns within each formation period and subsequently solve for the optimal portfolio during each period.

The algorithms are coded in Python and solved by Gurobi 9.1 running on a personal computer with 11th Gen Intel(R) Core(TM) i7-11700K @ 3.60GHz and 64.0 GB RAM. In our application, we adopt a naive approach for discretization scheme refinement in a way that divides all return sub-intervals in half. Computations for all optimization strategies are done based on brute-force use of Gurobi solver without further speeding up using more advanced algorithms (e.g., non-smooth convex optimization algorithms). On average the ASSD-enhanced optimization problem could be solved in about 2 minutes and 3 iterations. The average number of iterations of the cutting-plane Algorithm 1 is also less than 3.

5.1. Selection of τ

When implementing the proposed algorithm, an important task is to identify the value of τ, as it represents how small a violation of the SD rules would be allowed by decision makers. In the literature, the methods of identifying τ can be broadly categorized into two groups. The first group of methods use the individual selection data between two options and choose τ such that the considered dominance exists for most of the decision makers (Bali et al. 2013, Levy 2016, Lee et al. 2018, Bi and Zhu 2022, Han et al. 2023, Huang et al. 2021). However, due to differing definitions of ASSD in these studies compared to our framework, their estimated τ values cannot be directly applied to our model. The second approach for determining τ relies solely on historical data. For example, Post (2015) connects τ with the relative risk aversion (RRA) and the relative range b/a ($[a,b]$ is the range of the historical returns). This method leverages the extensive literature on RRA coefficients to derive the critical value of τ. However, this approach is not applicable to our case since it requires $0 < a < b$ and otherwise yielding τ value below 1.

In this study, we develop an intuitive data-driven selection rule for τ. In specific, we divide each formation period into two parts: a training period consisting of the first 9-month data and a testing period consisting of the last 3-month data. Based on the training dataset, the ASSD portfolio optimization is solved for each value of $\tau \in \{1.0001, 2, 4, 6, 8, 10, 12, 16, 20, 30, 50, 100, 500, 1000\}$. The optimal τ is then selected such that its enhanced portfolio achieves best performance in terms of a certain criterion in the testing dataset, e.g., the average return, the Sharpe ratio, or the certainty equivalent. This selected τ is then applied to the whole formation period which yields the optimal portfolio for the subsequent holding period. Table 3 shows the percentage of selected τ's by the criteria average return, Sharpe ratio and certainty equivalent for constant risk aversion utility functions with RAA equal to 5. As seen, the overall selection proportions of τ are quite consistent by different criteria. The most frequent choice is $\tau = 1.0001$, indicating the preference for solutions situated within the MV efficient frontier. Note that the optimal portfolios based on MV and ASSD strategy with τ approaching one could differ significantly, as discussed in Figure 5. On the other hand, $\tau = 1000$ is occasionally selected, implying that in such instances, the ASSD-enhanced portfolio is anticipated to perform similarly to the SSD-enhanced portfolio. Nevertheless, the moderate values of τ ($\tau \in [2, 500]$) account for approximately half of all the selected τ values, which underlines the distinctive value of using the ASSD-enhanced portfolio optimization.

5.2. Daily performance summary

In this section, we evaluate the performance of portfolios enhanced by the strategies under consideration. Table 4 and Table 5 show the in-sample and out-of-sample performances of ASSD-enhanced portfolios using three different selected τ values, along with portfolios enhanced by MV, SSD, SBK, SCTSD, and equal-weighted heuristics. We specifically evaluate the performances of these

Table 3 Percentage of different selected τ values

	1.0001	2	4	6	8	10	12	16	20	30	50	100	500	1000
τ-Mean	32.6	4.5	4.5	3.2	3.4	1.1	2.9	1.3	2.1	2.1	4.2	10.1	11.7	16.2
τ-Sharpe	39.5	5.6	6.4	4.8	2.9	2.7	1.6	1.9	3.2	2.9	3.2	6.1	10.6	8.8
τ-CE5	39.3	5.8	6.4	3.4	3.2	1.9	2.1	2.1	1.3	2.1	2.4	8.8	10.3	10.9

Notes: Shown are the percentage of different selected τ values based on the 377 formation period data. The first row represents the list of τ values. The last three rows represent the percentage of each τ by the criteria including daily average return, Sharpe ratio and certainty equivalent for constant risk aversion utility functions with degree of relative risk aversion 5 (CE5), respectively.

enhanced portfolios concerning excess returns over the risk-free Treasury bill (T-bill), and a series of performance metrics is presented. The in-sample performances involve the evaluation of portfolios using return data within the formation period. The reported values of performance metrics are the averages across 377 overlapping formation periods. Moreover, out-of-sample performance measures are computed using return data spanning from 1928 to 2022 across 377 holding periods.

Table 4 Daily in-sample performance based on excess over T-bill

	Mean	Std	VaR	CVaR	Skew	Sharpe	Sortino	CE2	CE5	CE10
MKT	0.0305	0.9457	-1.4881	-2.2376	-0.3487	0.0488	0.0461	0.0191	0.0018	-0.0278
EW_5	0.1406	1.1568	-1.7141	-2.6065	-0.3420	0.1359	0.2165	0.1243	0.0997	0.0575
EW_{10}	0.1147	1.0413	-1.5761	-2.4049	-0.4339	0.1262	0.1976	0.1016	0.0816	0.0474
EW_{15}	0.0984	0.9958	-1.5241	-2.3287	-0.4644	0.1159	0.1798	0.0863	0.0680	0.0365
MV	0.1275	0.9359	-1.3497	-2.0868	-0.3245	0.1627	0.2646	0.1164	0.0997	0.0713
SSD	0.1297	0.9714	-1.4010	-2.0972	-0.1120	0.1550	0.2585	0.1181	0.1007	0.0712
SBK_2	0.1278	0.9306	-1.3429	-2.0263	-0.1660	0.1577	0.2622	0.1173	0.1016	0.0749
SBK_3	0.1253	0.8999	-1.2959	-2.0095	-0.3340	0.1631	0.2658	0.1154	0.1003	0.0746
SBK_4	0.1223	0.8422	-1.2042	-1.8841	-0.3479	0.1610	0.2640	0.1139	0.1012	0.0797
SCTSD	0.1318	0.9886	-1.4321	-2.1346	-0.0912	0.1552	0.2587	0.1197	0.1016	0.0710
ASSD (τ-Mean)	0.1492	1.2195	-1.7593	-2.6422	-0.0755	0.1430	0.2355	0.1307	0.1027	0.0550
ASSD (τ-Sharpe)	0.1587	1.3026	-1.8778	-2.8060	-0.0600	0.1353	0.2205	0.1379	0.1067	0.0537
ASSD (τ-CE5)	0.1569	1.3025	-1.8736	-2.8200	-0.0605	0.1401	0.2304	0.1355	0.1031	0.0479

Notes: Shown are the average of performance measures across 377 formation periods for portfolios enhanced by different criteria. MKT represents the benchmark stock market, EW_m represents equal-weighted combination of Top-m industries, and SBK_m represents the stochastic bounding with reference set K_m. In terms of the performance measures, Mean and Std are mean and standard deviation of daily excess return over T-bill. VaR is 5% value at risk, which is the maximum possible loss when 5% percent of the left tail of the distribution is ignored. CVaR is 5% conditional value at risk, which equals the expected return on condition that the realized return belongs to the worst 5% of the distribution. Skew represents the skewness of daily excess return over T-bill. Sharpe ratio and Sortino ratio are risk-adjusted performance measures. The last three columns present the certainty equivalent for constant risk aversion utility functions with degree of relative risk aversion equal 2, 5 and 10 respectively.

Table 5 Daily out-of-sample performance based on excess over T-bill

Strategy	Mean	Std	VaR	CVaR	Skew	Sharpe	Sortino	CE2	CE5	CE10
MKT	0.0298	0.9067	-1.4004	-2.2893	-0.2214	0.0642	0.1210	0.0185	0.0014	-0.0275
EW_5	0.0381	0.9700	-1.4990	-2.4262	-0.1730	0.0680	0.1342	0.0258	0.0072	-0.0242
EW_{10}	0.0361	0.9289	-1.4381	-2.3493	-0.2441	0.0712	0.1397	0.0246	0.0072	-0.0223
EW_{15}	0.0351	0.9180	-1.4166	-2.3354	-0.2714	0.0713	0.1398	0.0237	0.0064	-0.0229
MV	0.0524	0.9668	-1.4801	-2.4674	-0.2796	0.0842	0.1617	0.0400	0.0212	-0.0109
SSD	0.0531	1.0014	-1.5172	-2.5336	-0.2106	0.0796	0.1543	0.0401	0.0204	-0.0133
SBK_2	0.0526	0.9670	-1.4764	-2.4564	-0.2246	0.0819	0.1595	0.0406	0.0224	-0.0088
SBK_3	0.0510	0.9363	-1.4424	-2.3979	-0.2821	0.0853	0.1635	0.0395	0.0221	-0.0078
SBK_4	0.0485	0.8928	-1.3747	-2.3032	-0.2826	0.0876	0.1667	0.0380	0.0221	-0.0053
SCTSD	0.0532	1.0174	-1.5432	-2.5721	-0.2076	0.0786	0.1526	0.0397	0.0193	-0.0152
ASSD (τ-Mean)	0.0573	1.2367	-1.8677	-3.1031	-0.1458	0.0731	0.1422	0.0368	0.0059	-0.0471
ASSD (τ-Sharpe)	0.0593	1.3096	-1.9795	-3.2730	-0.1119	0.0683	0.1341	0.0373	0.0039	-0.0528
ASSD (τ-CE5)	0.0581	1.3051	-1.9700	-3.2703	-0.1383	0.0692	0.1351	0.0356	0.0015	-0.0567

Notes: Shown are the performance measures of out-of-sample daily excess return over free-risk T-bill using data from January 1928 to March 2022 for portfolios enhanced by different criteria. MKT represents the benchmark stock market, EW_m represents equal-weighted combination of Top-m industries, and SBK_m represents the stochastic bounding with reference set K_m. In terms of the performance measures, Mean and Std are mean and standard deviation of daily excess return over T-bill. VaR is 5% value at risk, which is the maximum possible loss when 5% percent of the left tail of the distribution is ignored. CVaR is 5% conditional value at risk, which equals the expected return on condition that the realized return belongs to the worst 5% of the distribution. Skew represents the skewness of daily excess return over T-bill. Sharpe ratio and Sortino ratio are risk-adjusted performance measures. The last three columns present the certainty equivalent for constant risk aversion utility functions with degree of relative risk aversion equal 2, 5 and 10 respectively.

In the formation periods (see Table 4), the ASSD strategies stand out with notable characteristics in comparison to existing strategies. They show higher mean returns, accompanied by increased volatility as reflected by elevated standard deviations. Furthermore, the ASSD strategies generally present higher 5% value at risk (VaR) and conditional value at risk (CVaR) compared to other strategies, indicating a heightened likelihood of encountering more substantial losses during extreme market downturns. This pattern remains consistent when observed across holding periods (see Table 5). Among the three ASSD procedures, the ASSD strategy with optimized τ based on the Sharpe ratio exhibits larger means, both in-sample and out-of-sample.

To further compare the risk-return performances of portfolios, we present several commonly used risk-adjusted performance measures in Table 4 and Table 5: Sharpe ratio and Sortino ratio[6]. We find that while the ASSD strategies are generally comparable to the heuristic approaches (EW_5, EW_{10} and EW_{15}), they fall short when compared to strategies like MV, SSD, SBK and SCTSD across both in-sample and out-of-sample assessments. This discrepancy in performance metrics is crucial as it signifies potential differences in risk-return trade-offs between the ASSD and the other strategies. The lower Sharpe and Sortino ratios for ASSD portfolios indicate that they might offer relatively lower returns per unit of risk or fail to adequately manage downside risk compared to these alternative strategies.

Another important measure for assessing risk-adjusted performance is the certainty equivalent (CE), indicating the specific return level that provides the same expected utility as a risky portfolio. Here, we calculate CE using constant RRA utility functions, considering RRA values of 2, 5, and 10. At lower levels of risk aversion (CE2 and CE5), the ASSD strategies demonstrate higher certainty equivalent values in-sample. However, when evaluating out-of-sample CE5 and CE10, the ASSD strategies generally exhibit poorer performance compared to alternative strategies. This disparity between in-sample and out-of-sample CE values for ASSD strategies suggests that while they might have offered superior utility for risk-averse investors in historical data, their performance weakens when assessed using new data, particularly at higher levels of risk aversion. Notably, the out-of-sample CE values of all the considered portfolios at the RRA of 10 are negative, indicating that extremely risk-averse decision-makers are inclined to avoid the risky excess return provided by the enhanced portfolios.

The aforementioned analysis highlights the importance of evaluating ASSD-enhanced portfolios in terms of the excess over T-bill, taking into account both the potential for higher returns and the associated higher risks. On the other hand, when evaluating investment strategies, exploring their

[6] Sharpe ratio and Sortino ratio measure the expected excess return of a portfolio over a risk-free asset per unit of risk, using standard deviation and downside deviation as the measure of risk respectively. We adopt the one-month Treasury bill rate (from Ibbotson Associates) as risk-free rate.

performance in terms of excess returns over the benchmark, which in this context is the stock market return, becomes important. As demonstrated in Table 6, portfolios optimized using ASSD not only showcase increased mean returns but also demonstrate superior risk-adjusted performance, as evidenced by improved information and modified Sortino ratios akin to the Sharpe and Sortino ratios in the case of excess over T-bill. This outcome could be attributed to the ASSD approach, particularly when incorporating a significant proportion of moderate τ values, thereby imposing relatively looser constraints over the benchmark in the optimization process. Consequently, optimized portfolios might better encapsulate the underlying dynamics of the market. This argument is also partially supported by the diminishing performance of the stochastic bounding strategies as the reference set expands (from K_2 to K_4). In practical investment scenarios, individual preferences for distinct portfolios may vary based on investors' risk tolerance levels, with higher risk-tolerant investors leaning towards market-based excess returns, while more risk-averse investors might find T-bill-based excess returns more aligned with their investment goals and risk tolerance thresholds.

5.3. Further investigation of out-of-sample performance

We will delve deeper into the out-of-sample performance, a more practical consideration for investors. Figure 6 illustrates the cumulative performance of the enhanced portfolios by strategies including equal-weighted, MV, SBK, SSD, SCTSC, and ASSD, throughout the holding period from January 1928 to March 2022. To ensure clarity, among the equal-weighted, SBK, and ASSD strategies, we depict EW_5, SBK_4, and ASSD (τ-Sharpe) as they achieve the highest cumulative values. As seen, strategies employing ASSD, SCTSD, SSD, and MV exhibit significant outperformance compared to SBK, EW, and the benchmark. By March 2022, the ASSD strategy accumulates a value approximately 406.9 times higher than the benchmark, emphasizing the considerable potential of these advanced portfolio optimization approaches. When comparing ASSD, SCTSD, SSD, and MV, the ASSD strategy displays higher volatility in its performance trajectory. Initially lagging behind other criteria before 1943, it indicates a slower start or potentially less effectiveness during that period. Between 1943 and 1998, the ASSD strategy closely intersects and competes

Table 6 Daily out-of-sample performance based on excess over benchmark

Strategy	Mean	Std	VaR	CVaR	Skew	Information	Sortino	CE2	CE5	CE10
EW_5	0.0083	0.4466	-0.6808	-1.0829	-0.0143	-0.0144	-0.0010	0.0058	0.0020	-0.0044
EW_{10}	0.0063	0.3347	-0.5165	-0.8097	-0.0149	-0.0267	-0.0052	0.0049	0.0028	-0.0007
EW_{15}	0.0053	0.2894	-0.4432	-0.7056	-0.0268	-0.0348	-0.0074	0.0042	0.0026	0.0000
MV	0.0226	0.5869	-0.9026	-1.4377	-0.0642	0.0161	0.0241	0.0177	0.0105	-0.0018
SSD	0.0233	0.6210	-0.9415	-1.5126	-0.0414	0.0154	0.0253	0.0180	0.0101	-0.0034
SBK_2	0.0228	0.5843	-0.8824	-1.4330	-0.0688	0.0158	0.0248	0.0184	0.0117	0.0005
SBK_3	0.0212	0.5511	-0.8417	-1.3612	-0.0893	0.0148	0.0211	0.0171	0.0110	0.0008
SBK_4	0.0187	0.5766	-0.8925	-1.4346	-0.0985	0.0083	0.0124	0.0143	0.0076	-0.0036
SCTSD	0.0234	0.6424	-0.9726	-1.5670	-0.0333	0.0155	0.0256	0.0176	0.0090	-0.0057
ASSD (τ-Mean)	0.0275	0.8600	-1.2932	-2.0795	0.0455	0.0187	0.0333	0.0164	-0.0004	-0.0285
ASSD (τ-Sharpe)	0.0295	0.9339	-1.4115	-2.2626	0.0596	0.0166	0.0335	0.0174	-0.0009	-0.0315
ASSD (τ-CE5)	0.0283	0.9279	-1.3967	-2.2441	0.0673	0.0196	0.0341	0.0157	-0.0033	-0.0353

Notes: Shown are the performance measures of out-of-sample daily excess return over benchmark using data from January 1928 to March 2022 for portfolios enhanced by different criteria. MKT represents the benchmark stock market, EW_m represents equal-weighted combination of Top-m industries, and SBK_m represents the stochastic bounding with reference set K_m. In terms of the performance measures, Mean and Std are mean and standard deviation of daily excess return over the T-bill. VaR is 5% value at risk, which is the maximum possible loss when 5% percent of the left tail of the distribution is ignored. CVaR is 5% conditional value at risk, which equals the expected return on condition that the realized return belongs to the worst 5% of the distribution. Skew represents the skewness of daily excess return over the benchmark. Information ratio and modified Sortino ratio are risk-adjusted performance measures. The last three columns present the certainty equivalent for constant risk aversion utility functions with degree of relative risk aversion equal 2, 5 and 10 respectively.

with other strategies, demonstrating a phase of comparable performance and convergence. After 1998, the ASSD strategy consistently outperforms its counterparts, demonstrating sustained superior performance. By the end of March 2022, the ASSD-enhanced portfolio is at least 1.5 times more valuable than alternative portfolios.

Some other noteworthy out-of-sample performance measures are presented in Figure 7. Figure 7(a) displays the box plots representing the daily excess returns over free T-bill across different strategies. The return distributions appear to be symmetric overall, with the ASSD strategy exhibiting larger volatilities, aligning with our earlier observations. Moreover, Figure 7(b) illustrates the maximum drawdown during the out-of-sample period. The ASSD strategy experiences a smaller maximum drawdown compared to equal-weighted strategies but larger than other alternatives, further confirming the increased volatility of the ASSD-enhanced portfolios. Additionally, Figures 7(c) and 7(d) showcase the spreads of the number of industries and turnover across 377

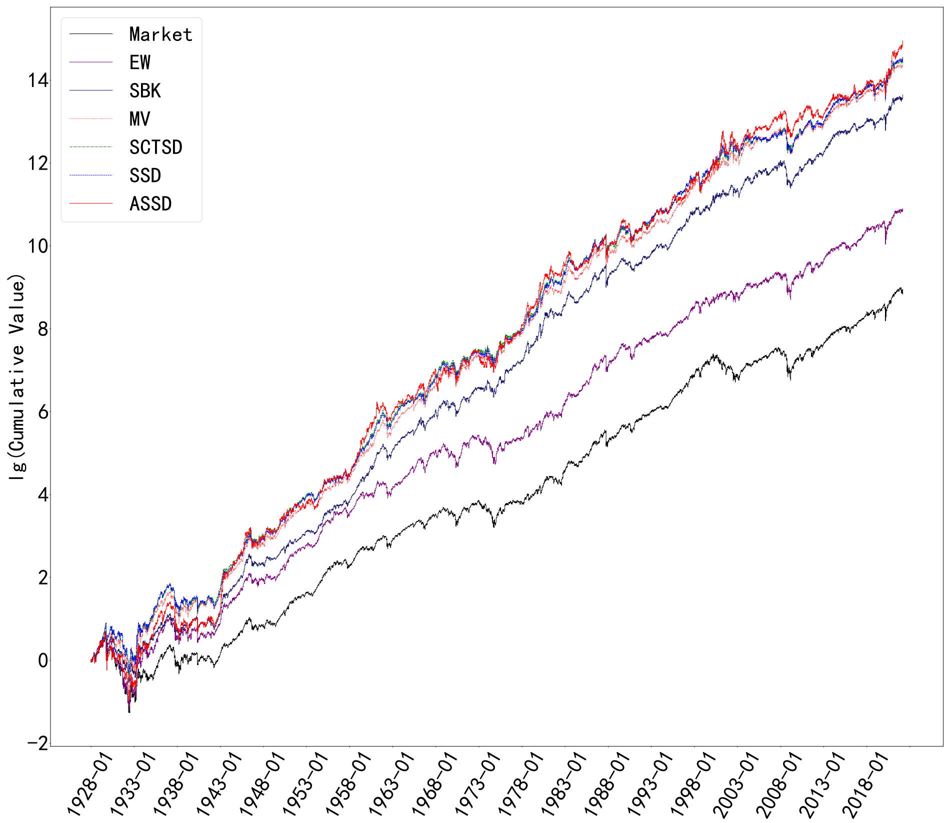
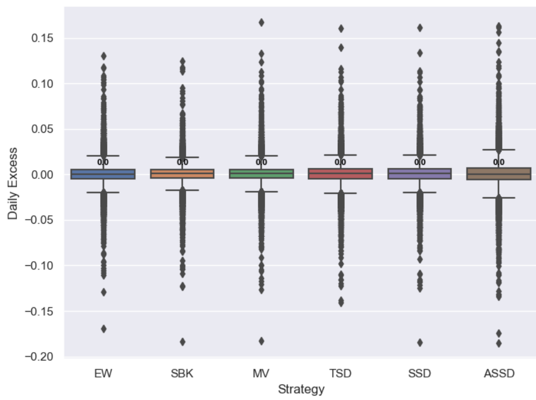


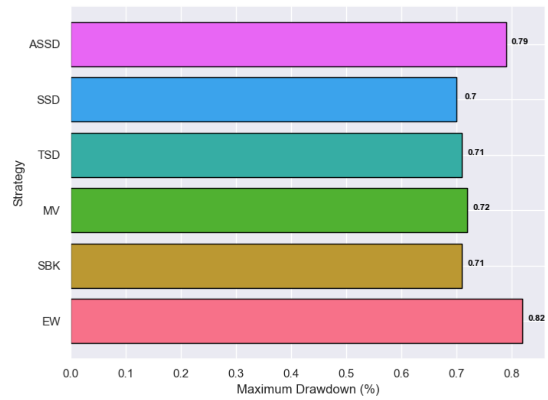
Figure 6 The solid black line represents the cumulative value of 1 unit of fund invested in the benchmark portfolio from January 1928. Additional lines denote portfolios employing various momentum-based strategies, including equal-weighted (EW), MV, SBK, SSD, SCTSD, and ASSD. Among the equal-weighted, SBK, and ASSD strategies, the graph highlights EW$_5$, SBK$_4$, and ASSD (τ-Sharpe) due to their achievement of the highest cumulative values. The graph is presented using a logarithmic scale.

out-of-sample periods. It is evident that the ASSD-enhanced portfolio tends to be less diversified and exhibits higher turnover compared to the SBK, SSD, MV, and SCTSD enhanced portfolios.

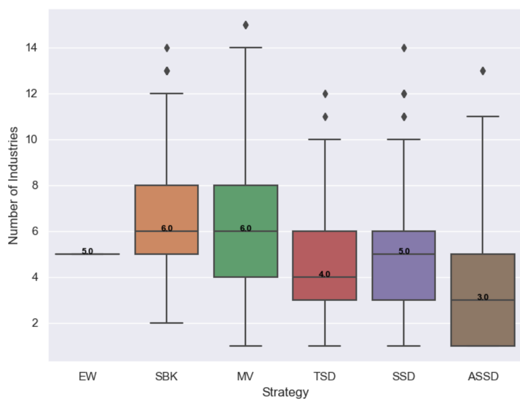
An important advantage of ASSD-enhanced portfolios lies in their higher excess returns compared to portfolios based on other strategies. This is evidenced by the increased expected value of daily excess returns, depicted in the first column of Table 4 and Table 5, and also in the cumulative returns shown in Figure 6. As daily excess returns tend to vary (as depicted in Figure 7(a)), we seek to ascertain the statistical significance of the observed out-of-sample improvement in expected excess returns through a paired t-test. Our null hypothesis posits that the mean of daily



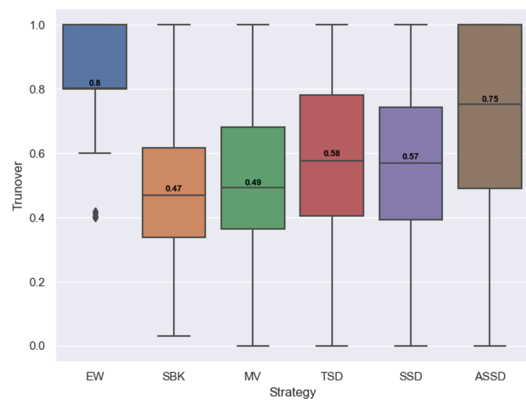
(a) Daily Excess



(b) Maximum Drawdown



(c) Number of Industries



(d) Turnover

Figure 7 Out-of-sample performance measures. Bars denote portfolios employing various momentum-based strategies, including equal-weighted (EW), MV-enhanced, SBK-enhanced, SSD-enhanced, SCTSD-enhanced, and ASSD-enhanced. Among the equal-weighted, SBK-enhanced, and ASSD-enhanced strategies, the graph highlights EW$_5$, SBK$_4$, and ASSD (τ-Sharpe) due to their achievement of the highest cumulative values.

excess returns of the ASSD-enhanced portfolio is not greater than that of portfolios based on other strategies. Employing out-of-sample return data spanning the entire holding period from 1928 to 2022, Table 7 reports the corresponding p-values obtained from the test results. The analysis indicates that the ASSD-enhanced portfolio, especially based on τ-Sharpe, consistently demonstrates higher mean returns compared to all other alternatives at the 0.1 significance level. The only exception is observed with the SCTSD, where the associated p-value slightly exceeds 0.1.

Conversely, portfolios enhanced by τ-Mean and τ-CE5 exhibit performance superiority compared to the benchmark and heuristic rules based on top industries. However, their differences with the MV-based and SD-based rules are not statistically significant.

Table 7 p**-values of tests on mean of out-of-sample daily excess.**

	ASSD (τ-Mean)	ASSD (τ-Sharpe)	ASSD (τ-CE5)
MKT	0.000	0.000	0.000
EW_5	0.003	0.002	0.003
EW_{10}	0.001	0.000	0.001
EW_{15}	0.000	0.000	0.001
MV	0.134	0.072	0.119
SSD	0.184	0.099	0.161
SBK_2	0.153	0.084	0.136
SBK_3	0.105	0.057	0.096
SBK_4	0.047	0.024	0.043
SCTSD	0.193	0.105	0.169

Note: The null hypothesis is that the expected value of out-of-sample daily excess of τ-ASSD enhanced portfolio is not greater than that of portfolios enhanced by other strategies.

The primary aim of our models is to identify portfolios that dominate the benchmark through ASSD. As per the propositions and theorems outlined in previous sections, ASSD is ensured to hold in the formation periods. However, it remains both interesting and essential to analyze whether the identified portfolio continues to dominate the benchmark in the holding periods. On the other hand, it is also intriguing to investigate whether portfolios enhanced by other strategies exhibit ASSD dominance in the holding periods. In this study, we adhere to the test procedures proposed by Guo et al. (2015) to assess the τ-ASSD in the holding periods, utilizing the optimized τ for each specific holding period. The percentage of non-ASSD dominance scenarios over the benchmark at the significance level of 0.05 is outlined in Table 8. Notably, approximately 6% of ASSD-enhanced portfolios display violations of the ASSD rule during the holding periods, while only 1% of portfolios enhanced by MV, SSD, SCTSD, and SBK strategies exhibit such violations. These findings are consistent with the results obtained from the out-of-sample performance evaluation using excess over the benchmark, as illustrated in Table 6. The relatively higher percentage of ASSD-enhanced portfolios violating the ASSD rule during the holding periods can be attributed to the moderate constraint imposed by ASSD with moderate values of τ in the optimization process, rendering the corresponding ASSD dominance more susceptible to being violated compared to other strategies.

Table 8 Percentage of no ASSD dominance in holding periods

EW_5	EW_{10}	EW_{15}	MV	SSD	SBK_2	SBK_3	SBK_4	SCTSD	ASSD (τ-Mean)	ASSD (τ-Sharpe)	ASSD (τ-CE5)
0.5	0.0	0.0	1.6	1.3	1.3	1.3	1.1	1.3	5.6	6.4	6.4

Notes: Shown are the percentages of no τ-ASSD dominance by strategies including equal-weighted (EW), MV-enhanced, SBK-enhanced, SSD-enhanced, SCTSD-enhanced, and ASSD-enhanced, utilizing the optimized τ in each holding period. The detailed test procedures can be found in Guo et al. (2015) and the critical values are obtained by bootstrap at significance level 0.05.

At last, we investigate the effects of return frequency on the out-of-sample performance. Specifically, we examine both weekly and monthly returns and the results are provided in Table 9. In consistent with the daily performance, the ASSD-enhanced portfolios continue to show superior performance in terms of average returns, albeit with escalated variations and risks. One significant observation is the considerable variance noticed in the weekly performance of ASSD, SCTSD, and MV-enhanced portfolios. Given their heightened risk and considerable fluctuations, employing these strategies with weekly data might pose significant risks. Conversely, the equal-weighted heuristics and stochastic bounding approaches appear as comparatively better options within the weekly data context.

6. Discussion

In this study, we introduce a straightforward data-driven approach to select τ by partitioning each formation period into training and test periods. This method provides a notable advantage over fixed τ strategies by dynamically adjusting τ during each formation period. This adaptability enables the method to more accurately adapt to shifts in investor preferences over time, resulting in a more precise reflection of the ASSD relationship. Among the three different selection strategies observed, τ-Sharpe generally demonstrates the largest out-of-sample returns, while τ-Mean shows the most favorable risk-adjusted performance. τ-CE5, on the other hand, displays intermediate

Table 9 Weekly and monthly out-of-sample performance based on excess over T-bill

	Weekly					Monthly				
	Mean	Std	VaR	Sortino	CE2	Mean	Std	VaR	Sortino	CE2
MKT	0.1500	2.4632	-3.7840	0.0770	0.0886	0.6675	5.3593	-7.9000	0.1682	0.3779
EW_5	0.3179	4.8004	-6.9710	0.0791	0.0751	1.1765	6.5111	-9.4400	0.2457	0.7389
EW_{10}	0.5056	7.5526	-11.037	0.0789	-0.1279	1.1765	6.5111	-9.4400	0.2457	0.7389
EW_{15}	0.7570	10.382	-14.956	0.0887	-0.5051	1.0201	5.7032	-8.3660	0.2343	0.6841
MV	1.8281	21.509	-31.897	0.1071	-2.0362	1.1069	6.2449	-8.8609	0.2390	0.7065
SSD	0.2984	3.8929	-5.5530	0.0990	0.1447	1.4503	7.7768	-10.994	0.2755	0.8393
SBK_2	0.3105	4.0956	-5.9320	0.0978	0.1396	1.3762	7.3152	-10.292	0.2740	0.8337
SBK_3	0.3360	4.8915	-6.7569	0.0842	0.0824	1.3762	7.3152	-10.292	0.2740	0.8337
SBK_4	0.2844	4.3750	-6.1109	0.0805	0.0878	1.2579	7.0983	-10.543	0.2500	0.7431
SCTSD	1.8688	21.693	-32.374	0.1086	-2.6022	1.4630	7.8371	-11.0400	0.2762	0.8431
ASSD (τ-Mean)	1.9272	21.431	-30.443	0.1086	-3.4878	1.5072	8.9546	-13.110	0.2482	0.6945
ASSD (τ-Sharpe)	1.9823	21.509	-30.509	0.1120	-3.6875	1.4829	9.0407	-13.295	0.2385	0.6500
ASSD (τ-CE5)	1.9379	21.739	-31.209	0.1088	-4.4697	1.5479	9.0818	-13.295	0.2524	0.7136

Notes: Shown are the performance measures of out-of-sample daily excess return over free-risk T-bill using data from January 1928 to March 2022 for portfolios enhanced by different criteria. MKT represents the benchmark stock market, EW_m represents equal-weighted combination of Top-m industries, and SBK_m represents the stochastic bounding with reference set K_m. In terms of the performance measures, Mean and Std are mean and standard deviation of daily excess return over T-bill. VaR (5%) is 5% value at risk, which is the maximum possible loss when 5% percent of the left tail of the distribution is ignored. Sortino ratio is a risk-adjusted performance measure. CE2 represents the certainty equivalent for constant risk aversion utility functions with degree of relative risk aversion equal 2.

performance characteristics. The practical selection among these three methods relies on the preferences of the decision makers. It is also worth noting that alternative criteria for selecting τ, such as maximizing the Sortino ratio or considering other CE values in the 3-month testing period, could also be considered in practice. Moreover, it is essential to acknowledge that the division of the formation period remains somewhat subjective, and there exists the possibility of developing more objective methods in future studies.

From our numerical investigations, it is evident that ASSD strategies generally outperform other strategies concerning returns. However, they also exhibit higher volatility and downside risks compared to MV, SBK, and other SD-based strategies. This might be attributed to relatively larger out-of-sample ASSD violations. To mitigate out-of-sample downside risk, one potential avenue is through robust optimization techniques that focus on minimizing downside risk measures. In our empirical application, we concentrated on a commonly used linear objective function aimed at maximizing expected return. Nonetheless, the preferences of investors can be catered to by adopting various other objective functions, linear or non-linear in nature. These could include minimizing left semi-variance, VaR, or CVaR, or maximizing metrics like Sharpe ratio or CE, as widely seen in existing portfolio optimization literature (for example, refer to Ghaoui et al. (2003), Quaranta and Zaffaroni (2008), Hodder et al. (2015)). Our algorithms and methods are adaptable to solving optimization problems with ASSD constraints, regardless of the objective functions used. However, addressing computational challenges arising from complex objective functions might require specific modifications. Additionally, more advanced optimization approaches like distributionally robust optimization can be adopted to consider a set of potential asset return distributions (Liesiö et al. (2020), Peng and Delage (2022)).

To assess the ASSD relation in the out-of-sample periods, we adopt the methodologies outlined in Guo et al. (2015). While hypothesis tests for SD have been extensively studied in the research community (e.g., Anderson 1996, Davidson and Duclos 2000, Barrett and Donald 2003, Linton et al. 2005, Donald and Hsu 2016, Linton et al. 2023, Beare and Clarke 2022, Lee et al. 2023), the literature on hypothesis testing for ASD remains relatively limited (Whang 2019, sec 5.4). The primary challenge lies in formulating a suitable test statistic and determining its distribution. Even though Guo et al. (2015) introduced consistent test procedures, the statistical power of these procedures has yet to be thoroughly validated. This may lead to high Type II error rates, consequently resulting in small ASSD violation percentages observed in Table 8. We believe addressing this aspect would require substantial efforts to develop a specialized test procedure tailored for more accurate ASSD assessments.

This study primarily delves into the distributional conditions of ASSD. Leveraging the properties of LPM, we introduce the concept of approximate ASSD and devise algorithms for portfolio optimization under ASSD constraints. These algorithms resolve the complex and intractable optimization problems through an iterative approach, transforming them into tractable QP and QCPs.

While ASSD can be interpreted from the perspective of preference conditions, it would be intriguing for future research to explore solving methods and analyze the optimality of this problem in the context of expected utility. This parallels existing studies focusing on optimization with SSD constraints (e.g., Dentcheva and Ruszczyński (2003, 2004a, 2006), Haskell et al. (2017)).

The computational challenges associated with optimization under ASSD constraints have limited the proposals in the literature. Our work addresses this challenge for the first time in the context of portfolio optimization. Although our focus lies in portfolio optimization, the algorithms and methodologies developed herein can be extended to address general optimization problems incorporating ASSD constraints. Furthermore, while our work primarily emphasizes the ASSD proposed by Leshno and Levy (2002) and Tzeng et al. (2013), these models and algorithms can be broadened to accommodate other formulations of ASSD, such as those presented by Luo and Tan (2020), through suitable modifications to the dominance constraints in optimization models. Future investigations might explore alternative degrees of ASD (Tsetlin et al. 2015, Liu and Meyer 2021) and relaxed formulations of SD (Müller et al. 2017, Huang et al. 2020) proposed in the existing literature.

7. Conclusion

The landscape of portfolio optimization has witnessed extensive research on SSD, yet its stringent nature incorporating extreme and 'unrealistic' utility functions has posed limitations for many decision makers. To counter this constraint, ASSD emerged as a promising solution in the literature. ASSD, permitting minor deviations from SSD, offers a more practical reflection of investment practices, thereby addressing some inadequacies unexplained by the SSD framework (Leshno and Levy 2002, Bali et al. 2009, 2013). The theoretical appeal of ASSD in portfolio optimization is undeniable. Nevertheless, its computational complexity has restricted its practical application. In this study, we have successfully overcome these computational barriers and, for the first time, presented a computationally viable approach to address portfolio optimization under ASSD constraints.

We developed optimization models and algorithms to identify portfolios dominating the benchmark portfolio by ASSD. To tackle the challenge of checking the ASSD condition in portfolio optimization with an infinite feasible solution space, we discretized the portfolio return range and propose an approach using QP and QCP. Specifically, we introduced an approximation condition for ASSD, serving as a sufficient criterion for ASSD. Utilizing this condition, we presented a cutting-plane algorithm to identify a candidate ASSD-enhanced portfolio satisfying the ASSD constraint, although not guaranteed to be optimal. Subsequently, an iterative algorithm refined the candidate portfolio by enhancing the discretization scheme of the return range iteratively until obtaining an optimal portfolio. Our approach mainly employed a QP model and QCP model with a single quadratic constraint, amenable to solving with commonly available business solvers. We demonstrated that both proposed algorithms are efficiently solvable with few iterations.

The empirical investigations conducted on optimizing combinations of the Fama and French 49 industry portfolios revealed the superiority of ASSD-enhanced portfolios over those enhanced by other strategies, such as equal-weighted heuristics, MV, and other SD-based strategies, in terms of excess returns. Nevertheless, it is crucial to acknowledge that while the ASSD-enhanced portfolios exhibit higher returns, they tend to be accompanied by amplified variations and risks. Consequently, the choice of strategies significantly relies on the risk tolerance levels of investors in practical investment scenarios.

Acknowledgement

We are grateful to the editors and two anonymous reviewers for their insightful comments that have led to a substantial improvement to an earlier version of the paper. This work was partially supported by Natural Science Foundation of China (72101070), Zhejiang Provincial Natural Science Foundation of China (LY23G010001), ONR grant N00014-18-1-2122 and AFOSR grant FA9550-23-1-0182.

Appendix A: Proofs

We first present the following lemma introduced in Luo and Tan (2020), which is important in the proofs.

Lemma 1. *Suppose random variable X are bounded between a and b. Then $Var(X) + (b - \mathbb{E}_F(X))^2 = 2\int_a^b F^{(2)}(x)dx$.*

Proof of Proposition 1. The proof follows from the definition of $\hat{F}^{(2)}_{\mathcal{D},\lambda}(x)$ and the convexity of $F^{(2)}_\lambda(x)$.$\square$

Proof of Proposition 2. It follows from the definition of expected value that $\mathbb{E}_{F_\lambda} = b - \int_a^b F_\lambda(x)dx = b - F^{(2)}_\lambda(b)$ and similarly $\mathbb{E}_{\hat{F}_{\mathcal{D},\lambda}} = b - \hat{F}^{(2)}_{\mathcal{D},\lambda}(b)$. By the definition of $\hat{F}^{(2)}_{\mathcal{D},\lambda}$, we have $\hat{F}^{(2)}_{\mathcal{D},\lambda}(b) = F^{(2)}_\lambda(b)$. Hence, $\mathbb{E}_{\hat{F}_{\mathcal{D},\lambda}} = \mathbb{E}_{F_\lambda}$. From Proposition 1, $\int_a^b \hat{F}^{(2)}_{\mathcal{D},\lambda}(x)dx \geq \int_a^b F^{(2)}_\lambda(x)dx$. Together with Lemma 1, we have $Var_{\hat{F}_{\mathcal{D},\lambda}} \geq Var_{F_\lambda}$. $\square$

Proof of Proposition 3. To prove Proposition 3, we prove that $\int_{\hat{S}_2\cap[d_s,d_{s+1}]} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)dt = A_{\mathcal{D},s,\lambda}$ for all $s, s = 1, 2, ..., T_\mathcal{D} - 1$.

Note that $\hat{F}^{(2)}_{\mathcal{D},\lambda}(t)$ and $G^{(2)}(t)$ are both linear in return sub-interval $[d_s, d_{s+1}]$, which follows from their definitions. Hence, under Case 1 and Case 4 it is straightforward that $\int_{\hat{S}_2\cap[d_s,d_{s+1}]} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)dt = A_{\mathcal{D},s,\lambda}$. Next, we prove this point for Case 2 and 3.

Under Case 2, $F^{(2)}_\lambda(d_s) \leq G^{(2)}(d_s)$ and $F^{(2)}_\lambda(d_{s+1}) \geq G^{(2)}(d_{s+1})$. By calculation, it can be found that

$$\hat{S}_2 \cap [d_s, d_{s+1}] = [\frac{(d_{s+1} - d_s)(F^{(2)}_\lambda(d_s) - G^{(2)}(d_s))}{G^{(2)}(d_{s+1}) - G^{(2)}(d_s) - F^{(2)}_\lambda(d_{s+1}) + F^{(2)}_\lambda(d_s)} + d_s, d_{s+1}].$$

Following from the fact that $\hat{F}^{(2)}_{\mathcal{D},\lambda}(t)$ and $G^{(2)}(t)$ are both linear in return sub-interval $[d_s, d_{s+1}]$, we have $\int_{\hat{S}_2\cap[d_s,d_{s+1}]} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)dt$ equals the area of the triangle between $F^{(2)}_\lambda(d_{s+1})$ and $G^{(2)}(d_{s+1})$ in the range of $\hat{S}_2 \cap [d_s, d_{s+1}]$, that is:

$$\begin{aligned}
&\int_{\hat{S}_2\cap[d_s,d_{s+1}]} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)dt \\
&= \frac{1}{2}(F^{(2)}_\lambda(d_{s+1}) - G^{(2)}(d_{s+1})) * (d_{s+1} - \frac{(d_{s+1} - d_s)(F^{(2)}_\lambda(d_s) - G^{(2)}(d_s))}{G^{(2)}(d_{s+1}) - G^{(2)}(d_s) - F^{(2)}_\lambda(d_{s+1}) + F^{(2)}_\lambda(d_s)} - d_s) \\
&= \frac{(d_{s+1} - d_s)\left(F^{(2)}_\lambda(d_{s+1}) - G^{(2)}(d_{s+1})\right)^2}{2\left(F^{(2)}_\lambda(d_{s+1}) - G^{(2)}(d_{s+1}) + G^{(2)}(d_s) - F^{(2)}_\lambda(d_s)\right)} \\
&= A_{\mathcal{D},s,\lambda}
\end{aligned}$$

Similar to Case 2, it can be proved that under Case 3, $\int_{\hat{S}_2\cap[d_s,d_{s+1}]} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right)dt = A_{\mathcal{D},s,\lambda}$.

In summary, $\int_{\hat{S}_2 \cap [d_s, d_{s+1}]} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt = A_{\mathcal{D},s,\lambda}$ holds under all possible cases, which shows that $\int_{\hat{S}_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt = \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}$. □

Proof of Theorem 1. If there exist a discretization scheme $\mathcal{D}$ such that X dominates Y by τ-AASSD, then:

$$\tau \leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}} + 1$$

$$\tau \leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\int_{\hat{S}_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt} + 1 \tag{12a}$$

$$\tau \leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt} + 1 \tag{12b}$$

$$\tau \leq \frac{\int_a^b \left(G^{(2)}(t) - F^{(2)}_{\lambda}(t)\right) dt}{\int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt} + 1 \tag{12c}$$

$$\tau \leq \frac{\int_{\bar{S}_2} \left(G^{(2)}(t) - F^{(2)}_{\lambda}(t)\right) dt}{\int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt} \tag{12d}$$

Equation (12a) and (12b) follows from Proposition 3 and 1 respectively. Equation (12c) follows from Lemma 1. Equation (12d) shows that F_λ dominates G by τ-ASSD. □

Proof of Proposition 4. The proof follows Lemma 1 and Proposition 2. □

Proof of Theorem 2. The proof of $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt \geq 0$ follows from Proposition 1 and 3. Next, we prove $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt \leq \delta_{\mathcal{D},\lambda}$.

$$\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt$$

$$= \int_{\hat{S}_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt - \int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt \tag{13a}$$

$$= \int_{\hat{S}_2 - S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt + \int_{S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt - \int_{S_2} \left(F^{(2)}_{\lambda}(t) - G^{(2)}(t)\right) dt \tag{13b}$$

$$= \int_{\hat{S}_2 - S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^{(2)}(t)\right) dt + \int_{S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - F^{(2)}_{\lambda}(t)\right) dt$$

$$\leq \int_{\hat{S}_2 - S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - F^{(2)}_{\lambda}(t)\right) dt + \int_{S_2} \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - F^{(2)}_{\lambda}(t) dt\right) \tag{13c}$$

$$\leq \int_a^b \left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - F^{(2)}_{\lambda}(t)\right) dt \tag{13d}$$

$$= \delta_{\mathcal{D},\lambda} \tag{13e}$$

Equation (13a) follows from Proposition 3. Let $\hat{S}_2 - S_2 = \{t : t \in \hat{S}_2 \ \textit{and} \ t \notin S_2\}$. It follows from Proposition 1 that $F^{(2)}_{\lambda}(t) \geq G^{(2)}(t)$ always implies $\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) \geq G^{(2)}(t)$. Hence, S_2 is a subset of $\hat{S}_2$, that is, $S_2 \subseteq \hat{S}_2$. This indicates that Equation (13b) holds. Equation (13c) follows from the fact that $\hat{S}_2 - S_2 \in \bar{S}_2$ where $G^{(2)}(t) \geq F_{\lambda}^{(2)}(t)$. Equation (13d) follows from Proposition 1. Equation (13e) follows from Proposition 4. $\square$

Proof of Corollary 1. When $\delta_{\mathcal{D},\lambda} \to 0$, it follows from Proposition 1 and 4 that $\hat{F}_{\mathcal{D},\lambda} \to F_{\lambda}$. Moreover, it follows from Theorem 2 that when $\delta_{\mathcal{D},\lambda} \to 0$, $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} \to \int_{S_2} F_{\lambda}^{(2)}(t) - G^{(2)}(t)dt$, which implies that $\frac{Var(Y)+(b-\mathbb{E}_G(Y))^2 - Var(X) - (b-\mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda}} + 1 \to \frac{\int_{\bar{S}_2}\left(G^{(2)}(t) - F^{(2)}(t)\right)dt}{\int_{S_2}\left(F^{(2)}(t) - G^{(2)}(t)\right)dt}$. Hence, τ-AASSD approaches τ-ASSD. $\square$

Proof of Proposition 5. The proof follows from the fact that $F_{\mathcal{D}^r,\lambda}^{(2)}(t) \leq F_{\mathcal{D},\lambda}^{(2)}(t), \forall t \in [a,b]$, which follows from the convexity of $F_{\lambda}^{(2)}(t)$. $\square$

Proof of Proposition 6. Proposition 6 follows from the definition of AASSD and Proposition 5. $\square$

Proof of Theorem 3. Note that

$$S_{2,s,\lambda} = S_{2,s,\lambda} \cap [d_s, d_{s+1}] = S_{2,s,\lambda} \cap (S_{2,s,\lambda'} \cup \bar{S}_{2,s,\lambda'}) = S_{2,s,\lambda} \cap S_{2,s,\lambda'} + S_{2,s,\lambda} \cap \bar{S}_{2,s,\lambda'}. \tag{14}$$

It follows from Definition 2 and the proof of Proposition 3 that:

$$\begin{aligned}
A_{\mathcal{D},s,\lambda} &= \int_{S_{2,s,\lambda}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt \\
&= \int_{S_{2,s,\lambda} \cap S_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt + \int_{S_{2,s,\lambda} \cap \bar{S}_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt && \text{(15a)} \\
&\geq \left(\int_{S_{2,s,\lambda'} \cap (S_{2,s,\lambda} + \bar{S}_{2,s,\lambda})} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'} \cap (S_{2,s,\lambda} + \bar{S}_{2,s,\lambda})} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ \\
& && \text{(15b)} \\
&= \left(\int_{S_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ \\
&= \left(\int_{S_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - \hat{F}_{\mathcal{D},\lambda'}^{(2)}(t) + \hat{F}_{\mathcal{D},\lambda'}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{(2)}(t)\right) dt\right)^+ \\
&= \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - \hat{F}_{\mathcal{D},\lambda'}^{2}(t)\right) dt\right)^+ + \left(\int_{\bar{S}_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) - G^{2}(t)\right) dt\right)^+ && \text{(15c)}
\end{aligned}$$

Equation (15a) follows from Equation (14). The first term in Equation (15b) follows from the fact that $\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) > G^{(2)}(t), \forall t \in S_{2,s,\lambda}$ and $\hat{F}_{\mathcal{D},\lambda}^{(2)}(t) \leq G^{(2)}(t), \forall t \in \bar{S}_{2,s,\lambda}$. The second term in Equation (15b) holds similarly. Equation (15c) follows from the fact that $A_{\mathcal{D},s,\lambda'} = \int_{S_{2,s,\lambda'}} \left(\hat{F}_{\mathcal{D},\lambda'}^{(2)}(t) - G^{(2)}(t)\right) dt$, which is shown in the proof of Proposition 3. $\square$

Proof of Corollary 2. The proof of Corollary 2 follows from the proof Theorem 3 with Equation (15b) to be equality rater than inequality. $\square$

Proof of Theorem 4. Given $S_{2,s,\lambda'} = [\underline{v}_s, \bar{v}_s]$, we have:

$$\begin{aligned}
&\int_{S_{2,s,\lambda'}} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t)dt \\
=&\int_{\underline{v}_s}^{\bar{v}_s} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t)dt \\
=&\int_{\underline{v}_s}^{\bar{v}_s} \left(F^{(2)}_{\lambda}(d_s) + \frac{F^{(2)}_{\lambda}(d_{s+1}) - F^{(2)}_{\lambda}(d_s)}{d_{s+1}-d_s}(x-d_s)\right)dx \qquad (16a)\\
=&\left(F^{(2)}_{\lambda}(d_s)\left(2d_{s+1}-\bar{v}_s-\underline{v}_s\right) + F^{(2)}_{\lambda}(d_{s+1})\left(\bar{v}_s+\underline{v}_s-2d_s\right)\right)\frac{\bar{v}_s-\underline{v}_s}{2(d_{s+1}-d_s)}. \qquad (16b)
\end{aligned}$$

Equation (16a) follows from the the definition of $\hat{F}^{(2)}_{\mathcal{D},\lambda}$, which implies that $\hat{F}^{(2)}_{\mathcal{D},\lambda}(t)$ is linear in $[d_s, d_{s+1}]$ and $\hat{F}^{(2)}_{\mathcal{D},\lambda}(d_s) = F^{(2)}_{\mathcal{D},\lambda}(d_s), \forall s = 1,2,...,T_{\mathcal{D}}$. It follows from Constraint (6c) that:

$$\begin{aligned}
h_{1,s} \geq &A_{\mathcal{D},s,\lambda'} + \left(\sum_{t=1}^{T} p_t\theta_{s,t}\left(2d_{s+1}-\bar{v}_s-\underline{v}_s\right) + \sum_{t=1}^{T} p_t\theta_{s+1,t}\left(\bar{v}_s+\underline{v}_s-2d_s\right)\right)\frac{\bar{v}_s-\underline{v}_s}{2(d_{s+1}-d_s)} \\
&-\left(\hat{F}^{(2)}_{\mathcal{D},\lambda'}(\bar{v}_s) + \hat{F}^{(2)}_{\mathcal{D},\lambda'}(\underline{v}_s)\right)\frac{\bar{v}_s-\underline{v}_s}{2} \\
\geq &A_{\mathcal{D},s,\lambda'} + \left(F^{(2)}_{\lambda}(d_s)\left(2d_{s+1}-\bar{v}_s-\underline{v}_s\right) + F^{(2)}_{\lambda}(d_{s+1})\left(\bar{v}_s+\underline{v}_s-2d_s\right)\right)\frac{\bar{v}_s-\underline{v}_s}{2(d_{s+1}-d_s)} \\
&-\left(\hat{F}^{(2)}_{\mathcal{D},\lambda'}(\bar{v}_s) + \hat{F}^{(2)}_{\mathcal{D},\lambda'}(\underline{v}_s)\right)\frac{\bar{v}_s-\underline{v}_s}{2} \qquad (17a)\\
= &A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t)dt + \int_{S_{2,s,\lambda'}} \hat{F}^{2}_{\mathcal{D},\lambda'}(t)dt \qquad (17b)
\end{aligned}$$

Equation (17a) follows from Constraints (6a) and (6b) which ensure that $\sum_{t=1}^{T} p_t\theta_{s,t} \geq F^{(2)}_{\lambda}(d_s)$ and $\sum_{t=1}^{T} p_t\theta_{s+1,t} \geq F^{(2)}_{\lambda}(d_{s+1})$ (following from Problem (1)) and the facts that $2d_{s+1}-\bar{v}_s-\underline{v}_s \geq 0$ and $\bar{v}_s + \underline{v}_s - 2d_s \geq 0$. Equation (17b) follows from Equation (16b) and the fact that $\hat{F}^{2}_{\mathcal{D},\lambda'}(t)$ is linear in $[d_s, d_{s+1}]$. Following from Equation (17b) and the constraint that $h_{1,s} \geq 0$, we have:

$$h_{1,s} \geq \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}}\left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{2}_{\mathcal{D},\lambda'}(t)\right)dt\right)^{+}. \qquad (18)$$

Hence, Constraints (6a), (6b), and (6c) guarantee that $h_{1,s} \geq \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}}\left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{2}_{\mathcal{D},\lambda'}(t)\right)dt\right)^{+}$.

Similarly, it can be proved that:

$$\int_{\bar{S}_{2,s,\lambda'}} \hat{F}^{(2)}_{\mathcal{D},\lambda}(t)dt = \left(F^{(2)}_{\lambda}(d_s)\left(2d_{s+1}-\bar{c}_s-\underline{c}_s\right) + F^{(2)}_{\lambda}(d_{s+1})\left(\bar{c}_s+\underline{c}_s-2d_s\right)\right)\frac{\bar{c}_s-\underline{c}_s}{2(d_{s+1}-d_s)}, \qquad (19)$$

and Constraints (6a), (6b), and (6d) guarantee that $h_{2,s} \geq \left(\int_{\bar{S}_{2,s,\lambda'}}\left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^2(t)\right)dt\right)^{+}$. Hence, Constraint (6e) implies that:

$$A_{\mathcal{D},s,\lambda} \geq h_{1,s} + h_{2,s} \geq \left(A_{\mathcal{D},s,\lambda'} + \int_{S_{2,s,\lambda'}}\left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - \hat{F}^{2}_{\mathcal{D},\lambda'}(t)\right)dt\right)^{+} + \left(\int_{\bar{S}_{2,s,\lambda'}}\left(\hat{F}^{(2)}_{\mathcal{D},\lambda}(t) - G^2(t)\right)dt\right)^{+},$$

which completes the proof. □

Proof of Proposition 7. For statement (1), it follows from Corollary 2 that if $\lambda^{k+1} \in \Omega(\lambda, \lambda^k)$, then $A_{\mathcal{D},s,\lambda^{k+1}}$ is bounded from below by its true value in Problem (7), which guarantees that λ^{k+1} dominates Y by τ-AASSD and thus is the optimal solution of Problem (5). On the other hand, if λ^{k+1} is not the optimal solution of Problem (5), then it does not dominate Y, which implies that $A_{\mathcal{D},s,\lambda^{k+1}}$ is bounded from below by a value smaller than its true value in master problem (7). This implies that $\lambda^{k+1} \notin \Omega(\lambda, \lambda^k)$.

Statement (2) also follows from Corollary 2. For all $\lambda \in \Omega(\lambda, \lambda^k)$, $A_{\mathcal{D},s,\lambda}$ is bounded from below by its true value in Problem (7). If it does not dominate Y by τ-AASSD, then it will be excluded by Constraint (7c). This completes the proof. □

Proof of Theorem 5. It follows from Theorem 2 and the definition of $\delta_{\mathcal{D}} := \max\{\delta_{\mathcal{D},\lambda}, \forall \lambda \in \Lambda\}$ that:

$$\int_{S_2} \left(F_\lambda^{(2)}(t) - G^{(2)}(t)\right) dt \geq \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D},\lambda} \geq \sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}}. \tag{20}$$

If X dominates Y by τ-ASSD, then:

$$\begin{aligned}
\tau &\leq \frac{\int_{\bar{S}_2} \left(G^{(2)}(t) - F^{(2)}(t)\right) dt}{\int_{S_2} (F^{(2)}(t) - G^{(2)}(t))\, dt} \\
&= \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\int_{S_2} (F^{(2)}(t) - G^{(2)}(t))\, dt} + 1 \qquad (21a) \\
&\leq \frac{Var(Y) + (b - \mathbb{E}_G(Y))^2 - Var(X) - (b - \mathbb{E}_{F_\lambda}(X))^2}{2\sum_{s=1}^{T_{\mathcal{D}}-1} (A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}})} + 1 \qquad (21b)
\end{aligned}$$

Equation (21a) follows from Lemma 1. Equation (21b) follows from Equation (20) and the given fact that $\sum_{s=1}^{T_{\mathcal{D}}-1} A_{\mathcal{D},s,\lambda} - \delta_{\mathcal{D}} > 0$. □

Proof of Proposition 8. It follows from Proposition 4 that:

$$\begin{aligned}
\delta_{\mathcal{D},\lambda} &= \int_a^b \hat{F}_{\mathcal{D},\lambda}^{(2)}(x)dx - \int_a^b F_\lambda^{(2)}(x)dx \qquad (22a) \\
&= \sum_{t=1}^{T_{\mathcal{D}}-1} (\hat{F}_{\mathcal{D},\lambda}^{(2)}(d_s) + \hat{F}_{\mathcal{D},\lambda}^{(2)}(d_{s+1}))\frac{d_{s+1} - d_s}{2} - \frac{1}{2}\left(\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov(x_i, x_j) + (b - \sum_{i=1}^n \lambda_i \mu_i)^2\right) \qquad (22b) \\
&= \sum_{s=1}^{T_{\mathcal{D}}-1} \left(\sum_{t=1}^T p_t \theta_{s,t} + \sum_{t=1}^T p_t \theta_{s+1,t}\right) \frac{d_{s+1} - d_s}{2} - \frac{1}{2}\left(\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Cov(x_i, x_j) + (b - \sum_{i=1}^n \lambda_i \mu_i)^2\right) \qquad (22c)
\end{aligned}$$

where $\theta_{s,t} = \max\{d_s - \sum_{i=1}^n x_{i,t}\lambda_i, 0\}$. The first term in Equation (22b) follows from the fact that $\hat{F}_{\mathcal{D},\lambda}^{(2)}(x)$ is linear in return sub-interval $[d_s, d_{s+1}], s = 1, 2, ..., T_{\mathcal{D}} - 1$. The second term in Equation (22b) follows from Lemma 1. Equation (22c) follows from Problem (1). Following from this property and the definition that $\delta_{\mathcal{D}} := \max\{\delta_{\mathcal{D},\lambda}, \forall \lambda \in \Lambda\}$, $\delta_{\mathcal{D}}$ equals the optimal value of objective function (10a), which completes the proof.

□

Appendix B: Algorithm

Algorithm 2 Algorithm for the ASSD-enhanced portfolio optimization problem

Step 1: Initialize the discretization scheme as $\mathcal{D} = \{d_1, d_2, ...d_{T_{\mathcal{D}}}\}$. Set lower bound $LB = -\infty$ and upper bound $UB = \infty$. Define a predefined acceptable optimality gap level γ.

Step 2: Given $\mathcal{D}$, solve Problem (5) with the cutting-plane Algorithm 1.

- If Problem (5) is feasible, let λ^* denote an optimal solution of Problem (5). Set $LB = f(\lambda^*)$, i.e., the optimal objective value of Problem (5).
- Otherwise, mark LB as 'infeasible'.

Step 3: Solve Problem (10) to obtain the value of $\delta_{\mathcal{D}}$.

Step 4: Given $\delta_{\mathcal{D}}$ obtained in last step, solve Problem (11) with the cutting-plane Algorithm 1.

- If Problem (11) is feasible with λ^U as its optimal solution, set $UB = f(\lambda^U)$.
- Otherwise, mark UB as 'infeasible'.

Step 5: Analyze the result.

- If neither LB nor UB is 'infeasible', calculate the optimality gap $UB - LB$.
  - If $UB - LB \leq \gamma$, terminate and return solution λ^*.
  - Otherwise, go to the next step.
- If both LB and UB are 'infeasible', terminate and return the result that the ASSD-enhanced portfolio optimization problem is infeasible.
- Otherwise, go to next step.

Step 6: Refine the discretization scheme $\mathcal{D}$ and **go back to Step 2**.

Step 7: Repeat Step 2-6 until one of the following conditions are satisfied:

- The optimality gap of λ^* is no greater than the optimality tolerance γ.
- Both LB and UB are marked as 'infeasible', which indicates that there exists no portfolio dominating the benchmark by τ-ASSD.

Remark: In the application of this work, the discretization scheme $\mathcal{D}$ is refined in Step 6 by adding points $\frac{d_s+d_{s+1}}{2}$ into $\mathcal{D}$ for all $d_s \in \mathcal{D}$, i.e., uniformly dividing all return sub-interval $[d_s, d_{s+1}]$ into two sub-intervals. Other refinement rules like dividing some selected sub-intervals may also be adopted to further accelerate the algorithm.

References

Anderson, G. (1996), "Nonparametric tests of stochastic dominance in income distributions," *Econometrica: Journal of the Econometric Society*, 1183–1193.

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

Bali, T. G., Brown, S. J., and Demirtas, K. O. (2013), "Do hedge funds outperform stocks and bonds?" *Management Science*, 59(8), 1887–1903.

Barrett, G. F. and Donald, S. G. (2003), "Consistent tests for stochastic dominance," *Econometrica*, 71(1), 71–104.

Beare, B. K. and Clarke, J. D. (2022), "Modified Wilcoxon-Mann-Whitney tests of stochastic dominance," *arXiv preprint arXiv:2210.08892*.

Bi, H. and Zhu, W. (2022), "Nonmonotonic risk preferences over lottery comparison," *European Journal of Operational Research*, 303(3), 1458–1468.

Davidson, R. and Duclos, J.-Y. (2000), "Statistical inference for stochastic dominance and for the measurement of poverty and inequality," *Econometrica*, 68(6), 1435–1464.

Donald, S. G. and Hsu, Y.-C. (2016), "Improving the power of tests of stochastic dominance," *Econometric Reviews*, 35(4), 553–585.

Guo, X., Levy, H., and Wong, W.-K. (2015), "Consistent tests for almost stochastic dominance," *Available at SSRN 2592205*.

Han, W., Newton, D., Platanakis, E., Sutcliffe, C., and Ye, X. (2023), "On the (almost) stochastic dominance of cryptocurrency factor portfolios and implications for cryptocurrency asset pricing," *European Financial Management*, *to appear*.

Huang, Y.-C., Kan, K., Tzeng, L. Y., and Wang, K. C. (2021), "Estimating the critical parameter in almost stochastic dominance from insurance deductibles," *Management Science*, 67(8), 4742–4755.

Lee, K., Linton, O., and Whang, Y.-J. (2023), "Testing for time stochastic dominance," *Journal of Econometrics*, 235(2), 352–371.

Lee, Y.-T., Kung, K.-L., and Liu, I.-C. (2018), "Profitability and risk profile of reverse mortgages: A cross-system and cross-plan comparison," *Insurance: Mathematics and Economics*, 78, 255–266.

Levy, H. (2016), "Aging population, retirement, and risk taking," *Management Science*, 62(5), 1415–1430.

Levy, H., Leshno, M., and Leibovitch, B. (2010), "Economically relevant preferences for all observed epsilon," *Annals of Operations Research*, 176(1), 153–178.

Liesiö, J., Xu, P., and Kuosmanen, T. (2020), "Portfolio diversification based on stochastic dominance under incomplete probability information," *European Journal of Operational Research*, 286(2), 755–768.

Linton, O., Maasoumi, E., and Whang, Y.-J. (2005), "Consistent testing for stochastic dominance under general sampling schemes," *The Review of Economic Studies*, 72(3), 735–765.

Linton, O., Seo, M. H., and Whang, Y.-J. (2023), "Testing stochastic dominance with many conditioning variables," *Journal of Econometrics*, 235(2), 507–527.

Markowitz, H. (2014), "Mean–variance approximations to expected utility," *European Journal of Operational Research*, 234(2), 346–355.

Post, T. (2015), "Critical values for almost stochastic dominance based on relative risk aversion," *Available at SSRN 2666209*.

Post, T. and Kopa, M. (2017), "Portfolio choice based on third-degree stochastic dominance," *Management Science*, 63(10), 3381–3392.

Post, T. and Potì, V. (2017), "Portfolio analysis using stochastic dominance, relative entropy, and empirical likelihood," *Management Science*, 63(1), 153–165.

Whang, Y.-J. (2019), *Econometric analysis of stochastic dominance: Concepts, methods, tools, and applications*, Cambridge University Press.

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

Bali, T. G., S. J. Brown, and K. O. Demirtas (2013). Do hedge funds outperform stocks and bonds? *Management Science 59*(8), 1887–1903.

Bali, T. G., K. O. Demirtas, H. Levy, and A. Wolf (2009). Bonds versus stocks: Investors' age and risk taking. *Journal of Monetary Economics 56*(6), 817–830.

Bawa, V. S., J. N. Bodurtha Jr, M. Rao, and H. L. Suri (1985). On determination of stochastic dominance optimal sets. *The Journal of Finance 40*(2), 417–431.

Bruni, R., F. Cesarone, A. Scozzari, and F. Tardella (2017). On exact and approximate stochastic dominance strategies for portfolio selection. *European Journal of Operational Research 259*(1), 322–329.

Dentcheva, D., R. Henrion, and A. Ruszczyński (2007). Stability and sensitivity of optimization problems with first order stochastic dominance constraints. *SIAM Journal on Optimization 18*(1), 322–337.

Dentcheva, D. and W. Römisch (2013). Stability and sensitivity of stochastic dominance constrained optimization models. *SIAM Journal on Optimization 23*(3), 1672–1688.

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

Dentcheva, D. and A. Ruszczyński (2004a). Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. *Mathematical Programming 99*(2), 329–350.

Dentcheva, D. and A. Ruszczyński (2004b). Semi-infinite probabilistic optimization: first-order stochastic dominance constrain. *Optimization 53*(5-6), 583–601.

Dentcheva, D. and A. Ruszczyński (2006). Portfolio optimization with stochastic dominance constraints. *Journal of Banking & Finance 30*(2), 433–451.

Dentcheva, D. and A. Ruszczyński (2010). Inverse cutting plane methods for optimization problems with second-order stochastic dominance constraints. *Optimization 59*(3), 323–338.

Denuit, M. M., R. J. Huang, and L. Y. Tzeng (2014). Bivariate almost stochastic dominance. *Economic Theory 57*(2), 377–405.

Fábián, C. I., G. Mitra, and D. Roman (2011). Processing second-order stochastic dominance models using cutting-plane representations. *Mathematical Programming 130*(1), 33–57.

Fang, Y. and T. Post (2017). Higher-degree stochastic dominance optimality and efficiency. *European Journal of Operational Research 261*(3), 984–993.

Ghaoui, L. E., M. Oks, and F. Oustry (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research 51*(4), 543–556.

Haskell, W. B., J. G. Shanthikumar, and Z. M. Shen (2017). Primal-dual algorithms for optimization with stochastic dominance. *SIAM Journal on Optimization 27*(1), 34–66.

Hodder, J. E., J. C. Jackwerth, and O. Kolokolova (2015). Improved portfolio choice using second-order stochastic dominance. *Review of Finance 19*(4), 1623–1647.

Hu, J., T. Homem-de Mello, and S. Mehrotra (2012). Sample average approximation of stochastic dominance constrained programs. *Mathematical Programming 133*(1), 171–201.

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

Huang, Y.-C., K. Kan, L. Y. Tzeng, and K. C. Wang (2021). Estimating the critical parameter in almost stochastic dominance from insurance deductibles. *Management Science 67*(8), 4742–4755.

Kallio, M. and N. D. Hardoroudi (2018). Second-order stochastic dominance constrained portfolio optimization: Theory and computational tests. *European Journal of Operational Research 264*(2), 675–685.

Kopa, M. and P. Chovanec (2008). A second-order stochastic dominance portfolio efficiency measure. *Kybernetika 44*(2), 243–258.

Kopa, M. and T. Post (2009). A portfolio optimality test based on the first-order stochastic dominance criterion. *Journal of Financial and Quantitative Analysis 44*(5), 1103–1124.

Kopa, M. and T. Post (2015). A general test for SSD portfolio efficiency. *OR Spectrum 37*(3), 703–734.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science 50*(10), 1390–1406.

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

Levy, H. (2016). *Stochastic dominance: Investment decision making under uncertainty*, Springer.

Levy, H., M. Leshno, and B. Leibovitch (2010). Economically relevant preferences for all observed epsilon. *Annals of Operations Research 176*(1), 153–178.

Levy, M. (2009). Almost stochastic dominance and stocks for the long run. *European Journal of Operational Research 194*(1), 250–257.

Liesiö, J., P. Xu, and T. Kuosmanen (2020). Portfolio diversification based on stochastic dominance under incomplete probability information. *European Journal of Operational Research 286*(2), 755–768.

Liu, L. and J. Meyer (2021). Almost stochastic dominance: Magnitude constraints on risk aversion. *Available at SSRN 4669556*.

Lizyayev, A. and A. Ruszczyński (2012). Tractable almost stochastic dominance. *European Journal of Operational Research 218*(2), 448–455.

Longarela, I. R. (2016). A characterization of the SSD-efficient frontier of portfolio weights by means of a set of mixed-integer linear constraints. *Management Science 62*(12), 3549–3554.

Luo, C. and C. H. Tan (2020). Almost stochastic dominance for most risk-averse decision makers. *Decision Analysis 17*(2), 169–184.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance 7*(1), 77–91.

Markowitz, H. (2014). Mean–variance approximations to expected utility. *European Journal of Operational Research 234*(2), 346–355.

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

Peng, C. and E. Delage (2022). Data-driven optimization with distributionally robust second order stochastic dominance constraints. *Operations Research*, *in press*.

Post, T. (2003). Empirical tests for stochastic dominance efficiency. *The Journal of Finance 58*(5), 1905–1931.

Post, T. (2017). A review of portfolio choice based on stochastic dominance. *Available at SSRN 2925222*.

Post, T., Y. Fang, and M. Kopa (2015). Linear tests for decreasing absolute risk aversion stochastic dominance. *Management Science 61*(7), 1615–1629.

Post, T., S. Karabatı, and S. Arvanitis (2018). Portfolio optimization based on stochastic dominance and empirical likelihood. *Journal of Econometrics 206*(1), 167–186.

Post, T. and M. Kopa (2013). General linear formulations of stochastic dominance criteria. *European Journal of Operational Research 230*(2), 321–332.

Post, T. and M. Kopa (2017). Portfolio choice based on third-degree stochastic dominance. *Management Science 63*(10), 3381–3392.

Post, T. and V. Potì (2017). Portfolio analysis using stochastic dominance, relative entropy, and empirical likelihood. *Management Science 63*(1), 153–165.

Quaranta, A. G. and A. Zaffaroni (2008). Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking & Finance 32*(10), 2046–2056.

Rockafellar, R. T. and S. Uryasev (2000). Optimization of conditional value-at-risk. *Journal of Risk 2*(3), 21–41.

Roman, D., K. Darby-Dowman, and G. Mitra (2006). Portfolio construction based on stochastic dominance and target return distributions. *Mathematical Programming 108*(2), 541–569.

Rudolf, G. and A. Ruszczyński (2008). Optimization problems with second order stochastic dominance constraints: duality, compact formulations, and cut generation methods. *SIAM Journal on Optimization 19*(3), 1326–1343.

Sun, H., H. Xu, R. Meskarian, and Y. Wang (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.

Tan, C. H. (2015). Weighted almost stochastic dominance: revealing the preferences of most decision makers in the st. petersburg paradox. *Decision Analysis 12*(2), 74–80.

Tan, C. H. and C. Luo (2017). Clear preferences under partial distribution information. *Decision Analysis 14*(1), 65–73.

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

Tsetlin, I., R. L. Winkler, R. J. Huang, and L. Y. Tzeng (2015). Generalized almost stochastic dominance. *Operations Research 63*(2), 363–377.

Tzeng, L. Y., R. J. Huang, and P.-T. Shih (2013). Revisiting almost second-degree stochastic dominance. *Management Science 59*(5), 1250–1254.