PORTFOLIO OPTIMIZATION WITH RELAXATION OF
STOCHASTIC SECOND ORDER DOMINANCE CONSTRAINTS
VIA CONDITIONAL VALUE AT RISK

MENG XUE
School of Economics and Management
Nanjing University of Science and Technology
Nanjing, 210094, China

YUN SHI
Department of Applied Mathematics
The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

HAILIN SUN*
Jiangsu Provincial Key Laboratory for NSLSCS
School of Mathematical Sciences
Nanjing Normal University
Nanjing, 210023, China

(Communicated by Ken Siu)

ABSTRACT. A portfolio optimization model with relaxed second order stochastic dominance (SSD) constraints is presented. The proposed model uses Conditional Value at Risk (CVaR) constraints at probability level $\beta \in (0, 1)$ to relax SSD constraints. The relaxation is justified by theoretical convergence results based on sample average approximation (SAA) method when sample size $N \to \infty$ and CVaR probability level $\beta$ tends to 1. SAA method is used to reduce infinite number of inequalities of SSD constraints to finite ones and also to calculate the expectation value. The proposed relaxation on the SSD constraints in portfolio optimization problem is achieved when the probability level $\beta$ of CVaR takes value less than but close to 1, and the model can then be solved by cutting plane method. The performance and characteristics of the portfolios constructed by solving the proposed model are tested empirically on three sets of market data, and the experimental results are analyzed and discussed. Furthermore, it is shown that with appropriate choices of CVaR probability level $\beta$, the constructed portfolios are sparse and outperform the portfolios constructed by solving portfolio optimization problems with SSD constraints, with either index portfolios or mean-variance (MV) portfolios as benchmarks.

2020 Mathematics Subject Classification. 91G10; 90C15.

Key words and phrases. Stochastic optimization, portfolio selection, second order dominance, CVaR approximation, sample average approximation.

The work is supported by National Natural Science Foundation of China grant 11871276, 11571178 and 11571056.

* Corresponding author: Hailin Sun.
1. **Introduction.** Portfolio selection is among one of the most popular research topics in mathematical finance. Investors evaluate the performance of a portfolio by its ability of maximizing anticipated returns or minimizing cost, and often subject to additional constraints. Therefore, it is natural to construct portfolio by solving constrained optimization problems, whose modelling reflect distinctive characteristics of economical and financial selection rules or objectives chosen by the investors. One class of the constraints is the requirement of simultaneous risk control, i.e., the confinement of undesirable profit variation while achieving expected returns. This idea leads to the development of model construction whose objective is set to balance the expected return and the risk for general risk-averse investors. This so-called mean-risk model is the most widely adapted portfolio selection problem construction. More specifically, such approach seeks the portfolio that balances the expected return and the potentially hazardous uncertainty. Rooting from the mean-variance (MV) rule proposed by Markowitz [23], its modifications are among the most popular classes of mean-risk models in mathematical finance research community. To list a few, there are mean semi-variance model [24], mean-absolute deviation model [20], mean-absolute deviation-skewness model [21] and so on.

Numerous risk measures besides variance have been proposed to evaluate and quantify the risks of portfolios. For example, Value at Risk (VaR) indicates the maximum possible loss when some percentage of the left tail distribution is ignored [17]. The Conditional Value at Risk (CVaR) [28, 1] is also widely used, which is more sensitive to the shape of the tail regarding the loss distribution. In this paper, the second order stochastic dominance (SSD) is considered as the choice of risk constraint, which is a fundamental concept in decision theory and economics, and has seen many uses in portfolio selection problems, see Mosler and Scarsini [25], Whitmore and Findlay [37]. The relation between mean-semi-deviation model and second order stochastic dominance was analyzed by Ogryczak and Ruszczyński [26] for the first time and they further showed that the mean-risk model is consistent with the second order stochastic dominance for several other risk measures [27]. It is also known that SSD is equivalent to a continuum of CVaR constraints for all probability level $\beta \in (0, 1]$, see [10]. Stochastic program with stochastic dominance constraints was proposed by Dentcheva and Ruszczyński and they analyzed different aspects including optimality and duality, see [7, 8, 9]. The convergence analysis of stochastic optimization problem with SSD constraints under sample average approximation (SAA) has been studied in several papers [16, 22, 34], while many efficient methods have been proposed to solve it. For example, Sun et. al. [32] proposed a smoothing penalized SAA method to solve this problem. Homem-de-Mello and Mehrota [15] proposed a sample average cutting surface algorithm for optimization problems with multidimensional polyhedral linear SSD constraints. Based on the equivalence result between SSD and a continuum of CVaR with all probability level $\beta \in (0, 1]$, Dentcheva and Ruszczyński [11] proposed an equivalent stochastic optimization problem with inverse dominance constraint and designed an inverse cutting plane methods for solving it. Rudolf and Ruszczyński [31], Fábián et. al. [13] and Sun et. al. [33] proposed cutting plane methods for solving stochastic programs with SSD constraints.

Adaptation of stochastic dominance constrained optimization in portfolio selection problem was introduced by Dentcheva and Ruszczyński [10, 12]. Roman et. al. [29] proposed a multi-objective portfolio selection model with SSD constraints. The rationale behind the usage of stochastic dominance constraints in portfolio selection
is to find the portfolio strategy whose distribution of the random returns dominates that of a benchmark portfolio strategy. For example, when the benchmark portfolio is to be a financial index, this portfolio optimization becomes an enhanced indexation model. An enhanced indexation portfolio manages to track and outperform a stock market index with limited number of stocks, which is desired from a portfolio management prospective [38]. An enhanced indexation model based on SSD was proposed by Roman et. al. [30] and produced consistently good returns, sparsity and with minimal rebalancing required in the constructed portfolios.

In this paper, a portfolio selection model is proposed based on the stochastic optimization problem with SSD constraints. The model cooperates the SAA method and CVaR approach proposed by Anderson et. al. [2]. More specifically, the SAA method is used to overcome the complicated calculations of expected value and to handle the infinite number of inequalities resulted from SSD constraints, followed by the CVaR approximation. Furthermore, when CVaR is used with probability level $\beta < 1$, the dominance constraints are relaxed and by solving the relaxed model, a less conservative portfolio may be obtained, see motivative Example 2.1. Indeed, as can be seen in later section of empirical studies, the resulting portfolio constructed by solving the proposed model demonstrates better expected return rate without jeopardizing the control over the risk. In short, contributions of this paper are as follows.

• A relaxation model of portfolio optimization problem with SSD constraints is proposed based on SAA method and CVaR approximation, where the relaxation is achieved with the choice of probability level $\beta < 1$. This is motivated by the observation on the conservative performance of portfolios constructed by solving the SSD constrained portfolio optimization problems, and solving the relaxation model can lead to portfolios with better returns under risk control. This is called the SAA-based CVaR-SSD relaxation model, whose solution is called the CVaR-SSD portfolio.

• The convergence analysis of the SAA-based CVaR-SSD relaxation problem to the original SSD problem is presented based on theoretical results of stochastic optimization problem with SSD constraints [34] with the re-proposed conditions in convex cases and CVaR approximation [2]. The convergence is also demonstrated empirically when the CVaR probability level $\beta$ approaches 1.

• The advantages of CVaR-SSD portfolios are demonstrated empirically. The performance of the constructed portfolios with probability level $\beta < 1$ is comparable to that of the SSD portfolio, and with appropriate choices of $\beta$, the performance of CVaR-SSD portfolio is always better than the SSD portfolio with the same benchmark. Empirical studies are carried out with the benchmark portfolios being both the index portfolio and the MV portfolio for three sets of historical market data, while both CVaR-SSD portfolios and SSD portfolio outperform the corresponding benchmarks in all the data sets. In addition, the resulting CVaR-SSD portfolios are sparse.

Note that the proposed SAA-based CVaR-SSD relaxation model is related to the stochastic optimization problem with inverse dominance constraint defined in [11]. The key differences are that,

1) the stochastic optimization problem with inverse dominance constraint reformulation is based on the equivalence relation between SSD constraints and a continuum of CVaR constraints;
2) the proposed SAA-based CVaR-SSD relaxation model is a relaxation model based on SAA method and CVaR relaxation of the max function proposed in [2] at some choices of \( \beta < 1 \).

The rest of this paper is organized as follows. In Section 2, constructions of SAA-based CVaR-SSD relaxation model and the SAA method for solving portfolio optimization with both SSD and CVaR constraints are presented. The convergence properties of the proposed model to the SSD model are analyzed in Section 3. We present our empirical studies to compare portfolios constructed by solving different models in terms of their performance when applied to three sets of historical market data. The detailed description and analysis are carried out in Section 4. The concluding remarks are summarized in Section 5.

2. CVaR approximation of portfolio optimization with SSD constraints.
In this section, we introduce the SAA-based CVaR-SSD relaxation model based on the SSD constrained portfolio optimization problem. More specifically, the CVaR constraint is treated as an approximation of the SSD constraints when CVaR probability level \( \beta \) tends to 1 under SAA. Relaxation is realized by lowering the CVaR probability level \( \beta \) to a value smaller than 1.

Let \( \xi(\omega) := (\xi_1(\omega), \cdots, \xi_n(\omega))^T \) denote a random vector of return rates of \( n \) assets, where \( \xi(\cdot) : \Omega \rightarrow \mathbb{R}^n \) is defined in probability space \( \mathcal{L}(\Omega, \mathcal{F}, P) \). To ease of notation, \( \xi \) is used to denote the random vector \( \xi(\omega) \). Let \( F_1(\xi; \eta) \) denote the cumulative distribution function of a random vector \( \xi \) evaluated at \( \eta \), i.e.,

\[
F_1(\xi; \eta) := P(\xi \leq \eta),
\]

and define

\[
F_2(\xi; \eta) := \int_{-\infty}^{\eta} F_1(\xi; t) dt, \quad \eta \in \mathbb{R}.
\]

The dominance is described for two random variables \( \xi_1 \) and \( \xi_2 \), and \( \xi_1 \) is said to dominate \( \xi_2 \) in first order, denoted by \( \xi_1 \succeq_1 \xi_2 \), if

\[
F_1(\xi_1; \eta) \leq F_1(\xi_2; \eta), \forall \eta \in \mathbb{R}.
\]

Similarly, \( \xi_1 \) is said to dominate \( \xi_2 \) in second order, denoted by \( \xi_1 \succeq_2 \xi_2 \), if

\[
F_2(\xi_1; \eta) \leq F_2(\xi_2; \eta), \forall \eta \in \mathbb{R}.
\] (1)

It is easy to observe that first order stochastic dominance implies second order stochastic dominance. Furthermore, second order dominance (1) can be reformulated as

\[
\mathbb{E}_\rho[(\eta - \xi_1)_+] \leq \mathbb{E}_\rho[(\eta - \xi_2)_+], \forall \eta \in \mathbb{R},
\] (2)

where \((x)_+ := \max(0, x)\), see [9]. With this expression, the SSD can be interpreted as partial order of distribution functions. It is claimed that the advantage of SSD over other comparing criteria is that it can fully use the information of random variables, therefore it can “reasonably” choose more realistically than, e.g., MV model.

The portfolio optimization problem with SSD constraints is formulated as

\[
\begin{align*}
\max_x \quad & \mathbb{E}_\rho[x^T \xi] \\
\text{s.t.} \quad & x^T \xi \succeq_2 Y(\xi), \forall \eta \in \mathbb{R}, \\
& x \in X_0,
\end{align*}
\] (3)

where \( x \in \mathbb{R}^n \) presents the fractions of capital invested in \( n \) assets, i.e., a portfolio, and thus the set of possible portfolios is denoted as \( X_0 = \{x \in \mathbb{R}^n_+, x_1 + \cdots + x_n = 1\} \).
Note that the constraint in (3) is the second order stochastic dominance of $x^T \xi$ over the return $Y(\xi)$ of a pre-specified benchmark portfolio $\bar{x} \in \mathbb{R}^n$, i.e., $Y(\xi) := \bar{x}^T \xi$. Therefore, the aim of the portfolio construction with SSD constraints is to find a portfolio $x^*$ that solves (3) in the sense of maximizing the expected return while the SSD constraint, $(x^*)^T \xi \succeq 2 \bar{x}^T \xi$, is satisfied. For purposes of approximation and relaxation, the construction starts by reformulating the SSD constrained portfolio optimization problem (3).

Let $H(x, \eta, \xi) := (\eta - x^T \xi)_+ - (\eta - Y(\xi))_+$ and $h(x, \eta) = \mathbb{E}_P[H(x, \eta, \xi)]$, and by the reformulation result (2), problem (3) can be written as

\[
\min_x \mathbb{E}_P[-x^T \xi] \\
\text{s.t.} \quad h(x, \eta) \leq 0, \forall \eta \in \mathbb{R}, \quad x \in X_0.
\] (4)

Both (3) and (4) are stochastic semi-infinite problems. From a numerical optimization perspective, the form of the constraints in problem (4) does not provide much convenience since it does not satisfy the checkable Slater-type constraint qualification (CQ). Consequently, one often considers a relaxed form of problem (4) [7]:

\[
\min_x -\mathbb{E}_P[x^T \xi] \\
\text{s.t.} \quad h(x, \eta) \leq 0, \forall \eta \in [a, b], \quad x \in X_0.
\] (5)

This is called the SSD problem.

In order to solve (5), SAA method is applied to handle the expectation functions in both its objective and constraints. Let $\{\xi_1, \cdots, \xi_N\}$ be the independent and identically distributed (i.i.d.) samples of random vector $\xi$ and the choice of $a$ and $b$ be such that $Y(\xi) \in [a, b]$, $\forall \xi \in \Xi$, then the SAA problem of (5) is written as,

\[
\min_x -\frac{1}{N} \sum_{i=1}^N x^T \xi_i \\
\text{s.t.} \quad \frac{1}{N} \sum_{i=1}^N ((\eta_j - x^T \xi_i)_+ - (\eta_j - Y(\xi_i))_+) \leq 0, \quad j = 1, \cdots, N,
\] (6)

where $\eta_j := Y(\xi_j)$ [7]. As described in the introduction section, the modelling is based on an approximation of (5) which will later be modified in such a way that the conservative SSD constraints in (5) are relaxed. The idea is that better portfolio construction may be achieved under appropriate relaxation of the SSD constraints.

Let $h_N(x, \eta) := \frac{1}{N} \sum_{i=1}^N ((\eta - x^T \xi_i)_+ - (\eta - Y(\xi_i))_+)$, and the SSD constraints in problem (6) is written as

\[
\max_{j \in \{1, \cdots, N\}} h_N(x, \eta_j) \leq 0.
\] (7)

Focusing on the constraint function (7) and we then show that it can be approximated by a CVaR formulation. To see this, be reminded that the definition of VaR of a random variable $\zeta$ is

\[
\text{VaR}_\beta(\zeta) := \min_{\gamma \in \mathbb{R}} \{ \gamma : \text{Prob}\{\zeta \leq \gamma\} \geq \beta \},
\]

where $\beta$ is the VaR probability level. The definition of CVaR as in [28] is adapted as follows,

\[
\text{CVaR}_\beta(\zeta) := \min_{\gamma \in \mathbb{R}} \left( \gamma + \frac{1}{1 - \beta} \mathbb{E}[(\zeta - \gamma)_+] \right).
\]
Then, if $\eta$ is treated as a random variable, inequilities (7) can be expressed as
\[
\text{CVaR}_\beta(h_N(x, \eta)) = \min_{\zeta \in \mathbb{R}} (\zeta + \frac{1}{1-\beta} \mathbb{E}_{P_{\eta}}[(h_N(x, \eta) - \zeta)_+]) \leq 0, 
\] (8)
where $P_{\eta}$ denotes the discrete probability measure of random variable $\eta$. Then, problem (6) is approximated by
\[
\min_x -\frac{1}{N} \sum_{i=1}^{N} x^T \xi_i \\
\text{s.t.} \quad \text{CVaR}_\beta(h_N(x, \eta)) \leq 0, \\
x \in X_0. 
\] (9)

In addition, when we take $\eta$ such that $\text{Prob}\{\eta = Y(\xi_i)\} = \frac{1}{N}$, the constraint function of problem (9) can be expressed as,
\[
\text{CVaR}_\beta^N(h_N(x, \eta)) = \min_{\zeta \in \mathbb{R}} (\zeta + \frac{1}{1-\beta} \frac{1}{N} \sum_{j=1}^{N} (h_N(x, \eta_j) - \zeta)_+) \\
= \min_{\zeta \in \mathbb{R}} \left( \zeta + \frac{1}{1-\beta} \frac{1}{N} \sum_{j=1}^{N} \left( \frac{1}{N} \sum_{i=1}^{N} (\eta_j - x^T \xi_i)_+ - (\eta_j - Y(\xi_i))_+ \right) \right). 
\] (10)

In turn, problem (6) takes the form
\[
\min_x -\frac{1}{N} \sum_{i=1}^{N} x^T \xi_i \\
\text{s.t.} \quad \text{CVaR}_\beta^N(h_N(x, \eta)) \leq 0, \\
x \in X_0. 
\] (11)

This is called the \textit{SAA-based CVaR-SSD relaxation problem}.

Similar approximation scheme was recently considered by Anderson et. al. [2]. In the context of portfolio optimization, the SSD constraint (7) is regarded an extremely robust risk measure, and CVaR would be a relaxation of SSD, depending on the choice of probability level $\beta$. This can also be viewed from a numerical solution perspective. In that case, instead of a single sample at the extremely robust constraint (7), the CVaR approximation with specific choices of $\beta$ permits to take more samples near extremum constraint and in that way “smooth up” or “stabilize” the numerical calculation when SAA method is used.

As described above, the relaxation of the SSD constraint in the proposed model is achieved by choosing the probability level $\beta < 1$. The interpretation is that since SSD requires all choices of $\beta \in (0, 1]$ to be satisfied, it is relaxed if some inequalities of the constraints of problem (6) with index $j \in \{1, \cdots, N\}$ can be violated. This enlarges the feasible set of the original SSD problem (3) by omitting part of the support set of $\eta$. Then a less conservative solution, in comparison with the full obedience of SSD, that violates some of the inequalities may result in possibly better portfolio. To demonstrate the effects of this relaxation, a very simple yet illustrative example is presented as follows.

\textbf{Example 1.} Consider the problem
\[
\min_z \mathbb{E}_{P}[G(z, \xi)] \\
\text{s.t.} \quad \mathbb{E}_{P}(\eta - G(z, \xi))_+ \leq \mathbb{E}_{P}[(\eta - Y(\xi))_+], \forall \eta \in \mathbb{R}, \\
z \in [1, 2], 
\] (12)
where $\xi \sim \mathcal{U}[0, 1]$, is uniformly distributed on $[0, 1]$, $G(z, \xi) = z\xi$ and
Then for all $\eta < 0.1$ and all $z \in [1, 2]$, the constraint $\mathbb{E}_P[(\eta - G(z, \xi))_+] \leq \mathbb{E}_P[(\eta - Y(\xi))_+]$ cannot be satisfied. Then problem (12) has no feasible solution. However, by construction, if $\eta$ is treated as a random variable $\eta$ with uniform distribution over $[0,1]$, there is only 10% chance that $Y(\xi)$ cannot be “dominated” by $G(z, \xi)$. And if those “10% chance” is omitted, we can get the result at $z = 2$ and the expected loss $-\mathbb{E}_P[G(2, \xi)] = -1$, better than the benchmark loss $-\mathbb{E}_P[Y(\xi)] = -0.5$. Note that at solution $z = 2$, the relaxed CVaR constraint (8) holds.

3. Convergence analysis and the cutting plane method. In this section, the convergence analysis is presented for the optimal values and for the behaviour of optimal solutions of SAA-based CVaR-SSD relaxation problem (11) as CVaR probability level $\beta \to 1$ and sample size $N \to \infty$. Then, a cutting plane algorithm is adapted to solve it.

Starting by reviewing some of the basics in measure theory, let $\mathcal{C}([a, b])$ denote the space of continuous functions defined on $[a, b]$ with maximum norm. By the Riesz’ representation theorem, the space dual to $\mathcal{C}([a, b])$, denoted by $\mathcal{C}^*([a, b])$, is the space of regular countably additive measures on $[a, b]$ having finite variation, see [4, Example 2.63], [7] and the references therein. Let $\mathcal{C}^+_*([a, b])$ denote the subset of $\mathcal{C}^*([a, b])$ of positive measures and $||\mu||$ the induced norm of map $\int_a^b \cdot d\mu$ : $\mathcal{C}([a, b]) \to \mathbb{R}$. Then, for $\mu \in \mathcal{C}^*([a, b])$, $||\mu|| = \int_a^b \mu(d\eta) = \mu([a, b])$, which is the total variation measure of $\mu$ on $[a, b]$, see [14, section 3] and [4, Example 2.63].

3.1. Convergence analysis. The convergence analysis between (5) and its SAA problem (6) is considered first. Note that the convergence analysis of SAA problem of general stochastic optimization problem with SSD has been studied in several papers [16, 22, 32, 34]. Moreover, in [34], the convergence and exponential rate of convergence based on SAA method has been built under mild conditions. Although the results are constructed for stationary points, which may be more interesting in nonconvex case, better results can be achieve with similar conditions in convex case. In this section, the convergence result in [34] is applied to problem (5) with the re-proposed condition in convex cases.

The SAA problem (6) is equivalent to

$$
\min_x \; -\frac{1}{N} \sum_{i=1}^N x^T \xi_i \\
\text{s.t.} \; \frac{1}{N} \sum_{i=1}^N ((\eta - x^T \xi_i)_+ - (\eta - Y(\xi_i))_+) \leq 0, \; \eta \in [a, b],
$$

Based on this expression, the convergence properties of optimal values and optimal solutions of SAA problem (14) as $N \to \infty$ are considered first.

Assumption 1. Slater-type CQ is satisfied in problem (5), i.e., there exists a point $x_0 \in X_0$ such that

$$
\sup_{\eta \in [a, b]} h(x, \eta) < 0.
$$

Let $\mu \in \mathcal{C}^+_*([a, b])$ and define the Lagrange function of problem (5) and (14) as:

$$
\mathcal{L}(x, \mu) := \mathbb{E}_P[-x^T \xi] + \int_a^b h(x, \eta) \mu(d\eta),
$$
and
\[ \mathcal{L}_N(x, \mu) := -\frac{1}{N} \sum_{i=1}^{N} x^T \xi_i + \int_a^b h_N(x, \eta) \mu(d\eta), \]
respectively.

**Lemma 3.1.** Suppose \( \mathbb{E}[|\xi|] < \infty \), then for all \( x \in X_0 \) and \( \eta \in [a, b] \), \( \mathbb{E}[|x^T \xi|] < \infty \), \( h(x, \eta) < \infty \) and \( \mathbb{E}[\|
abla_x (x^T \xi)\|] < \infty \).

The proof of the above lemma is straightforward, yet the result guarantees the properties of the objective and the constraint functions required for the convergence proofs. Also note that the “\( \infty \)” in the above lemma has the corresponding dimensions. Recall that the Bouligand tangent cone to a set \( X \subset \mathbb{R}^n \) at a point \( x \in X \) is defined by,
\[ T_X(x) := \{ u \in \mathbb{R}^n : d(x + tu, X) = o(t), t \geq 0 \}, \]
and the normal cone \( \mathcal{N}_X(x) \) is defined as the polar of the tangent cone,
\[ \mathcal{N}_X(x) := \{ \zeta \in \mathbb{R}^n : \zeta^T u \leq 0, \forall u \in T_X(x) \}, \]
and \( \mathcal{N}_X(x) = \emptyset \) if \( x \notin X \). Then, the following first order necessary conditions of (5) and (14) follow from [34, Section 2], [4, Theorem 5.107] and [7, Theorem 4.2].

**Theorem 3.2** (First order necessary conditions). Suppose \( \mathbb{E}[|\xi|] < \infty \) and Assumptions 1 holds. Let \( x^* \in X_0 \) be an optimal solution to the problem (5). Then there exists \( \mu^* \in \mathcal{C}^*_+([a, b]) \) such that
\[ \begin{cases} 0 \in -\mathbb{E}[\xi] + \int_a^b \partial_x h(x^*, \eta) \mu^*(d\eta) + \mathcal{N}_{X_0}(x^*), \\ h(x^*, \eta) \leq 0, \forall \eta \in [a, b], \\ \int_a^b h(x^*, \eta) \mu^*(d\eta) = 0, \end{cases} \tag{15} \]
where \( \partial_x h(\cdot, \eta) \) denotes the Clarke subdifferential [5] of \( h \), and
\[ \int_a^b \partial_x h(x, \eta) \mu(d\eta) = \left\{ \int_a^b \phi(\eta) \mu(d\eta) : \phi(\eta) \in \partial_x h(x, \eta) \text{ and } \phi(\eta) \text{ is integrable} \right\}. \]

Similarly, the optimality conditions and first order necessary conditions of SAA problems can be shown as follows.

**Theorem 3.3** (First order necessary conditions). Let \( x_N \in X_0 \) be any optimal solution to the problem (14). Suppose \( \mathbb{E}[|\xi|] < \infty \) and Assumptions 1 holds. Then, w.p.1 problem (14) satisfies the Slater-type CQ and there exists \( \mu_N \in \mathcal{C}^*_+([a, b]) \) such that
\[ \begin{cases} 0 \in -\frac{1}{N} \sum_{i=1}^{N} \xi_i + \int_a^b \partial_x h_N(x_N, \eta) \mu_N(d\eta) + \mathcal{N}_{X_0}(x_N), \\ h_N(x_N, \eta) \leq 0, \forall \eta \in [a, b], \\ \int_a^b h_N(x_N, \eta) \mu_N(d\eta) = 0, \end{cases} \tag{16} \]
A tuple \( (x_N, \mu_N(\cdot)) \) satisfying (16) is called a Karush-Kuhn-Tucker (KKT) pair of problem (14), where \( x_N \) is a Clarke stationary point and \( \mu_N(\cdot) \) the corresponding Lagrange multiplier.

**Proposition 3.1.** Consider the SAA problem (14). Suppose \( \mathbb{E}[|\xi|] < \infty \) and Assumption 1 holds. Then, the sequence of the Lagrange multipliers \( \{\mu_N\} \) is bounded w.p.1.
Proof. Note that the dual problem of (14) is
\[
\max_{\mu \in \mathcal{C}_+(\mathbb{R})} \mathcal{J}_N(\mu) := \min_{x \in X_0} \mathcal{L}_N(x, \mu).
\] (17)

It is sufficient to show (17) satisfies the boundedness condition of upper level set in the weak topology w.p.1 when sample size \( N \) sufficiently large, i.e., there exists \( m \) such that \( \mathcal{W}(m, N) := \{ \mu \in \mathcal{C}_+(\mathbb{R}) : \mathcal{J}_N \geq m \} \) is a bounded and nonempty set for \( N \) sufficiently large w.p.1. Let
\[
S_0 := \{ x \in X_0, h(x, \eta) \leq 0, \forall \eta \in [a, b] \}
\]
be the feasible set of problem (5). Assume for the sake of a contradiction that the boundedness result is not true. Then, for some \( m_0 \in \mathbb{R} \) such that \( \mathcal{W}(m_0, N_0) \) is nonempty for \( N_0 \) sufficiently large w.p.1, there exists a sequence \( \{m_k\} \) satisfies \( m_k \geq m_0 \), a sequence \( (N_k, \mu_k) \) such that \( N_k \to \infty \), \( \mu_k \in \mathcal{W}(m_k, N_k) \) for each \( k \) and \( \|\mu_k\| \to \infty \). Since \( \mu_k \in \mathcal{W}(m_k, N_k) \), then by (17), w.p.1
\[
\mathcal{J}_{N_k}(x, \mu_k) \geq \mathcal{J}_{N_k}(\mu_k) \geq m_k \geq m_0, \forall x \in S_0.
\]

Let \( \hat{\mu}_k = \mu_k / \|\mu_k\| \) and since \( \hat{\mu}_k \) is a Borel measure defined on compact set \([a, b]\), by Helly-Bray’s theorem (see Theorems 9.2.1-9.2.3 and Remark 9.2.1 in [3]), it has a weakly convergent subsequence. Take a subsequence if necessary, assuming without loss of generality that \( \hat{\mu}_k \to \hat{\mu} \), which yields \( \|\hat{\mu}\| = 1 \). Dividing both sides of the above inequality by \( \|\hat{\mu}_k\| \) and driving \( k \to \infty \), and it is easy to observe that w.p.1
\[
\frac{1}{N_k} \sum_{i=1}^{N_k} x^T \xi^i \to \mathbb{E}_P[x^T \xi], \quad h_N(x, \eta) \to h(x, \eta) \quad \text{uniformly w.r.t.} \quad \eta \in [a, b]
\]
and then
\[
\int_a^b h(x, \eta) \hat{\mu}(d\eta) \geq 0, \forall x \in S_0.
\]

Note that since Slater-type CQ holds, there exists \( x_0 \in S_0 \) such that \( h(x_0, \eta) < 0 \) for all \( \eta \in [a, b] \). Then we have
\[
\int_a^b h(x_0, \eta) \hat{\mu}(d\eta) \geq 0,
\]
which implies \( \|\hat{\mu}\| = 0 \). This contradicts the fact that \( \|\hat{\mu}\| = 1 \). \( \Box \)

Sun and Xu [34] investigated the boundedness of Lagrange multipliers \( \{\mu_N\} \) under nonzero abnormal multipliers constraint qualification (NNAMCQ) for the general case. Because of the convexity of problem (5), by Proposition 3.1, the easily checkable Slater-type CQ can be used to prove the following.

**Theorem 3.4.** Assume \( \xi \) is a continuous random variable, \( \mathbb{E}[\|\xi\|] < \infty \), and Assumption 1 holds. Furthermore, assume that the support set \( \Xi \) is bounded. Then, sequence \( \{x_N, \mu_N\} \) has cluster points and any cluster point of the sequence, \((x^*, \mu^*)\) is a KKT pair of the problem (5) defined by (15) w.p.1.

Moreover, for any \( \alpha > 0 \), there exist positive constants \( C(\alpha) \) and \( \beta(\alpha) \) independent of \( N \) such that
\[
\text{Prob}\{d(x_N, X^*) \geq \alpha\} \leq C(\alpha) e^{-N\beta(\alpha)},
\]
where \( X^* \) denotes the set of Clarke stationary points characterized by (15).

The results follow from [34, Theorem 5 and Theorem 7]. With the benefit of convexity of problem (5) and linearity of \( x^T \xi \), by Proposition 3.1, we can simplify the conditions of [34, Theorem 7]. Note also that the condition of the continuously differentiability of \( h(x, \eta) \) w.r.t. \( x \) for every \( \eta \in [a, b] \) in [34, Theorem 5] can
be weaken to the continuity of $h^o(x, \eta, u)$ w.r.t. $(x, u)$, which is guaranteed by [34, Proposition 4], where $h^a(x, \eta, u)$ denotes the Clarke directional derivative with direction $u \in \mathbb{R}^n$, see [5].

Let $\beta$ in problem (11) tend to 1, and let $v_N$ and $X_N$ denote the optimal value and optimal solution sets of (6) respectively. Similarly, let $v_N(\beta)$ and $X_N(\beta)$ be the optimal value and optimal solution sets of (11) and $v^*$ denote the optimal value of problem (5).

**Theorem 3.5.** Assume $\xi$ is a continuous random variable, $\mathbb{E}[|\xi|] < \infty$, and Assumption 1 holds. Furthermore, assume that the support set $\Xi$ is bounded. Then, w.p.1

$$\lim_{N \to \infty} \lim_{\beta \to 1} v_N(\beta) \to v^*,$$

(18) and

$$\limsup_{N \to \infty} \limsup_{\beta \to 1} X_N(\beta) \subset X^*.$$  

(19)

**Proof.** By Assumption 1, for $N$ sufficiently large, problem (6) satisfies Slater-type CQ. Then by [2, Theorem 4], w.p.1, $v_N(\beta) \to v_N$ as $\beta \to 1$ and $\limsup_{\beta \to 1} X_N(\beta) \subset X_N$.

The statement (18) and (19) can be achieved by combining the above result with Theorem 3.4. □

3.2. **Cutting plane algorithm.** The cutting plane method, proposed by Kelley [18], is used to solve SAA-based CVaR-SSD relaxation problem (11). By (8) and (11), let the problem be re-written as the following,

$$\min_{x, \zeta} - \frac{1}{N} \sum_{i=1}^{N} x^T \xi_i$$

s.t.  

$$\zeta + \frac{1}{(1-\beta)N} \sum_{j=1}^{N} \left( \frac{1}{N} \sum_{i=1}^{N} ((\eta_j - x^T \xi_i)_+) - (\eta_j - Y(\xi_i))_+ \right)_+ \leq 0,$$

$$x \in X_0, \zeta \in \mathbb{R}.$$ 

For simpler notation, let $\Gamma_i(x, \zeta) := \left( \frac{1}{N} \sum_{j=1}^{N} ((\eta_j - x^T \xi_i)_+) - (\eta_j - Y(\xi_i))_+ \right)_+$, and define

$$\Psi(x, \zeta) := \zeta + \frac{1}{(1-\beta)N} \sum_{j=1}^{N} \Gamma_j(x, \zeta).$$

Functions required for the usage in the cutting plane algorithm, $\rho(x, \zeta) \in \partial_x \Psi(x, \zeta)$, $\varrho(x, \zeta) \in \partial_\zeta \Psi(x, \zeta)$, $b_j(x, \zeta) \in \partial_x \Gamma_j(x, \zeta)$ and $d_j(x, \zeta) \in \partial_\zeta \Gamma_j(x, \zeta)$, can be computed as follows,

$$\rho(x, \zeta) = \frac{1}{(1-\beta)N} \sum_{j=1}^{N} b_j(x, \zeta), \quad \varrho(x, \zeta) = 1 + \frac{1}{(1-\beta)N} \sum_{j=1}^{N} d_j(x, \zeta),$$

$$b_j(x, \zeta) = \begin{cases} 0, & \Gamma_j(x, \zeta) = 0, \\ \sum_{i=1}^{N} \gamma_{ij}, & \text{others} \end{cases}, \quad d_j(x, \zeta) = \begin{cases} 0, & \Gamma_j(x, \zeta) = 0, \\ -1, & \text{others}. \end{cases}$$

$$\gamma_{ij}(x, \zeta) = \begin{cases} 0, & \eta_j - x^T \xi_i \leq 0, \\ -\xi_j, & \eta_j - x^T \xi_i > 0, \end{cases}$$
Similar to our previous notation, “0” has the corresponding dimension in the above expressions.

**Algorithm 1** Cutting Plane Method

**Initialization:** Set iteration $t = 0$, tolerance $\epsilon \geq 0$, $S_0 = \{(x, \zeta) | x \in X_0, \zeta \in \mathbb{R}\}$.

**Step 1:** Solve the linear programming problem

$$
\min_{(x, \zeta)} -\frac{1}{N} \sum_{i=1}^{N} x^T \xi_i
$$

s.t. $$(x, \zeta) \in S_t.$$

Let $(x_t, \zeta_t)$ denote the optimal solution of problem (20).

**Step 2:** If $\Psi(x_t, \zeta_t) \leq \epsilon$, $(x_t, \zeta_t)$ is the optimal solution, stop; otherwise, go to Step 3.

**Step 3:** Set

$$S_{t+1} = S_t \cap \{(x, \zeta) | \Psi(x_t, \zeta_t) + (\rho(x_t, \zeta_t)^T \cdot g(x_t, \zeta_t))^T (x - x_t)^T, \zeta - \zeta_t \leq 0\}.$$

**Step 4:** $t = t + 1$, go to Step 1.

Note that Algorithm 1 is the standard cutting plane algorithm proposed in [18], see reference [18] for the details and convergence analysis of the algorithm. Note also that, when $\epsilon \downarrow 0$, the solution calculated from Algorithm 1 converges to the real solution of problem (11).

4. **Empirical studies.** Both the motivation and the aim of this project are to construct better portfolios for risk-averse investors. In this section, the results of the empirical studies on SAA-based CVaR-SSD relaxation model in applications of solving portfolio selection problems are presented. The methodology of identifying advantageous strategy is to compare the portfolios in terms of their performance, and in turn, to analyze the strategies of constructing them. All the portfolios under comparison are obtained by solving the corresponding portfolio optimization problems, e.g., index tracking, MV model, SSD model and SAA-based CVaR-SSD relaxation model. More specifically, index portfolios came straight from data sets as the perfect trackings of closing value of indices, while MV portfolios are obtained by solving the standard non-shorting MV model (See problem (4.1)). For each data set, both index portfolio and MV portfolio are used as benchmarks for solving SSD problem and SAA-based CVaR-SSD relaxation problem. SSD portfolios are the solutions of SSD problem (5) with corresponding benchmarks, and CVaR-SSD portfolios are constructed by solving SAA-based CVaR-SSD relaxation problem (11) with both benchmarks and different probability level $\beta$. The performance of portfolio is quantified in terms of average daily returns, cumulated returns and risk control abilities, measured by standard deviation, Sharpe Ratio and Sortino Ratio$^1$. The main observation from the empirical study is that the CVaR-SSD portfolio obtained by solving the proposed SAA-based CVaR-SSD relaxation problem with probability level $\beta < 1$ demonstrates comparable performance with respect to SSD portfolio using the same benchmark. Moreover, with appropriate choice of probability level $\beta < 1$ demonstrates comparable performance with respect to SSD portfolio using the same benchmark.

---

$^1$Sortino Ratio [36] is a modification of the Sharpe Ratio but penalizes only those returns falling below a required rate of return, while the Sharpe Ratio penalizes both upside and downside volatility equally. The ratio $S$ is calculated as $S = \frac{R - T}{DR}$, where $R$ is portfolio expected return, $T$ is target rate of return, and $DR$ is target semi-deviation.
it outperforms SSD portfolio in most cases. A rule of thumb is to choose the value of $\beta$ be less but close to 1, while the “best fit” value depends on the data set as well as the choice of benchmarks.

4.1. **Scope of the study and data.** In this subsection, the specifications of the numerical experiments are explained and the description of data sets is presented. The aim is to construct portfolios by solving various optimization problems subject to constraints, which reflect the underlining strategies of portfolio construction models. The analysis mainly investigates the following aspects of constructed portfolios:

- validity of approximation results is tested empirically by comparing CVaR-SSD portfolios and SSD portfolios in-sample;
- comparisons of the mean, standard deviation, Sharpe Ratio, Sortino Ratio of daily ex-post returns, and cumulated returns among SSD portfolios, CVaR-SSD portfolios and their corresponding benchmark portfolios;
- sparsity results of CVaR-SSD portfolios, SSD portfolios and MV portfolios.

The data sets used are the historical data of three financial indices:

- National Association of Securities Dealers Automated Quotations 100 Index (NDX),
- Standard and Poor’s (S&P 500),
- Financial Times Stock Exchange 100 Index (FTSE 100),

within the period 01/03/2016 – 30/09/2016.

These indices are treated as (benchmark) portfolios while all other portfolios are constructed based on real data of daily closing prices of the component stocks of the corresponding index. More specifically, as on 30/09/2016, NDX is calculated from the weighted sum of 100 largest non-financial companies based on market capitalization listed on NASDAQ. The daily closing prices of these stocks are publicly available from 01/03/2016 to 30/09/2016, and they form the pool of assets from which MV portfolios, SSD portfolios and CVaR-SSD portfolios are constructed from. Similarly, data sets of daily closing prices are obtained for all the 500 component stocks contributing to S&P 500 index, excepts for Fortive, whose was listed within the investigation period after spinning off from Danaher in July 2016, and Westrock, who was merged with MeadWestvaco before listing in The New York Stock Exchange. 100 assets listed in the London Stock Exchange contributing to FTSE 100 are available for portfolio constructions, except B share of Royal Dutch Shell. It is not included because of the lack of full pricing data due to different taxation issues than the A share of the company, which is also one of the 101 contributing stocks to the index.

Besides the usage of indices as the choice of benchmark, constructions of the classical MV portfolios served as alternative choices. For $n$ risky securities, the classical non-shorting MV portfolio has minimal variance for a given expected return $\rho$ and can be obtained by solving the following problem:

\[
\begin{align*}
\min & \quad x^T C x \\
\text{s.t.} & \quad x^T \mu = \rho, \\
& \quad x^T 1_n = 1, \\
& \quad x \geq 0,
\end{align*}
\]
where $1_n$ is an $n$-dimensional vector with all entries being one, random variable $r_i : \Omega \rightarrow \mathbb{R}$ is the return of the $i$th security, $\mu_i = \mathbb{E}[r_i]$ is its expected return and $r(\xi) = (r_1, r_2, \cdots, r_n)^T$, $\mu = (\mu_1, \mu_2, \cdots, \mu_n)^T$, the covariance matrix of the returns $C = \mathbb{E}[(r - \mu)(r - \mu)^T]$. Note that $C$ is an $n \times n$ positive semi-definite matrix.

The MV portfolios are constructed and revised at daily basis with rolling window sampling strategy.

All the numerical tests are carried out in MATLAB R2014a installed on an Lenovo PC with Windows 7 operating system and Intel Core i7 processor. IBM ILOG CPLEX Studio 12.4 solver is used for solving the subproblems (20) with cutting plane algorithm. The stopping criteria is set to be $\epsilon = 10^{-5}$ for all cases in our empirical studies.

4.2. In-sample verification of convergence. We devote the first part of our empirical studies to verify our theoretical analysis on the convergence of SAA-based CVaR-SSD relaxation problem to SSD problem as CVaR probability level $\beta \rightarrow 1$. Historical data within the period 01/03/2016 - 31/05/2016 is used as the in-sample data set and we back test the returns of the portfolios obtained by solving SSD problem (5) and our proposed SAA-based CVaR-SSD relaxation problem (11) within this period. With in the period, we took 64 observations for the in-sample tests of NDX and S&P 500, while there are 62 scenarios for the data set of FTSE 100. The daily returns, calculated from the observation on the difference between the closing prices of adjacent trading days, are used to form the observation of random vector of return rates in-sample. Figure 1 shows the results in terms of returns in-sample with NDX index as benchmark. CVaR-SSD portfolios obtained by solving SAA-based CVaR-SSD problems with different probability levels $\beta = 0.9, 0.8, 0.7$ are shown, and the SSD portfolio in-sample returns are also plotted with the same choice of benchmark for comparison. It can be observed that as the probability level $\beta$ tends to 1, the return plots of CVaR-SSD portfolios converge to that of the SSD portfolio, as demonstrated theoretically in Section 3. In-sample tests are also carried out with both S&P 500 and FTSE 100 data sets and we arrive at the same conclusion regarding the in-sample convergence in term of the tendency of return plots. Due to the similarity of the plots, only the in-sample test for the NDX data set is shown in Figure 1.

4.3. Out-of-sample test. In this subsection, demonstrations of the results of ex-post testing for three sets of market data and with different choices of probability levels $\beta$ for the SAA-based CVaR-SSD relaxation problem (11) are presented. At our disposal, the historical data covers the period 01/03/2016 - 30/09/2016.

The comparisons of ex-post performance among portfolios are presented in daily basis. For all the portfolio constructions in this subsection, a “rolling one day” investment period is adapted. Portfolio optimization problems are consistent in the way that on any given date, a fixed size sampling window prior to the decision date is used as training data. More specifically, the daily closing prices of each of the component stocks and the corresponding benchmark portfolio are used within the training data as scenarios to construct current day portfolio. Based on these, we solve different portfolio optimization problems and their corresponding solutions are the decided portfolios for the trading day following the last day in the training sampling window. The indices are used as benchmarks and for the first day of ex-post portfolio decisions on 01/06/2016, the training data is within the period
01/03/2016 - 31/05/2016. Within the first period, there are 64 trading days for NDX and S&P 500, while there are 62 trading days for FTSE 100. The portfolio obtained with data in this sampling window is evaluated using the closing prices of the component stocks on that day. Thus, the daily return of the first out-of-sample portfolio can be obtained. For the next trading day, the real data on 01/06/2016 becomes known, and we use the real data within period 02/03/2016 - 01/06/2016 as the training data set and the portfolio obtained with this data set is evaluated as the second out-of-sample portfolio. By analogy, with a daily rolling window sampling method, training data set will always have 64 or 62 scenarios, based on which the next-day portfolio is constructed and evaluated. Thus, for each portfolio optimization model, either 86 portfolio optimization problems or 87 portfolio optimization problems are solved and the performance of portfolios obtained are evaluated on the next working day’s historical closing price. Correspondingly, each data set results in 86 or 87 out-of-sample daily portfolio decisions, covering the period 01/06/2016 - 30/09/2016. Concerning the benchmarks, within the training data set, the corresponding indices are available and the MV portfolios are standard to construct. For both benchmarks, the results are presented.

Table 1 shows the average daily return rate, standard deviation, Sharpe Ratio and Sortino Ratio of the ex-post returns of 8 constructed portfolios as well as two benchmark portfolios using NDX data. With both benchmarks, SSD portfolios and CVaR-SSD portfolios perform better than the corresponding benchmarks with higher mean, Sharpe Ratio and Sortino Ratio of returns, although with slightly less attractive standard deviation. Note that with NDX index as benchmark, solving SAA-based CVaR-SSD relaxation problem (11) with $\beta = 0.7, 0.9$ result in portfolios with slightly larger mean, Sharpe Ratio and Sortino Ratio than SSD portfolio. When $\beta = 0.8$, the performance of CVaR-SSD portfolio is comparable to that of SSD portfolios using our comparison criteria. In the case of MV benchmark,
the CVaR-SSD portfolios behave better than the SSD portfolio for all choices of probability levels for NDX data.

The ex-post compounded returns are computed over the period 01/06/2016 - 30/09/2016 and both SSD portfolios and CVaR-SSD portfolios perform better than the corresponding benchmark portfolios as shown in Figure 2 and Figure 3, respectively. In particular, both the index portfolios and MV portfolios have very small accumulative return, less than 8% for index and less than 5% for MV portfolio, while the CVaR-SSD portfolios perform consistently better throughout the whole out-of-sample period, thus having a cumulated gain of over 30% for our choices of $\beta$ with index benchmark. In addition, throughout almost all the period 01/06/2016 - 30/09/2016, CVaR-SSD portfolios with $\beta = 0.9$ have the best compounded returns in the index benchmark case. For the benchmark being the MV portfolio returns,

| Benchmark       | mean  | std  | Sharpe Ratio | Sortino Ratio |
|-----------------|-------|------|--------------|---------------|
| Benchmark: index| 0.0009| 0.0088| 0.1030       | 0.1389        |
| SSD             | 0.0032| 0.0118| 0.2717       | 0.4501        |
| $CVaR_{\beta=0.9}$ | 0.0033| 0.0118| 0.2784       | 0.4670        |
| $CVaR_{\beta=0.8}$ | 0.0032| 0.0117| 0.2724       | 0.4486        |
| $CVaR_{\beta=0.7}$ | 0.0033| 0.0119| 0.2810       | 0.4652        |

| Benchmark       | mean  | std  | Sharpe Ratio | Sortino Ratio |
|-----------------|-------|------|--------------|---------------|
| Benchmark: MV   | 0.0005| 0.0078| 0.0658       | 0.0841        |
| SSD             | 0.0026| 0.0102| 0.2545       | 0.4129        |
| $CVaR_{\beta=0.9}$ | 0.0027| 0.0118| 0.2696       | 0.4442        |
| $CVaR_{\beta=0.8}$ | 0.0030| 0.0117| 0.2931       | 0.4916        |
| $CVaR_{\beta=0.7}$ | 0.0030| 0.0102| 0.2943       | 0.5004        |

| Benchmark       | mean  | std  | Sharpe Ratio | Sortino Ratio |
|-----------------|-------|------|--------------|---------------|
| Benchmark: index| 0.0009| 0.0088| 0.1030       | 0.1389        |
| SSD             | 0.0032| 0.0118| 0.2717       | 0.4501        |
| $CVaR_{\beta=0.9}$ | 0.0033| 0.0118| 0.2784       | 0.4670        |
| $CVaR_{\beta=0.8}$ | 0.0032| 0.0117| 0.2724       | 0.4486        |
| $CVaR_{\beta=0.7}$ | 0.0033| 0.0119| 0.2810       | 0.4652        |

Table 1. NDX: average daily return, standard deviation, Sharpe Ratio, Sortino Ratio

Figure 2. NDX: ex-post compounded daily returns (01/06/2016 - 30/09/2016), index returns as benchmark
the portfolios with $\beta = 0.8, 0.7$ behave better than the SSD portfolio and the CVaR-SSD portfolio with $\beta = 0.9$ behaves similar to SSD portfolio in the cumulated return plots.

Figure 3. NDX: ex-post compounded daily returns (01/06/2016 - 30/9/2016), MV returns as benchmark

Table 2 shows the ex-post performance of comparison portfolios with S&P 500 data. Similar to the results of the NDX data, both SSD portfolios and CVaR-SSD portfolios have better performance than the corresponding benchmarks. In the index benchmark case, CVaR-SSD portfolios with probability level $\beta = 0.9$ produce better results than that of other CVaR-SSD portfolios. Indeed, Figure 4 shows that CVaR-SSD portfolios with probability level $\beta = 0.9$ also have the best compounded return of 17.1% on the final day of test period, which is much better than the index return and maintaining advantageous over the other portfolios including the SSD portfolio. The plots for MV benchmark portfolios shown in Figure 5 are rather different in the sense that the MV portfolio, despite being the benchmark, behaves better than other portfolios in earlier stage of ex-post testing period. Both SSD portfolio and CVaR-SSD portfolios behave better than MV portfolio when long term observation is of consideration. Moreover, all CVaR-SSD portfolios with different values of $\beta$ taken behave better than the SSD portfolios by the end of the ex-post testing period. It is observed that even the SSD portfolios seems to “fail” for relatively short investment horizon and only gain in a longer investment window. This is due to the structure of SSD constraints (respectively, the SAA-based CVaR-SSD constraints) takes account over a period of volatility of stock prices while remains not very sensitive to sudden and rapid changes in price fluctuation. In this particular example, on 23/06/2016, the sudden event occurred when people of the United Kingdom voted to leave EU, catching the financial markets by surprise, resulted in days of stock prices’ free-fall.

Similar behaviours can also be observed in the case of the FTSE 100 market data, shown in Table 3, Figure 6 and Figure 7. The SSD portfolio and CVaR-SSD portfolios with 3 different $\beta$ have better ex-post performance than the benchmark portfolios. The best of such, the CVaR-SSD portfolio with confidence $\beta = 0.7$ has the best performance with 18.5% cumulated gain in the end of the index benchmark
case. As for the case of MV portfolio returns as benchmark, the CVaR-SSD portfolios with $\beta = 0.8, 0.9$ behave comparable to the SSD portfolio and the relaxation leads to better portfolio when $\beta = 0.7$. This observation suggests that although with $\beta < 1$ a relaxation model of portfolio optimization problem can produce better portfolios than the conservative SSD portfolio. The “best” choice of $\beta$, or “level” of relaxation, is not universal across different data sets and benchmarks. From the results of empirical tests, it seems that the choice of $\beta$ is subject to different data structure and benchmark choices.

The SSD portfolios and CVaR-SSD portfolios have been demonstrated to have an overall better performance than the corresponding benchmarks, especially the CVaR-SSD portfolios with probability levels $\beta$ less but close to 1. CVaR-SSD portfolios with $\beta = 0.9$ is the best for NDX and S&P 500 data sets, and with $\beta = 0.7$, the portfolio outperforms the rest in the case of FTSE 100 data. With MV benchmarks, the portfolios corresponding to $\beta = 0.7$ behave the best over all three

| Benchmark: index | mean   | std    | Sharpe Ratio | Sortino Ratio |
|------------------|--------|--------|--------------|---------------|
| SSD              | 0.0017 | 0.0112 | 0.1490       | 0.2264        |
| $CVaR_{\beta=0.9}$ | 0.0018 | 0.0111 | 0.1606       | 0.2444        |
| $CVaR_{\beta=0.8}$ | 0.0016 | 0.0114 | 0.1442       | 0.2171        |
| $CVaR_{\beta=0.7}$ | 0.0017 | 0.0115 | 0.1477       | 0.2241        |

Table 2. S&P 500: average daily return, standard deviation, Sharpe Ratio, Sortino Ratio
sets of data. To summarize, the CVaR-SSD portfolios always behave comparable to that of the SSD portfolios and, in most cases, much better with the same training sets.

*Sparse portfolios.* Another observation of the constructed portfolios is the number of stocks in the composition of the efficient portfolios (Table 4). It is observed that the cardinalities of the portfolios have slight variations among different models within the same market. For example, considering the portfolios constructed by solving the SAA-based CVaR-SSD relaxation model with probability level $\beta = 0.9$ in the case of NDX data set with index as benchmark, 19 stocks are “selected” to construct the portfolios over the out-of-sample investment horizon of 86 trading days. On average, daily optimal portfolio has 4.6511 component stocks with the largest basket has 10 stocks. Baskets of similar size are traded for all SSD portfolios and CVaR-SSD portfolios with different probability levels $\beta$ with both benchmarks in all three data sets, see Table 4.
Although MV portfolios can also achieve index-like performance with less components, desired under e.g., enhanced indexation setting [38]. Both SSD portfolios and CVaR-SSD portfolios contain much less number of stocks in their compositions, typically less than one tenth of the available stocks, while dominate the corresponding benchmarks in terms of performance.

We finish this section by commenting on the choice of benchmarks, which is of essential importance for portfolio selection via our proposed model. By definition of SSD, solving the SSD constrained portfolio selection problem with a proper benchmark is to choose the “optimal” out of the portfolios already “better” than benchmark in the sense of risk-aversing described by SSD. This can be observed.
from our empirical experiment that the solution of SAA-based CVaR-SSD relaxation problem with higher accumulated return benchmark has greater returns at the end of investment horizons. Similar results can be seen from our empirical experiment results for all three data sets. The feature of the proposed SSA-based CVaR-SSD relaxation model is that for a given benchmark portfolio, the portfolio obtained as the solution of the problem may provide better than benchmark asset allocation choice. The performance of our approach is by default influenced by the specific choice of benchmark, which is either user specified or pre-determined at the modelling stage of the problem. The choice of benchmark is of great interest on its own along with the proper decision criteria for identifying good portfolios. In practice, such choices are usually based on the investor pre-acquired knowledge on the market as well as investment preferences.

4.4. **Empirical study results.** To conclude this section, it is found empirically that:

1. the portfolios obtained by solving SAA-based CVaR-SSD relaxation problem (11) converge to that obtained by solving problem (5) as $\beta \rightarrow 1$;
2. both SSD portfolios and CVaR-SSD portfolios have overall better performance than the corresponding benchmarks;
3. the performance of CVaR-SSD portfolios are always comparable to that of SSD portfolios, and when $\beta$ is less but close to 1, the performance of CVaR-SSD portfolios is better than the more conservative SSD portfolios in most cases;
4. both SSD portfolios and CVaR-SSD portfolios are sparse portfolios.

| Index       | MV       |
|-------------|----------|
|             | avg.     | min.     | max.     | avg.     | min.     | max.     |
| NDX (100)   | SSD      | 4.60     | 3        | 9        | 5.55     | 3        | 9        |
|             | CVaR$_{\beta=0.9}$ | 4.65     | 3        | 10       | 5.53     | 2        | 9        |
|             | CVaR$_{\beta=0.8}$ | 4.65     | 3        | 10       | 5.51     | 2        | 8        |
|             | CVaR$_{\beta=0.7}$ | 4.64     | 3        | 9        | 5.50     | 3        | 8        |
| FTSE (100)  | SSD      | 4.05     | 2        | 9        | 5.16     | 2        | 9        |
|             | CVaR$_{\beta=0.9}$ | 3.98     | 2        | 9        | 5.11     | 2        | 9        |
|             | CVaR$_{\beta=0.8}$ | 3.90     | 2        | 8        | 5.01     | 2        | 8        |
|             | CVaR$_{\beta=0.7}$ | 3.91     | 3        | 9        | 5.17     | 2        | 9        |
| S&P (500)   | SSD      | 5.87     | 3        | 10       | 6.62     | 4        | 11       |
|             | CVaR$_{\beta=0.9}$ | 6.07     | 3        | 11       | 6.49     | 3        | 12       |
|             | CVaR$_{\beta=0.8}$ | 6.07     | 3        | 12       | 6.70     | 4        | 12       |
|             | CVaR$_{\beta=0.7}$ | 6.30     | 3        | 11       | 6.74     | 4        | 12       |

**Table 4.** Average, minimum and maximum of daily traded basket sizes of different models with both benchmarks in three data sets.
5. **Conclusions.** In this paper, we propose a SAA-based CVaR-SSD relaxation model for portfolio optimization problem based on an approximation to the portfolio optimization problem with SSD constraints based on SAA method and CVaR approach. SAA method is used to simplify the calculations of expected value and also to handle the infinite number of inequalities in SSD constraints, followed by the SSD constraints been approximated and relaxed by CVaR. The convergence results for SAA-based CVaR-SSD relaxation problem (11) are shown based on recent convergence analysis results of stochastic optimization problem [34, 2]. The effectiveness of the constructed portfolios are tested as portfolio selection strategies, using three data sets: NDX, S&P 500 and FTSE 100, with both indices and MV portfolios as benchmarks. By empirical studies, it is observed that the performance of CVaR-SSD portfolios changes with different choices CVaR probability levels $\beta$. More importantly, when the value of $\beta$ is chosen to be less but close to 1, CVaR-SSD portfolios perform comparably and often outperform the SSD portfolios, while the advantages of the former are demonstrated and analyzed. Furthermore, the portfolios obtained by solving SSD problem (5) and SAA-based CVaR-SSD relaxation problem (11) are sparse, much less than the total number of available stocks. Based on the good performance of CVaR-SSD portfolios in mean value, standard deviation, Sharpe Ratio, Sortino Ratio of ex-post returns, and the ex-post compounded returns, it is concluded that solving SAA-based CVaR-SSD relaxation problem is an efficient method to construct portfolios.

**REFERENCES**

[1] G. J. Alexander and A. M. Baptista, A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model, *Manag. Sci.*, 50 (2004), 1261–1273.

[2] E. Anderson, H. Xu and D. Zhang, Confidence levels for CVaR risk measures and minimax limits, manuscript, The University of Sydney, 2014. Available from: [http://hdl.handle.net/2123/9943](http://hdl.handle.net/2123/9943).

[3] K. B. Athreya and S. N. Lahiri, *Measure Theory and Probability Theory*, Springer Texts in Statistics, Springer, New York, 2006.

[4] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Series in Operations Research, Springer-Verlag, 2000.

[5] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983.

[6] N. A. Canakgoz and J. E. Beasley, Mixed-integer programming approaches for index tracking and enhanced indexing, *European J. Oper. Res.*, 196 (2009), 384–399.

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

[8] D. Dentcheva and A. Ruszczyński, Semi-infinite probabilistic constraints: Optimality and convexification, *Optimization*, 53 (2004), 583–601.

[9] D. Dentcheva and A. Ruszczyński, Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints, *Math. Program.*, 99 (2004), 329–350.

[10] D. Dentcheva and A. Ruszczyński, Portfolio optimization with stochastic dominance constraints, *J. Bank. Financ.*, 30 (2006), 433–451.

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

[12] D. Dentcheva and A. Ruszczyński, Risk-averse portfolio optimization via stochastic dominance constraints, in *Handbook of Quantitative Finance and Risk Management* (eds. C. Lee, A. Lee and J. Lee), Springer, New York, (2015), 2281–2302.

[13] C. Fábián, G. Mitra and D. Roman, Processing second-order stochastic dominance models using cutting-plane representations, *Math. Program.*, 130 (2011), 33–57.

[14] M. Gugat, A parametric review on the Mangasarian-Fromovitz constraint qualification, *Math. Program.*, 85 (1999), 643–653.

[15] T. Homem-de-Mello and S. Mehrota, A cutting surface method for uncertain linear programs with polyhedral stochastic dominance constraints, *SIAM J. Optim.*, 20 (2009), 1250–1273.
[16] J. Hu, T. Homen-De-Mello and S. Mehrotra, Sample average approximation of stochastic dominance constrained programs, *Math. Program.*, **133** (2012), 171–201.

[17] P. Jorion, *Value at Risk: The New Benchmark for Controlling Market Risk*, McGraw-Hill Inc., US, 2006.

[18] J. E. Kelley, The cutting-plane method for solving convex programs, *SIAM J. Appl. Math.*, **8** (1960), 703–712.

[19] P. Kolm, R. Tütüncü and F. Fabozzi, 60 Years of portfolio optimization: Practical challenges and current trends, *European J. Oper. Res.*, **234** (2014), 356–371.

[20] H. Konno and H. Yamazaki, Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market, *Manag. Sci.*, **37** (1991), 519–531.

[21] H. Konno, H. Shirakawa and H. Yamazaki, A mean-absolute deviation-skewness portfolio optimization model, *Ann. Oper. Res.*, **45** (1993), 205–220.

[22] Y. Liu and H. Xu, Stability and sensitivity analysis of stochastic programs with second order dominance constraints, *Math. Program.*, **142** (2013), 435–460.

[23] H. Markowitz, *Portfolio Selection*, *J. Finance*, (1952), 77–91.

[24] H. Markowitz, *Portfolio Selection*, John Wiley & Sons, New York, 1959.

[25] K. Mosler and M. Scarsini, (eds.), *Stochastic Orders and Decision under Risk*, Institute of Mathematical Statistics, Hayward, CA, 1991.

[26] W. Ogryczak and A. Ruszczyński, From stochastic dominance to mean-risk models: Semideviations as risk measures, *Euro. J. Oper. Res.*, **116** (1999), 33–50.

[27] W. Ogryczak and A. Ruszczyński, Dual stochastic dominance and related mean-risk models, *SIAM J. Optim.*, **13** (2002), 60–78.

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

[29] D. Roman, K. Darby-Dowman and G. Mitra, Portfolio construction based on stochastic dominance and target return distributions, *Math. Program.*, **108** (2000), 541–569.

[30] D. Roman, G. Mitra and V. Zverovich, Enhanced indexation based on second-order stochastic dominance, *European J. Oper. Res.*, **228** (2013), 273–281.

[31] G. Rudolf and A. Ruszczyński, Optimization problems with second order stochastic dominance constraints: Duality, compact formulations, and cut generation methods, *SIAM J. Optim.*, **19** (2008), 1326–1343.

[32] H. Sun, H. Xu and Y. Wang, A smoothing penalized sample average approximation method for stochastic programs with second order stochastic dominance constraints, *Asia-Pac. J. Oper. Res.*, **30** (2013), 1340002 (25 pages).

[33] H. Sun, H. Xu, R. Meskarian and Y. Wang, Exact penalization, level function method and modified cutting-plane method for stochastic programs with second order stochastic dominance constraints, *SIAM J. Optim.*, **23** (2013), 602–631.

[34] H. Sun and H. Xu, Convergence analysis of stationary points in-sample average approximation of stochastic programs with second order stochastic dominance constraints, *Math. Program.*, **143** (2014), 31–59.

[35] W. E. Sharpe, *The Sharpe Ratio*, *J. Portfolio Manage.*, **21** (1994), 49–58.

[36] F. A. Sortino and L. N. Price, Performance measurement in a downside risk framework, *J. Invest.*, **3** (2009), 59–64.

[37] G. A. Whitmore and M. C. Findlay, (eds.), *Stochastic Dominance: An Approach to Decision-Making Under Risk*, D.C.Heath, Lexington, MA., 1978.

[38] F. Xu, M. Wang, Y. H. Dai and D. Xu, A sparse enhanced indexation model with chance and cardinality constraints, *J. Glob. Optim.*, **70** (2018), 5–25.

Received December 2017; revised March 2019.

E-mail address: xue.njust@163.com
E-mail address: yun.shi@connect.polyu.hk
E-mail address: hlsun@njust.edu.cn