Research Article

Yijia Zhou and Lijun Xu*

A new distributionally robust reward-risk model for portfolio optimization

https://doi.org/10.1515/math-2024-0010
received June 28, 2023; accepted March 25, 2024

Abstract: A new distributionally robust ratio optimization model is proposed under the known first and second moments of the uncertain distributions. In this article, both standard deviation (SD) and conditional value-at-risk (CVaR) are used to measure the risk, avoiding both fat-tail and volatility. The new model can be reduced to a simple distributionally robust model under assumptions on the measurements of reward, CVaR and SD. Furthermore, it can be rewritten as a tractable semi-definite programming problem by the duality theorem under partially known information of the uncertain parameters. Finally, the model is tested on portfolio problems and verified from numerical results that it can give a reasonable decision under only the first and second moments.

Keywords: distributionally robust, portfolio optimization, CVaR, standard deviation, semidefinite programming

MSC 2020: 90-10, 90B50, 90C17, 65C20, 65K10

1 Introduction

The studies on robust optimization have been explosively increased. It has been a popular methodology to deal with practical problems involving uncertain parameters. Robust optimization immunizes deterministic optimization problems against perturbations of parameters by inducing approaches to tackle uncertainty with probability distributions. From computational viewpoint, robust optimization usually provides a tractable methodology based on fundamental analysis. We refer to [1–3] for comprehensive guides to robust optimization.

Robust optimization problems contain uncertain parameters but do not require the specific distributions in models. It is also regarded as the main difference from stochastic programming for optimization problems with uncertainties [4,5]. For stochastic programming referring to [6], if the specific distribution of uncertain parameters is assumed differently from the actual distribution, the results may perform badly. Conversely, if the actual distribution is precisely known, approaches of robust optimization would be overly conservative and unnecessary. For balancing the specificity of the former and the conservatism of the latter, a dominant approach is the min-max optimization method, which is considered as finding optimal solutions in the worst case scenario under all possible probability distributions. Possible distributions are usually defined by certain properties for example the moment information.

Distributionally robust optimization problems have been developed for about 70 years and the earliest work can date back to Scarf [7] in 1950s. Most methods about distributionally robust optimization utilize the duality theory for some problems with moment information (e.g. [8–10]). Many approaches have been extended to distributionally robust problems with chance constraints which are linear and conic. In this

*** Corresponding author: Lijun Xu**, School of Science, Dalian Maritime University, Dalian, Liaoning, 116026, Dalian, China, e-mail: lijun_xu@dlmu.edu.cn
Yijia Zhou: School of Computer & Software, Dalian Neusoft University of Information, Dalian 116023, Dalian, China, e-mail: zhouyijia@neusoft.edu.cn

case, some distribution properties like average, covariance, and the other supports are known (e.g. [11–14]). In [15], they transform the linear programming with uncertainty into a deterministic convex programming by inducing the expectation bounds of distributionally robust. A unified framework for remodelling and computing distributionally robust problems is proposed in [16]. El Ghaoui et al. [17] used a conic programming approach to deal with portfolio problems with distributionally robust optimization. For robust portfolio problems, they develop the Value-at-Risk bounds in the worst case which rely on the known bounds on the means and covariance. Delage and Ye [18] studied distributionally robust problems with uncertain moments using techniques of stochastic programming and applied the approach to data-driven problems. Then, some research arise in distributionally robust ratio optimization. For example, Zhao et al. [19] studied the problem where the ambiguity on the distributions is defined through Wassertein metric and Ji et al. [20] focused on the problem with linearized stable tail adjusted return ratio (STARR) performance measure. In [21], the distributionally robust optimization approach is developed to set up a new risk-limiting dispatch model. Liu et al. proposed a distributionally robust reward-risk ratio model varied from the Sharpe ratio where the ambiguity set is constructed through prior moment information and the return function is not necessarily linear in [22]. In [23], robust portfolio optimization models for reward–risk ratios are proposed by utilizing Omega, semi-mean absolute deviation ratio, and weighted STARR. Moreover, Guo and Xu considered distributionally robust version of the shortfall risk measure in [24]. Recently, Gotoh et al. studied the out-of-sample properties of robust empirical optimization problems with smooth phi-divergence penalties and smooth concave objective functions in [25]. In [26], a class of fractional distributionally robust optimization problems are investigated and studied computationally. Xu et al. [27] studied the distributionally robust bi-level programming based on worst conditional value at risk (WCVaR) in 2022.

The aim of this article is seeking a robust solution under uncertain distributions for reward-risk ratio problems. The solution is optimal and feasible for the worst case scenario under all possible probability distributions. The outline of this article is organized as follows. We first devote to present a distributionally robust reward-risk model involving both conditional value-at-risk (CVaR) and standard deviation (SD) with partial information of uncertainty in the next section. By introducing the properties of the new ratio model, the equivalent model without ratio is proposed in Section 3. In Section 4, the new model with certain moments can be rewritten as a semidefinite programming which is computationally tractable. In Section 5, numerical experiments are conducted to validate the new model is robust and comparable for portfolio problems.

2 Distributionally robust reward-risk optimization model

In this work, we focus on the reward-risk model with both CVaR and SD. Recently, a new model considering both CVaR and SD is presented for robust reward-risk problem in [28]. The model is used to formulate a large class of financial problems with uncertain parameters. The general model is

$$\begin{aligned} \max_{x} \quad & x^T\xi \\ \text{s.t.} \quad & \mathrm{CVaR}_F(X) + \mathrm{std}_F(X) \leq R^*, \\ & x \in \mathscr{X}, \end{aligned}$$

where $\xi \in R^n$ is a random return variable with respect to decision variable $x \in \mathscr{X} \subseteq R^n$, F stands for the true distribution of ξ, and $\mathscr{X}$ denotes as a decision set with many kinds of expressions. In this work, we dedicate to the nonempty box decision set with the following version:

$$\mathscr{X} = \{x : e^T x = 1, L \leq Bx \leq U\}, \tag{1}$$

where L, U, B, and e are known lower bound, upper bound, associated matrix, and all one vector, respectively. It is obvious that $\mathscr{X}$ is convex. In this constraint set, the distribution F is fully and accurately required. However, the distribution F can only be obtained approximately from historical data and then is uncertain in many practical situations. Typically, partial information of F can be estimated from data sets, e.g. the first and second moments etc. Since the uncertain distribution F is replaced by an estimate $\hat{F}$ with some measurement errors, the obtained solutions with $\hat{F}$ often fail to satisfy the constraint.

In the following of this article, denote the vector $\mu \in R^n$ as the mean return and let the matrix $\Sigma \in S^n$ be the covariance of ξ under the true probability distribution F. Therefore, the second-order moments of F are supposed to be finite. Without loss of generation, Σ is also assumed to be positive definite. Furthermore, P is defined as a set of probability distributions, of which the first and second moments are equal to that of F. That is

$$P = \left\{ p \in M_+ : \int_{R^n} p(\mathrm{d}\xi) = 1, \int_{R^n} \xi p(\mathrm{d}\xi) = \mu, \int_{R^n} \xi\xi^T p(\mathrm{d}\xi) = \Sigma + \mu\mu^T \right\},$$

where M_+ is the cone set of nonnegative Borel measure on R^n. For simplicity, we denote

$$\Omega = \begin{bmatrix} \Sigma + \mu\mu^T & \mu \\ \mu^T & 1 \end{bmatrix}$$

as the matrix of second-order moment of ξ and assume p is the probability measure which satisfies $p \in P$.

In order to evaluate the model, we define a probability distribution set P which contains the true distribution F. The model is reformulated to consider the worst case on expected reward and risk in this set. Therefore, we introduce the definitions of robust reward functions and robust risk measurements under the worst case of the distribution set P.

Definition 2.1. Denote $x \in R^n$ as the decision vector, ξ as the corresponding reward vector, and P as a given distribution set including the probability measure p of ξ. Let the reward measure function, CVaR measurement, and SD measurement be $E_p(*)$, $\mathrm{CVaR}_p(*)$, and $\mathrm{std}_p(*)$, respectively. Then robust reward, robust risk including both CVaR and SD are defined as follows:

$$E^W(*) \equiv \inf_{p \in P} E_p(*), \; [\mathrm{CVaR}(*) + \mathrm{std}(*)]^W \equiv \sup_{p \in P} [\mathrm{CVaR}_p(*) + \mathrm{std}_p(*)],$$

where the "W" in superscript represents the "worst case."

Based on the assumption of compactness on the set P, the infimum and supremum can be rewritten by minimum and maximum, respectively. That is

$$E^W(*) \equiv \min_{p \in P} E_p(*), \; [\mathrm{CVaR}(*) + \mathrm{std}(*)]^W \equiv \max_{p \in P} [\mathrm{CVaR}_p(*) + \mathrm{std}_p(*)].$$

Combining with the classical robust reward-risk model in [29–31] and mean-variance optimization (MVO) model [32], we present a new model of robust reward-risk optimization under an uncertain probability distribution,

$$\begin{aligned} \max_x \quad & E^W(x) \\ \text{s.t.} \quad & [\mathrm{CVaR}(x) + \mathrm{std}(x)]^W \le R^*, \\ & x \in \mathscr{X}, \end{aligned} \tag{2}$$

where both CVaR and SD are considered as risk measurements.

The model protects against the probability distribution in the worst case scenario simultaneously taken in objective and constraints. As presented in [31,33], it is assumed to be overconservative.

From the model above, the constant R^* suppose to be priori known at first. Actually, for pursuing maximum returns, a minimum acceptable expected return may not be determined by investors. Similarly, a maximum risk aversion bound may not be estimated as well. So it is tough to choose a suitable value in practice. To address this problem, such ratio optimization models are proposed, see Sharpe ratio model for example. We will study the ratio optimization model in the next section.

3 Ratio model of distributionally robust optimization

We first present a ratio model to avoid choosing the prior parameters in (2). Referring to the idea in [31] and the optimal Sharpe ratio model, we introduce the ratio optimization model associated with model (2). Since the set P is compact, there is a type of ratio optimization model as follows:

$$\begin{aligned} \max_x \quad & \frac{\min_{p\in P} E_p(x)}{\max_{p\in P}[\mathrm{CVaR}_p(x) + \mathrm{std}_p(x)]} \\ \text{s.t.} \quad & x \in \mathscr{X}. \end{aligned} \tag{3}$$

Unlike model (2), the ratio model does not involve the bounds of minimum acceptable expected return or maximum risk aversion, which need decision-makers to decide. According to Markowitz' mean-variance theorem, the ratio model should give an efficient solution. In other words, the decision should be on the efficient frontier of original model [34].

Next, we show that the ratio model involving both CVaR and SD is efficient in Markowitz' mean-variance principle and it decomposes the difficulty of trade-off over reward and risk.

Proposition 3.1. *For any distribution $p \in P$, the distributionally robust ratio model* (3) *gives a solution on the efficient frontier of the corresponding robust mean-variance model.*

Proof. On the contrary, we assume that the obtained portfolio x is not on the efficient frontier [34]. Then, a portfolio $\hat{x} \in \mathscr{X}$ satisfies $E_p(x) \le E_p(\hat{x})$ and $\mathrm{std}_p(x) \ge \mathrm{std}_p(\hat{x})$, and at least one inequality holds strictly, for example

$$E_p(x) < E_p(\hat{x}). \tag{4}$$

Assuming the probability distribution $F(\cdot)$ is standard normal and $f(\cdot)$ is its density function, we have [35]

$$\mathrm{CVaR}_p = T\mathrm{std}_p^2 - E_p,$$

where, for a confidence level α, the constant $T = f[F^{-1}(\alpha)]/(1-\alpha)$. Hence, the following inequality is easily obtained

$$\mathrm{CVaR}_p(x) + \mathrm{std}_p(x) > \mathrm{CVaR}_p(\hat{x}) + \mathrm{std}_p(\hat{x}). \tag{5}$$

Combining (4) with (5), we obtain

$$\begin{gathered} E^W(x) < E^W(\hat{x}), \\ [\mathrm{CVaR}(x) + \mathrm{std}(x)]^W > [\mathrm{CVaR}(\hat{x}) + \mathrm{std}(\hat{x})]^W. \end{gathered}$$

That is

$$\frac{E^W(x)}{[\mathrm{CVaR}_\alpha(x) + \mathrm{std}(x)]^W} < \frac{E^W(\hat{x})}{[\mathrm{CVaR}_\alpha(\hat{x}) + \mathrm{std}(\hat{x})]^W},$$

which is contrary to the assumption that x is the optimum. So model (3) gives a solution on robust efficient frontier. □

Obviously, the computation of the aforementioned ratio model is complicate, but reasonable in practical finance problems. Next, we will analyse the model and transform it to a tractable model. First, we introduce some concave and convex properties of robust reward, CVaR and SD. Referring to the propositions in [30,36], we present an assumption according to distributionally robust optimization principle.

Proposition 3.2. (Proposition 2 in [36]) *If μ is a concave function and ρ is a convex function, then*

(a) *the ratio μ/ρ is quasi-concave*;

(b) *the ratio ρ/μ is quasi-convex*;

(c) *the following relationship holds*

$$\arg\max_x \frac{\mu(x)}{\rho(x)} = \arg\min_x \frac{\rho(x)}{\mu(x)}.$$

Proposition 3.3. (Proposition 2 in [30]) *If ρ is a coherent risk measure associated with crisp probability measure p, then the worst case risk measure ρ_W associated with ambiguous probability measure P remains a coherent risk measure.*

Here, a risk measure that satisfies the four desirable axioms, that is Subadditivity, Positive homogeneity, Monotonicity, and Translation invariance, is called a coherent risk measure.

Assumption 1. Denote $E_p(\cdot)$, $\mathrm{CVaR}_p(\cdot)$, and $\mathrm{std}_p(\cdot)$ as measure functions of the common reward, CVaR and SD with measure p, respectively. Assume that the following three statements are satisfied:
(a) $E_p(\cdot)$ is a positive function on R^n and has properties of positive homogeneity and convexity;
(b) $\mathrm{CVaR}_p(\cdot)$ is a positive function on R^n and satisfies positive homogeneity and sub-additivity;
(c) The SD measure function is a positive function as well as $\mathrm{CVaR}_p(\cdot)$.

According to the specific form of reward function, many reward measures satisfy this assumption.

Since we have assumed the compactness of distribution set P, the robust measures in the worst case retain inherent properties of the above three measures.

Proposition 3.4. *Suppose the above three measurement functions satisfy Assumption 1 with a compact distribution set P, then the worst case reward E^W and robust CVaR and SD $[\mathrm{CVaR}+\mathrm{std}]^W$ satisfy*

$$E^W(tx) = tE^W(x), \quad t > 0,$$
$$[\mathrm{CVaR}(tx) + \mathrm{std}(tx)]^W = t[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W, \quad t > 0.$$

Proof. It is easy to know that "inf" and "sup" operations in robust measures are equivalent to "min" and "max" since P is compact. Referring to Proposition 3.3, we can obtain the conclusion about E^W straightforwardly. In addition, according to Assumption 1,

$$\begin{aligned}[\mathrm{CVaR}(tx) + \mathrm{std}(tx)]^W &= \sup_{p\in P}[\mathrm{CVaR}_p(tX) + \mathrm{std}_p(tX)] \\ &= \max_{p\in P} t[\mathrm{CVaR}_p(X) + t\mathrm{std}_p(X)] \\ &= t\max_{p\in P}[\mathrm{CVaR}_p(X) + \mathrm{std}_p(X)] \\ &= t[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W.\end{aligned}$$

□

Let the constraint set $\mathcal{X}$ in (1) be convex. Then, the convexity of the worst case measurements and objective functions can be induced.

Proposition 3.5. *Let E_p, CVaR_p, and std_p: $\mathcal{X} \subset R^n \to R^{++}$ with respect to $p \in P$ and P be a compact distribution set satisfying Assumption 1, the concavity and convexity of E^W and $[\mathrm{CVaR}+\mathrm{std}]^W$ are straightforward, respectively. In addition, it holds*

$$\textit{the ratio } \frac{E^W(x)}{[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W} : \mathcal{X} \to R^{++} \textit{ is quasi-concave.}$$

Proof. We know that CVaR_p and std_p satisfy sub-additive from Assumption 1. Let

$$f_p(x) = \mathrm{CVaR}_p(x) + \mathrm{std}_p(x).$$

For $\forall \alpha \in [0,1]$, we have

$$\begin{aligned} f_p(\alpha x_1 + (1-\alpha)x_2) &= \mathrm{CVaR}_p(\alpha x_1 + (1-\alpha)x_2) + \mathrm{std}_p(\alpha x_1 + (1-\alpha)x_2) \\ &\leq \mathrm{CVaR}_p(\alpha x_1) + \mathrm{CVaR}_p((1-\alpha)x_2) + \mathrm{std}_p(\alpha x_1) + \mathrm{std}_p((1-\alpha)x_2). \end{aligned}$$

And CVaR_p and std_p satisfy positive homogeneity, we have

$$\begin{aligned} f_p(\alpha x_1 + (1-\alpha)x_2) &\leq \alpha\,\mathrm{CVaR}_p(x_1) + (1-\alpha)\mathrm{CVaR}_p(x_2) + \alpha\mathrm{std}_p(x_1) + (1-\alpha)\mathrm{std}_p(x_2) \\ &= \alpha(\mathrm{CVaR}_p(x_1) + \mathrm{std}_p(x_1)) + (1-\alpha)(\mathrm{CVaR}_p(x_2) + \mathrm{std}_p(x_2)) \\ &= \alpha f(x_1) + (1-\alpha)f(x_2). \end{aligned}$$

Through the definition of convex function, the conclusion on $[\mathrm{CVaR}_p + \mathrm{std}_p]$ is easily obtained with measure p. On the other hand, we can easily obtain the concavity of E_p. According to Proposition 3.3, the concavity of E^W and convexity of $[\mathrm{CVaR} + \mathrm{std}]^W$ are easily induced. The final conclusion on the ratio can be straightforwardly obtained in terms of Proposition 3.2. □

According to the above propositions, the following statement can be presented that the ratio model (3) can be further transformed into a reduced form.

Proposition 3.6. *Suppose E_p, CVaR_p, and std_p satisfy Assumption 1 with a compact set P. Then distributionally robust ratio model* (3) *can be rewritten equivalently to*

$$\begin{aligned} \max_{w,t} \quad & E^W(w) \\ s.t. \quad & [\mathrm{CVaR}(w) + \mathrm{std}(w)]^W \leq 1, \\ & e^T w = t, \;\; tL_b \leq Bw \leq tU_b, \;\; t \geq 0. \end{aligned} \tag{6}$$

Denote the pair (w^*, t^*) as a solution to (6), *then $x^* = \frac{w^*}{t^*}$ solves the ratio model* (3). *Conversely, if x^* is the optimal decision of* (3), *and let $t^* = \frac{1}{[\mathrm{CVaR}(x^*) + D(x^*)]^W}$, then (t^*x^*, t^*) solves the equivalent problem* (6).

Proof. First, we verify the equivalence of models (3) and (6). It would be proven in two steps. In the first step, model (3) is clarified to equivalent to

$$\begin{aligned} \max_{w,t} \quad & E^W(w) \\ \text{s.t.} \quad & [\mathrm{CVaR}(w) + \mathrm{std}(w)]^W = 1, \\ & e^T w = t, \;\; tL_b \leq Bw \leq tU_b, \;\; t \geq 0. \end{aligned} \tag{7}$$

Let $t = \frac{1}{[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W}$ which is obviously not less than 0. For any feasible point x of model (3), let $w = tx$ and according to the positive homogeneity of E_p, CVaR_p, and std_p, we can straightforwardly obtain that (w, t) is feasible for model (7). Therefore,

$$\max_{x \in \mathscr{X}} \frac{E^W(x)}{[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W} \leq \max_{(w,t) \in \mathscr{X}_1} E^W(w),$$

where $\mathscr{X}$ and $\mathscr{X}_1$ are the respective sets of constraints for (3) and (7). On the other hand, suppose one feasible solution of (7) is (w, t), we can obtain $x = \frac{w}{t}$ is feasible for (3) and

$$\max_{x \in \mathscr{X}} \frac{E^W(x)}{[\mathrm{CVaR}(x) + \mathrm{std}(x)]^W} \geq \max_{(w,t) \in \mathscr{X}_1} E^W(w).$$

Combining the above inequalities, we can easily obtain that the optimal function values of (3) and (7) are the same. In addition, the optimal solutions of models (3) and (7) are formed as x^* and (t^*x^*, t^*) (or $x^* = \frac{w^*}{t^*}$ and (w^*, t^*)), respectively, where $t^* = \frac{1}{[\mathrm{CVaR}(x^*) + \mathrm{std}(x^*)]^W}$.

Second, the equivalent of problems (6) and (7) should be proved. We focus on a solution of (6) denoted as (w^*, t^*), and then prove $[\mathrm{CVaR}(w^*) + \mathrm{std}(w^*)]^W = 1$. It should be noted that two problems (6) and (7) are different only in the first constraint (i.e. one is inequality and the other is equality). Denote (w^0, t^0) as one optimal solution of problem (6). On the contrary, suppose (w^0, t^0) satisfies

$$[\mathrm{CVaR}(w^0) + \mathrm{std}(w^0)]^W < 1.$$

According to Assumption 1, there exists a constant a satisfying

$$([\mathrm{CVaR}(w^0) + \mathrm{std}(w^0)]^W)^{-1} \geq a > 1.$$

In addition, since the positive homogeneity of measures in 3.4, we obtain

$$a[\mathrm{CVaR}(w^0) + \mathrm{std}(w^0)]^W = [\mathrm{CVaR}(aw^0) + \mathrm{std}(aw^0)]^W \leq 1.$$

Obviously, (aw^0, at^0) is feasible for problem (6).

We focus on the objective value of (6) at point (aw^0, at^0) with $a > 1$, we have

$$E^W((aw^0)^T\beta) = aE^W((w^0)^T\beta) > E^W((w^0)^T\beta).$$

Clearly, it is contrary to the assumption that (w^0, t^0) is one optimum of problem (6). So the first constraint is equality at the optimum. Therefore, the equivalence between (6) and (7) is proved. In summary, we prove two models (3) and (6) are equivalent. For more details refer to [36]. □

The equivalent model (6) has good properties (like convexity) to analysis, reformulation, and computation for distributionally robust problems. We will discuss in next part from optimization viewpoint.

4 An semi-definite programming (SDP) reformulation for the distributionally robust model

In this part, we will use moment theorem to transform the objective and constraints to the form of semi-definite programming which is tractable in computation.

In Section 1, it is known that P denotes a probability distribution set, in which the distributions have the same first and second moments as the true distribution F of the assets' reward ξ. Referring to [30], the reward objective, SD and CVaR function can be written as

$$E(X) = \int_{R^n} x^T\xi p(\mathrm{d}\xi), \quad \mathrm{std}(X) = \|x^T\Sigma^{1/2}\|, \quad \mathrm{CVaR}(X) = \min_{\beta} \beta + \frac{1}{1-\alpha}\int_{R^n}(-x^T\xi - \beta)^+ p(\mathrm{d}\xi).$$

Then model (6) has the following form:

$$\begin{aligned} &\max_{w,t} \ \min_{p\in P} \int_{R^n} w^T\xi p(\mathrm{d}\xi) \\ &\text{s.t.} \ \max_{p\in P}\min_{\beta} \beta + \frac{1}{1-\alpha}\int_{R^n}(-w^T\xi - \beta)^+ p(\mathrm{d}\xi) + \|w^T\Sigma^{1/2}\| \leq 1, \\ &\qquad w^Te = t, \ t \geq 0, \ tL_b \leq Bw \leq tU_b, \end{aligned} \tag{8}$$

which is the so-called robust "counterpart" in optimization.

From the above model, we easily know that the equivalent problem has complicated parts with max-min in objective and constraints. Therefore, optimization techniques will be used to simplify the model in next part. The dual variables $h_0 \in R$, $h \in R^n$, and $H \in R^{n\times n}$ are introduced to the equality constraints in the objective function of (8), respectively. Similarly, introduce dual variables $k_0 \in R$, $k \in R^n$, and $K \in R^{n\times n}$ to the equalities of the risk constraint. We transfer the minimization in objective and maximization in constraints to their dual problems, respectively.

Proposition 4.1. *The bi-level max-min problem* (8) *can be reformulated to*

$$\begin{aligned}\max_{M_1,w,t}\quad & \langle\Omega, M_1\rangle\\ \text{s.t.}\quad & \varphi(w)\le 1,\\ & \begin{bmatrix} 0 & \frac{1}{2}w\\ \frac{1}{2}w^T & 0\end{bmatrix} - M_1 \ge 0,\\ & w^Te=t,\ \ t\ge 0,\ \ tL_b\le Bw\le tU_b,\end{aligned}\tag{9}$$

where

$$M_1=\begin{bmatrix} H & \frac{1}{2}h\\ \frac{1}{2}h^T & h_0\end{bmatrix}\in R^{(n+1)\times(n+1)},\quad h_0\in R,\quad h\in R^n,\quad \text{and}\quad H\in R^{n\times n}.$$

$\varphi(w)$ is the minimal value of the following problem:

$$\begin{aligned}\min_{\beta,M_2}\quad & \beta+\|w^T\Sigma^{1/2}\|+\frac{1}{1-\alpha}\langle\Omega, M_2\rangle\\ & \text{s.t. } M_2+\begin{bmatrix} 0 & \frac{1}{2}w\\ \frac{1}{2}w^T & \beta\end{bmatrix}\ge 0,\quad M_2\ge 0,\end{aligned}\tag{10}$$

where

$$M_2=\begin{bmatrix} K & \frac{1}{2}k\\ \frac{1}{2}k^T & k_0\end{bmatrix}\in R^{(n+1)\times(n+1)},\quad k_0\in R,\ \ k\in R^n,\ \ \text{and } K\in R^{n\times n}.$$

Proof. First, we focus on the value of objective function in robustness. The objective function in the worst case scenario is actually a minimization problem over the probability variable p,

$$\begin{aligned}\min_{p\in M_+}\quad & \int_{R^n} w^T\xi p(\mathrm{d}\xi)\\ \text{s.t.}\quad & \int_{R^n} p(\mathrm{d}\xi)=1,\\ & \int_{R^n} \xi p(\mathrm{d}\xi)=\mu,\\ & \int_{R^n} \xi\xi^T p(\mathrm{d}\xi)=\Sigma+\mu\mu^T,\end{aligned}\tag{11}$$

where M_+ denotes the nonnegative Borel measure cone in R^n.

Actually, the optimal variable of the semi-infinite programming (13) is also a nonnegative measure p. Specifically, the first equality ensures that p is a probability measure. The other two equalities are consistent with the known first and second moments, respectively.

From the dual theorem (Lemma A.1 in [14]), the dual problem of (11) can be obtained,

$$\begin{aligned}\max_{h_0,h,H}\quad & h_0+h^T\mu+\langle H,\Sigma+\mu\mu^T\rangle\\ \text{s.t.}\quad & h_0+h^T\xi+\langle H,\xi\xi^T\rangle\le w^T\xi,\ \ \forall\xi\in R^n,\end{aligned}\tag{12}$$

where $h_0 \in R, h \in R^n$, and $H \in R^{n\times n}$. Since $\Sigma > 0$, the property of strong duality holds [37]. Then the primal (11) and dual (12) are solvable and have the equal optimal objective value. By defining the combined variable

$$M_1 = \begin{bmatrix} H & \frac{1}{2}h \\ \frac{1}{2}h^T & h_0 \end{bmatrix} \in R^{(n+1)\times(n+1)}, \quad h_0 \in R,$$

we rewrite the dual form (12) by matrix inner production,

$$\begin{aligned} \max_{M_1 \in R^{(n+1)\times(n+1)}} \quad & \langle \Omega, M_1 \rangle \\ \text{s.t.} \quad & [\xi^T 1] M_1 [\xi^T 1]^T \leq w^T \xi, \quad \forall \xi \in R^n. \end{aligned} \tag{13}$$

It should be noted that the inequality in (13) can be transformed to the SDP form:

$$[\xi^T \ 1]\left(\begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & 0 \end{bmatrix} - M_1\right)[\xi^T \ 1]^T \geq 0, \ \forall \xi \in R^n \Leftrightarrow \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & 0 \end{bmatrix} - M_1 \geq 0.$$

Therefore, the objective part in (8) can be reformulated to

$$\begin{aligned} \max_{M_1} \quad & \langle \Omega, M_1 \rangle \\ \text{s.t.} \quad & \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & 0 \end{bmatrix} - M_1 \geq 0. \end{aligned} \tag{14}$$

Second, we consider the robust constraint function

$$\max_{p\in P} \min_{\beta} \beta + \frac{1}{1-\alpha} \int_{R^n} (-w^T\xi - \beta)^+ p(\mathrm{d}\xi) + \|w^T \Sigma^{1/2}\|.$$

According to the classical saddle point theorem, the constraint function can be equivalently expressed as

$$\begin{aligned} &\max_{p\in P} \min_{\beta} \beta + \frac{1}{1-\alpha} \int_{R^n} (-w^T\xi - \beta)^+ p(\mathrm{d}\xi) + \|w^T \Sigma^{1/2}\| \\ &= \min_{\beta} \max_{p\in P} \beta + \frac{1}{1-\alpha} \int_{R^n} (-w^T\xi - \beta)^+ p(\mathrm{d}\xi) + \|w^T \Sigma^{1/2}\| \\ &= \min_{\beta} \beta + \|w^T \Sigma^{1/2}\| + \frac{1}{1-\alpha} \max_{p\in P} \int_{R^n} (-w^T\xi - \beta)^+ p(\mathrm{d}\xi). \end{aligned}$$

It should be clarified that the interchange of “max” and “min” is allowed according to the saddle point theorem proposed by Shapiro and Kleywegt [38] (see also Natarajan et al. [39] or Delage and Ye [18], for example).

Next, we derive an SDP reformulation to the “max” part of the above equation, which is an expectation maximizing problem on the worst case probability distribution:

$$\begin{aligned} \max_{p\in M_+} \quad & \int_{R^n} (-w^T\xi - \beta)^+ p(\mathrm{d}\xi) \\ \text{s.t.} \quad & \int_{R^n} p(\mathrm{d}\xi) = 1, \\ & \int_{R^n} \xi p(\mathrm{d}\xi) = \mu, \\ & \int_{R^n} \xi\xi^T p(\mathrm{d}\xi) = \Sigma + \mu\mu^T, \end{aligned} \tag{15}$$

where M_+ represents the nonnegative Borel measures on R^n. Similar to the above derivation, we derive the dual form of (15) by the dual theorem

$$\begin{aligned} \min_{k_0,k,K} \quad & k_0 + k^T\mu + \langle K, \Sigma + \mu\mu^T\rangle \\ \text{s.t.} \quad & k_0 + k^T\xi + \langle K, \xi\xi^T\rangle \geq (-w^T\xi - \beta)^+, \ \ \forall \xi \in R^n, \end{aligned}$$

where $k_0 \in R$, $k \in R^n$, and $K \in R^{n\times n}$. So the constraint on the worst case CVaR and SD can be equivalent to the following SDP:

$$\begin{aligned} \min_{\beta,M_2} \quad & \beta + \|w^T\Sigma^{1/2}\| + \frac{1}{1-\alpha}\langle \Omega, M_2\rangle \\ \text{s.t.} \quad & [\xi^T \ 1]M_2[\xi^T \ 1]^T \geq (-w^T\xi - \beta)^+, \ \ \forall \xi \in R^n, \end{aligned} \tag{16}$$

where

$$M_2 = \begin{bmatrix} K & \frac{1}{2}k \\ \frac{1}{2}k^T & k_0 \end{bmatrix} \in R^{(n+1)\times(n+1)}.$$

Furthermore, the inequality in (16) can be rewritten as two inequality constraints:

$$\begin{aligned} &[\xi^T \ 1]M_2[\xi^T \ 1]^T \geq -w^T\xi - \beta, \ \ \forall \xi \in R^n, \\ &[\xi^T \ 1]M_2[\xi^T \ 1]^T \geq 0, \ \ \forall \xi \in R^n. \end{aligned}$$

Equivalently,

$$[\xi^T \ 1]M_2[\xi^T \ 1]^T \geq -w^T\xi - \beta, \ \ \forall \xi \in R^n \Leftrightarrow M_2 + \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & \beta \end{bmatrix} \geq \ 0.$$

In summary, the constraint's dual form is obtained as

$$\begin{aligned} \min_{\beta,M_2} \quad & \beta + \|w^T\Sigma^{1/2}\| + \frac{1}{1-\alpha}\langle \Omega, M_2\rangle \\ \text{s.t.} \quad & M_2 + \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & \beta \end{bmatrix} \geq \ 0, \\ & M_2 \geq 0. \end{aligned} \tag{17}$$

By substituting (14) and (17) into the objective and the first constraint of model (8), respectively, we can easily obtain model (9). □

Note that the subproblem (10) is still very complicated. In the next part, we further simplify the subproblem. Denote

$$u = (w, t, M_1), \quad U = (M_1, M_2, w, t, \beta).$$

The single-layer minimization problem is presented as follows:

$$\begin{aligned}
\max_{M_1,M_2,w,t,\beta} \quad & \langle \Omega, M_1 \rangle \\
\text{s.t.} \quad & \beta + \|w^T \Sigma^{1/2}\| + \frac{1}{1-\alpha}\langle \Omega, M_2 \rangle \le 1, \\
& M_2 + \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & \beta \end{bmatrix} \succeq 0, \quad M_2 \succeq 0, \\
& \begin{bmatrix} 0 & \frac{1}{2}w \\ \frac{1}{2}w^T & 0 \end{bmatrix} - M_1 \succeq 0, \\
& w^T e = t, t \ge 0, \ tL_b \le Bw \le tU_b.
\end{aligned} \tag{18}$$

Obviously, (18) is more tractable in computation than (9). The following proposition is presented to illustrate solutions of two models referring to Proposition 4.1 in [29].

Proposition 4.2. *If* $U^* = (M_1^*, M_2^*, w^*, t^*, \beta^*)$ *is an optimum of* (18)*, then* $u^* = (M_1^*, w^*, t^*)$ *is an optimum of* (9). *Conversely, if* $u' = (w', t', M_1')$ *is an optimum of* (9)*, then* $U' = (M_1', M_2', w', t', \beta')$ *is an optimum of* (18)*, where* (M_2', β') *is an optimum of model* (10) *with* $(M_2, \beta) = (M_2', \beta')$.

Proof. Let $U^* = (M_1^*, M_2^*, w^*, t^*, \beta^*)$ be an optimum solution of model (18), then the pair $u^* = (M_1^*, w^*, t^*)$ can be proved to be an optimum of problem (9). First, we prove that $u^* = (M_1^*, w^*, t^*)$ is feasible as follows.

Since $U^* = (M_1^*, M_2^*, w^*, t^*, \beta^*)$ is an optimum solution of model (18), it can be obtained that

$$\begin{aligned}
& \beta^* + \|w^{*T} \Sigma^{1/2}\| + \frac{1}{1-\alpha}\langle \Omega, M_2^* \rangle \le 1, \\
& M_2^* + \begin{bmatrix} 0 & \frac{1}{2}w^* \\ \frac{1}{2}w^{*T} & \beta^* \end{bmatrix} \succeq 0, \quad M_2^* \succeq 0.
\end{aligned}$$

It is obvious that $\varphi(w^*) \le 1$ because of definition of $\varphi(w^*)$. In addition to other constraints in model (18), we can easily show that $u^* = (M_1^*, w^*, t^*)$ is feasible for problem (9).

Second, we show that $u^* = (M_1^*, w^*, t^*)$ is an optimum solution of (9). Suppose $u^* = (M_1^*, w^*, t^*)$ is not an optimum solution of (9), there must exist an optimum $\hat{u} = (\hat{w}, \hat{t}, \hat{M}_1)$. Denote values of objective function in maximization problem (9) as $\phi(w, t, M_1)$, then

$$\phi(\hat{w}, \hat{t}, \hat{M}_1) \ge \phi(w^*, t^*, M_1^*). \tag{19}$$

Therefore, for the first inequality of (9) related to subproblem (10), there exists a solution $(\hat{\beta}, \hat{M}_2)$ of the subproblem (10) with $w = \hat{w}$. It is obvious that $\hat{U} = (\hat{w}, \hat{t}, \hat{M}_1, \hat{\beta}, \hat{M}_2)$ is feasible to model (18). But the inequality of objective values (19) implies that U^* is not a solution of (18). It is a contradiction to prior assumption. So $u^* = (M_1^*, w^*, t^*)$ is an optimum solution of (9).

Conversely, denote $u' = (M_1', w', t')$ as an optimum solution of model (9), $U' = (M_1', M_2', w', t', \beta')$ is proven to be a solution of (18). Here, (β', M_2') is an optimum solution of subproblem (10) with $w = w'$.

In fact, suppose that $U' = (M_1', M_2', w', t', \beta')$ is not an optimum solution of problem (18), there exists a solution denoted as $U'^* = (M_1'^*, M_2'^*, w'^*, t'^*, \beta'^*)$. According to the above proof, we obtain that the pair $(M_1'^*, w'^*, t'^*)$ is a solution of (9), and the following inequality of objective values

$$\phi(M_1'^*, w'^*, t'^*) \ge \phi(M_1', w', t'),$$

which contradicts to that $u' = (M_1', w', t')$ is an optimum of problem (9). Hence, the conclusion is obtained. □

It is obvious that the final version (18) is actually a SDP model, which is tractable in computation.

5 Application in portfolio problems

In this section, the final distributionally robust reward-risk model (18), which considers both CVaR and SD in risk measurement, is used to solve portfolio problems. In the first part, the proposed model is experimented on two market data to validate its robustness and effectiveness. Next, the corresponding Mean-CVaR ratio model without SD is compared to show the distinction with the proposed model through simulation analysis.

5.1 Model verification on two market data

In this test, we verify the proposed model (18) on the portfolio problem with two market. First, the market data and parameters are given. Second, we output the results of the distributionally robust model vs the robust model under the ellipsoid uncertainty. Finally, the proposed model is compared with the Mean-CVaR ratio model without SD in [30,31,33] to show the validation of our motivation.

5.1.1 Market data and settings

We use the examples of real Hang Seng Index data referring to [28,30,40]. The two sectoral sub-indices: Hang Seng Finance Index (HSNF) and Hang Seng Utilities Index (HSNU) are selected as the financial assets to construct the portfolios. The history data (1,000 samples) of their daily returns are chosen from January 3, 2017 to January 22, 2021.

– HSNF: the first 500 samples in the history data have mean return 1.1455 with 0.4589 SD; the latter half 500 daily returns have the average 0.7264 with 0.4389 SD.
– HSNU: the 500 samples in the first period of history data have mean return 0.7522 with 0.1412 SD; the latter half 500 ones have mean return 0.3648 with 0.1270 SD.

Conveniently, according to the feasibility of portfolio problems. The parameters in problem (18) are set as:

$$B = \begin{bmatrix} 22 & 4 \\ 5 & 26 \end{bmatrix}, \quad L_b = (0, 0)^T, \quad U_b = (14, 13)^T, \quad \alpha = 0.95.$$

Besides, define P as the distribution set of the measure p on R^n with the same first- and second-order moments as the true distribution F. In this test, the first-order moments μ is estimated by calculating the mean history returns of two assets. Since the estimation errors between the sample covariance matrix $\hat{\Sigma}$ and the true Σ may probably accumulate. We use the shrinkage method in [41,42] to estimate the second-order moments Σ, which linearly combines the sample covariance matrix $\hat{\Sigma}$ and a simple definite matrix.

5.1.2 Contrast with the robust model under the ellipsoid uncertainty P_π

In this comparison, we focus on the robust reward-risk ratio models in [28,30,31,33] under an ellipsoid uncertainty P_π. It should be noted that the robust models in [30,31,33] do not contain SD as a part of risk measure. For the sake of fairness, we add this part to their models as follows:

$$\begin{aligned} \max_{w,t} \quad & \min_{p\in P_\pi} E_p(w) \\ \text{s.t.} \quad & \max_{p\in P_\pi} \mathrm{CVaR}_p(w) + \mathrm{std}_p(w) \le 1, \\ & e^T w = t, \ \ tL_b \le Bw \le tU_b, \ \ t \ge 0, \end{aligned} \tag{20}$$

where $P_\pi = \{p : p = \pi^0 + A\eta, \pi^0 + A\eta \geq 0, e^T A\eta = 0, \|\eta\| \leq 1\}$. Actually, the contrasting model (20) is similar with the first case of robust reward-risk model in [28]. We can see that model (20) and our model (18) have the same economic meaning but under the different uncertainty sets of assets' reward variable ξ.

The parameters in (20) are chosen as follows:

$$A = \begin{bmatrix} 3 & 7 \\ 4 & 6 \end{bmatrix}, \quad \pi^0 = (0.5, 0.5)^T.$$

The other settings are the same as the above in 5.1.1.

The semidefinite programming problems (20) and (18) can be solved by SDPT3 solver. The allocation obtained with the proposed model is (0.6172, 0.3828), which is different from the solution (0.5257, 0.4743) of model (18). In other words, for the same problem, the portfolio choice will be different under different uncertainty distribution sets. So the investors would choose some preferable uncertainty distribution sets of variables with personal preference and practical situation.

5.1.3 Comparison with distributionally robust Mean-CVaR ratio model

In this experiment, we compare (18) with the following Mean-CVaR ratio model,

$$\begin{aligned} \max_{w,t} \quad & E^W(w^T\xi - tr_b) \\ \text{s.t.} \quad & \mathrm{CVaR}^W(w^T\xi - tr_b) \leq 1, \\ & e^T w = t, \ \ tL_b \leq Bw \leq tU_b, \ \ t \geq 0. \end{aligned} \tag{21}$$

which does not involve standard variance in [31]. Here, the parameter r_b represents the risk-free reward.

Since the risk-free reward is not concerned in this article, we set $r_b = 0$ for consistency with our model. Then, (21) is reduced to the problem with the uncertainty distribution set P,

$$\begin{aligned} \max_{w,t} \quad & \min_{p \in P} E_p(w) \\ \text{s.t.} \quad & \max_{p \in P} \mathrm{CVaR}_p(w) \leq 1, \\ & e^T w = t, \ \ tL_b \leq Bw \leq tU_b, \ \ t \geq 0. \end{aligned} \tag{22}$$

We firstly evaluate the performance of models (22) and (18) on history rewards in Section 5.1.1. Numerically, model (22) gives an allocation of assets (0.6580, 0.3420). Meanwhile, the result obtained with our model is (0.5257, 0.4743).

Now we apply the above two allocations to future data (45 days) of HSNF and HSNU, which are shown in Figure 1.

From Figure 1, it is obvious that the volatilities of two returns are distinct to the history data. Furthermore, we compute the expected returns of the next 45 transactions with the above two allocations, respectively. Figure 2 outputs the comparison of the next 45 daily expected returns obtained by model (22) and our model (18).

As shown in Figure 2, two allocations obtained by (22) and (18) result in similar mean returns around 0.75. However, the corresponding volatilities are distinct apparently. Furthermore, the corresponding mean returns, variance, and CVaR of the obtained portfolios are calculated as shown in Table 1, respectively. It should be noted that both variance and CVaR by the proposed model (18) are smaller than the distributionally robust Mean-CVaR model (22). This indicates that the risk measurement combining SD and CVaR is actually feasible and effective.

In other words, the rewards in the worst case scenario with respect to the proposed model are merely a little bit smaller than that of model (22). It also shows that the robustness of our model. Consequently, the proposed model may be applicable as a competitive and reasonable portfolio strategy for conservative investors.

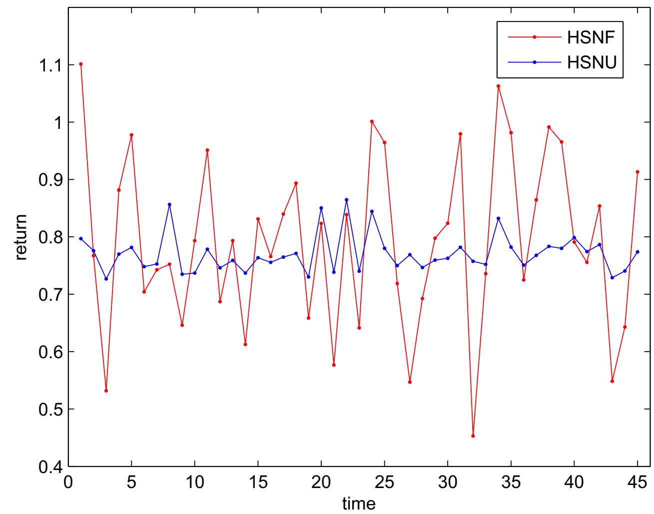

Figure 1: Future returns of two assets in next 45 transactions.

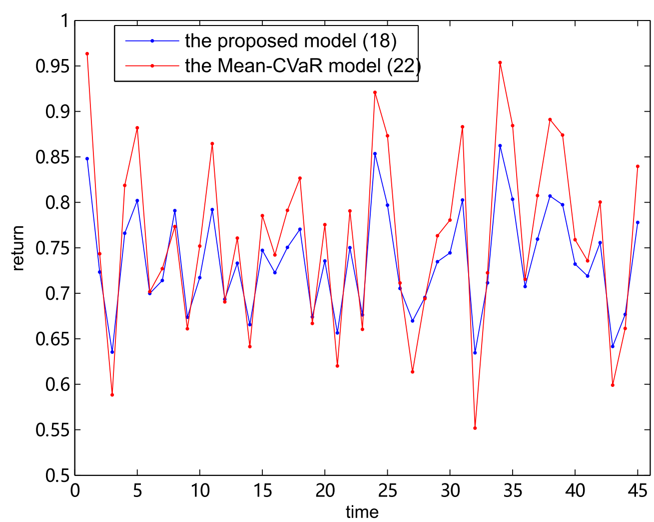

Figure 2: Comparison on the predicted returns of models (18) and (22).

Table 1: Comparison of mean return, variance, and CVAR of two portfolios for future returns

Method	Mean return	Variance	CVaR
The proposed model (18)	0.7485	0.1908	0.8609
The Mean-CVaR model (22)	0.7747	0.2889	0.9416

5.2 Simulation analysis

In this subsection, we present two examples to illustrate the distributionally robust portfolio problems with vs without considering SD. Here, we only focus on the comparison with the distributionally robust Mean-CVaR model (22) through market data simulation and Monte Carlo simulation.

5.2.1 Market data simulation analysis

To verify the effectiveness of the proposed model, we introduce four real-world datasets, including Ken French's Fama French with 25 assets (FF25) and 100 assets (FF100), SP500 index, and Shanghai Stock Exchange Composite Price Index (SHCI). The four assets are collected from the datasets FF25, FF100, SP500, and SHCI from the time period from February 7, 2022 to January 31, 2023. The mean of daily returns and the covariance matrix in this example are given in Tables 2 and 3, respectively.

We solve the semidefinite programming problem (18) and (22). Here, we omit the box constraint on weights of assets. The obtained optimal portfolios are displayed in Table 4, respectively.

Then, for the confidence level 0.95 (α = 0.05), we calculate the mean return, variance, VaR, and CVaR of two portfolios, respectively. The obtained results are listed in Table 5.

From Table 5, we can find out that the presented model leads to a presumptive variance (about 10% smaller) and very similar mean return, VaR, and CVaR. This observation is consistent with the corresponding optimal portfolios in Table 4. For instance, since the variance of SHCI is more than twice as much as the others, the weight obtained by our model is indeed much lower than that by the robust Mean-CVaR model. It verifies the effectiveness on involving the standard deviance in the proposed model.

Table 2: Portfolio mean returns

	Mean return
FF25	0.0002689
FF100	0.0003391
SP500	0.0002141
SHCI	0.0004857

Table 3: Covariance matrix of returns

	FF25	FF100	SP500	SHCI
FF25	0.0003479	0.0002463	−0.0000228	0.0000210
FF100	0.0002463	0.0002370	−0.0000118	0.0000322
SP500	−0.0000228	−0.0000118	0.0002450	−0.0000368
SHCI	0.0000210	0.0000322	−0.0000368	0.0008837

Table 4: Optimal portfolios with models (18) and (22)

Method	FF25	FF100	SP500	SHCI
The proposed model (18)	0.06230	0.4674	0.3383	0.1320
The Mean-CVaR model (22)	0.0000	0.4737	0.3443	0.1820

Table 5: Mean return, variance, VaR, and CVaR obtained with models (18) and (22)

Method	Mean return	Variance	VaR	CVaR
The proposed model (18)	0.0003118	0.0001073	0.0169	0.0211
The Mean-CVaR model (22)	0.0003228	0.0001185	0.0170	0.0212

Table 6: Mean return, variance, VaR, and CVaR obtained with models (18) and (22) of Monte Carlo simulations

Method	Mean return	Variance	VaR	CVaR
The proposed model (18)	0.0003599	0.0001083	0.01694	0.02107
The Mean-CVaR model (22)	0.0003641	0.0001196	0.01698	0.02124

5.2.2 Monte Carlo simulation analysis

In the example, the discrete sample space of random returns consists of 10,000 samples are generated via the Monte Carlo approach by assuming the same joint normal distribution. The expected values and the covariance matrix of returns of these four assets are the same as in Tables 2 and 3, respectively. Then, mean return, variance, VaR, and CVaR of the samples can be calculated in the same way. The corresponding results of the Monte Carlo simulation are shown in Table 6.

Upon comparing the results in Table 6 to those that correspond to the optimal portfolio in Tables 5, we see that the CVaR values differ by only a few percentage points, depending upon the number of samples, and likewise for the VaR values. In addition, the behaviours of mean return and variance in Monte Carlo simulation are also similar to those in Table 5.

6 Conclusion

In this article, motivated by distributionally robust optimization models, we propose a new distributionally robust ratio model which concerns both CVaR and SD as risk measures. By introducing measurement assumptions of reward, CVaR, and SD, the ratio model can be reduced to a distributionally robust model without ratios. Given the first and second moments of distribution, it can be equivalent to a semidefinite programming problem by the duality theory. Compared with the existed distributionally robust models, the new model not only considers CVaR, but also involves SD in measuring the risk. It should give an allocation without too large volatility. For investors, it would be a conservative case with reasonable rewards. However, like the other distributionally robust models, the proposed model is sensitive to different distribution forms of uncertainties as well. On the other hand, it indicates that distributionally robust models usually depend on the chosen distribution parameters. Numerical experiments validate that the new model is effective and reasonable by comparing with the robust model under ellipsoid uncertainty and the corresponding Mean-CVaR model.

Acknowledgement: We thank the anonymous referee for the careful reading and many helpful comments.

Funding information: This research was funded by Dalian Science and Technology Star Project (2021RQ064), the Fundamental Research Funds for the Central Universities (No. 3132024), and China Postdoctoral Science Foundation (No. 2020M670717).

Author contributions: Conceptualization of the article was given by Y.Z. and L.X.; methodology by Y.Z. and L.X.; writing–original draft preparation by Y.Z.; software and validation by L.X.; writing–review and editing by Y.Z. and L.X. All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Conflict of interest: The authors state no conflict of interest.

Data availability statement: All data generated or analysed during this study are included in this article.

References

[1] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Math. Oper. Res. **23** (1998), no. 4, 769–805, DOI: https://doi.org/10.1287/moor.23.4.769.

[2] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton University Press, New Jersey, United States, 2009.

[3] D. Bertsimas, D. B. Brown, and C. Caramanis, *Theory and applications of robust optimization*, SIAM Rev. **53** (2011), no. 3, 464–501, DOI: https://doi.org/10.1137/080734510.

[4] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer, New York, 1997.

[5] A. Ruszczynski and A. Shapiro, *Stochastic programming models*, in: A. Ruszczynski, and A. Shapiro (Eds.), Handbooks in Operations Research and Management Science, Stochastic Programming, Vol. 10, 2003, pp. 1–64, DOI: https://doi.org/10.1016/S0927-0507(03)10001-1.

[6] D. Bertsimas and A. Thiele, *A robust optimization approach to inventory theory*, Oper. Res. **54** (2006), no. 1, 150–168, DOI: https://doi.org/10.1287/opre.1050.0238.

[7] H. E. Scarf, *A min-max solution of an inventory problem*, in: K. J. Arrow, S. Karlin, and H. E. Scarf (Eds.), Studies in the Mathematical Theory of Inventory and Production, Stanford University Press, California, 1958, pp. 201–209.

[8] K. Isii, *On sharpness of Tchebycheff-type inequalities*, Ann. Inst. Statist. Math. **14** (1962), no. 1, 185–197, DOI: https://doi.org/10.1007/BF02868641.

[9] A. Shapiro, *On duality theory of conic linear problems*, in: M. Á. Goberna, M. A. López (Eds.), Semi-Infinite Programming, Nonconvex Optimization and Its Applications, Vol. 57, Springer, Boston, MA, 2001, pp. 135–165, DOI: https://doi.org/10.1007/978-1-4757-3403-4_7.

[10] D. Bertsimas and I. Popescu, *Optimal inequalities in probability theory: A convex optimization approach*, SIAM J. Optim. **15** (2004), no. 3, 780–804, DOI: https://doi.org/10.1137/S1052623401399903.

[11] G. C. CalaLore and L. El Ghaoui, *On distributionally robust chance-constrained linear programs*, J. Optim. Theory Appl. **130** (2006), no. 1, 1–22, DOI: https://doi.org/10.1007/s10957-006-9084-x.

[12] W. Chen, M. Sim, J. Sun, and C.-P. Teo, *From CVaR to uncertainty set: Implications in joint chance constrained optimization*, Oper. Res. **58** (2009), no. 2, 470–485, DOI: https://doi.org/10.1287/opre.1090.0712.

[13] S.-S. Cheung, A. Man-Cho So, and K. Wang, *Linear matrix inequalities with stochastically dependent perturbations and applications to chance-constrained semidefinite optimization*, SIAM J. Optim. **22** (2012), no. 4, 1394–1430, DOI: https://doi.org/10.1137/110822906.

[14] S. Zymler, D. Kuhn, and B. Rustem, *Distributionally robust joint chance constraints with second-order moment information*, Math. Program. **137** (2013), 167–198, DOI: https://doi.org/10.1007/s10107-011-0494-7.

[15] J. Goh and M. Sim, *Distributionally robust optimization and its tractable approximations*, Oper. Res. **58** (2010), no. 4, 902–917, DOI: https://doi.org/10.1287/opre.1090.0795.

[16] W. Wiesemann, D. Kuhn, and M. Sim, *Distributionally robust convex optimization*, Oper. Res. **62** (2014), no. 6, 1358–1376, DOI: https://doi.org/10.1287/opre.2014.1314.

[17] L. El Ghaoui, M. Oks, and F. Oustry, *Worst-case value-at-risk and robust portfolio optimization: a conic programming approach*, Oper. Res. **51** (2003), no. 4, 543–556, DOI: https://doi.org/10.1287/opre.51.4.543.16101.

[18] E. Delage and Y. Ye, *Distributionally robust optimization under moment uncertainty with application to data-driven problems*, Oper. Res. **58** (2010), no. 3, 595–612, DOI: https://doi.org/10.1287/opre.1090.0741.

[19] Y. Zhao, Y. Liu, J. Zhang, and X. Yang, *Distributionally robust reward-risk ratio programming with Wasserstein metric*, Pac. J. Optim. **15** (2019), no. 1, 69–90, DOI: http://manu71.magtech.com.cn/Jwk3_pjo/EN/Y2019/V15/I1/69.

[20] R. Ji, M. A. Lejeune, and Z. Fan, *Distributionally robust portfolio optimization with linearized STARR performance measure*, Quant. Finance **22** (2022), no. 1, 113–127, DOI: https://doi.org/10.1080/14697688.2021.1993623.

[21] X. Tong, X. Luo, H. Yang, and L. Zhang, *A distributionally robust optimization-based risk-limiting dispatch in power system under moment uncertainty*, Int. Trans. Electr. Energ. Syst. **27** (2017), no. 8, e2343, DOI: https://doi.org/10.1002/etep.2343.

[22] Y. Liu, R. Meskarian, and H. Xu, *Distributionally robust reward-risk ratio optimization with moment constraints*, SIAM J. Optim. **27** (2017), no. 2, 957–985, DOI: https://doi.org/10.1137/16M106114X.

[23] R. Sehgal and A. Mehra, *Robust reward-risk ratio portfolio optimization*, Int. Trans. Oper. Res. **28** (2021), no. 4, 2169–2190, DOI: https://doi.org/10.1111/itor.12652.

[24] S. Guo and H. Xu, *Distributionally robust shortfall risk optimization model and its approximation*, Math. Program. **174** (2019), 473–498, DOI: https://doi.org/10.1007/s10107-018-1307-z.

[25] J. Gotoh, M. J. Kim, and A. E. B. Lim, *Calibration of distributionally robust empirical optimization models*, Oper. Res. **69** (2021), no. 5, 1630–1650, DOI: https://doi.org/10.1287/opre.2020.2041.

[26] R. Ji and M. A. Lejeune, *Data-driven optimization of reward-risk ratio measures*, INFORMS J. Comput. **33** (2020), no. 3, 1120–1137, DOI: https://doi.org/10.1287/ijoc.2020.1002.

[27] L. Xu, R. Shen, and Z. Meng, *Distributionally robust bilevel programming based on worst conditional value-at-risk*, Pac. J. Optim. **18** (2022), no. 4, 825–839, DOI: http://www.yokohamapublishers.jp/online-p/PJO/vol18/pjov18n4p825.pdf.

[28] L. Xu and Y. Zhou, *New robust reward-risk ratio models with CVaR and standard deviation*, J. Math. **2022** (2022), 8304411, DOI: https://doi.org/10.1155/2022/8304411.

[29] X. J. Tong, F. Wu, and L. Q. Qi, *Worst-case CVaR based portfolio optimization models with applications to scenario planning*, Optim. Methods Softw. **24** (2009), no. 6, 933–958, DOI: https://doi.org/10.1080/10556780902865942.

[30] S. S. Zhu and M. Fukushima, *Worst-case conditional value-at-risk with application to robust portfolio management*, Oper. Res. **57** (2009), no. 5, 1155–1168, DOI: https://doi.org/10.1287/opre.1080.0684.

[31] X. J. Tong and F. Wu, *Robust reward-risk ratio optimization with application in allocation of generation asset*, Optimization **63** (2014), no. 11, 1761–1779, DOI: https://doi.org/10.1080/02331934.2012.672419.

[32] H. M. Markowitz, *Portfolio selection*, J. Finance **7** (1952), no. 1, 77–91, DOI: https://doi.org/10.1111/j.1540-6261.1952.tb01525.x.

[33] D. Goldfarb and G. Iyengar, *Robust portfolio selection problems*, Math. Oper. Res. **28** (2003), no. 1, 1–38, DOI: https://doi.org/10.1287/moor.28.1.1.14260.

[34] S. J. Kim and S. Boyd, *Robust efficient frontier analysis with a separable uncertainty model*, Technical Report, Stanford University, 2007, https://optimization-online.org/?p=10294.

[35] X. Lin, *Researches on the efficient frontier of mean-CVaR under normal distribution condition*, 2009 International Conference on Management and Service Science, Beijing, China, 2009, pp. 1–4, DOI: https://doi.org/10.1109/ICMSS.2009.5305151.

[36] S. V. Stoyanov, S. T. Rachev, and F. J. Fabozzi, *Optimal financial portfolios*, Appl. Math. Finance **14** (2007), no. 5, 401–436, DOI: https://doi.org/10.1080/13504860701255292.

[37] K. Isii, *The extrema of probability determined by generalized moments (I) bounded random variables*, Ann. Inst. Stat. Math. **12** (1960), no. 2, 119–134.

[38] A. Shapiro and A. Kleywegt, *Minimax analysis of stochastic problems*, Optim. Methods Softw. **17** (2002), no. 3, 523–542, DOI: https://doi.org/10.1080/1055678021000034008.

[39] K. Natarajan, D. Pachamanova, and M. Sim, *Constructing risk measures from uncertainty sets*, Oper. Res. **57** (2009), no. 5, 1129–1141, DOI: https://doi.org/10.1287/opre.1080.0683.

[40] Y. Zhou, L. Yang, L. Xu, and B. Yu, *Inseparable robust reward risk optimization models with distribution uncertainty*, Japan J. Indust. Appl. Math. **33** (2016), 767–780, DOI: https://doi.org/10.1007/s13160-016-0230-z.

[41] O. Ledoit and M. Wolf, *A well-conditioned estimator for large-dimensional covariance matrices*, J. Multivariate Anal. **88** (2004), no. 2, 365–411, DOI: https://doi.org/10.1016/S0047-259X(03)00096-4.

[42] Z. Dai and F. Wen, *Some improved sparse and stable portfolio optimization problems*, Finance Res. Lett. **27** (2018), 46–52, DOI: https://doi.org/10.1016/j.frl.2018.02.026.