Data-driven robust mean-CVaR portfolio selection under distribution ambiguity

ZHILIN KANG†‡, XUN LI§, ZHONGFEI LI*¶ and SHUSHANG ZHU¶

†School of Mathematical Sciences, Huaqiao University, Fujian 362021, P.R. China
‡School of Mathematics, Sun Yat-sen University, Guangzhou 510275, P.R. China
§Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, P.R. China
¶Sun Yat-sen Business School, Sun Yat-sen University, Guangzhou, 510275, P.R. China

(*Received 00 Month 20XX; in final form 00 Month 20XX*)

In this paper, we present a computationally tractable optimization method for a robust mean-CVaR portfolio selection model under the condition of distribution ambiguity. We develop an extension that allows the model to capture a zero net adjustment via the linear constraint in the mean return, which can be cast as a tractable conic program. Also, we adopt a nonparametric bootstrap approach to calibrate the levels of ambiguity and show that the portfolio strategies are relatively immune to variations in input values. Finally, we show that the resulting robust portfolio is very well diversified and superior to its non-robust counterpart in terms of portfolio stability, expected returns and turnover. The results of numerical experiments with simulated and real market data shed light on the behavior of our distributionally robust optimization model established.

Keywords: Portfolio selection; Distributionally robust optimization; Zero net adjustment; Bootstrap; Conic programs

JEL Classification: G11, C14, C52, C61, C63

1. Introduction

Quantile-based risk measures are practically important in various fields such as financial engineering, financial management and economics. Recently, there has been a dramatic increase in the interest of studying this family of risk measures and its financial applications. Value-at-risk (VaR) and conditional value-at-risk (CVaR), which are concerned with the probability or magnitude of losses, are the most popular quantile-based risk measures. VaR reflects the maximum potential loss of an asset or portfolio in a given period and at a confidence level, i.e., it provides information on losses that cannot be exceeded with a certain probability. As a measure of risk, however, VaR has several limitations. For example, it does not satisfy subadditivity, i.e., diversification may result in greater risk, hence it is not a coherent risk measure (Artzner *et al.* 1999). To overcome the drawbacks of VaR, researchers propose a modified version, namely, CVaR, which is defined as mean of its tail distribution exceeding VaR. CVaR, which is a coherent risk measure, is more appealing than VaR because it takes into account the contribution from the very rare but very large losses.

To the best of our knowledge, most of observations and discussions of CVaR in the portfolio optimization literature are concerned with the formulation and tractability of model, and full knowledge

*Corresponding author. Email: lnslzf@mail.sysu.edu.cn. Tel.: +86 20 84111989; fax: +86 20 84114823.

of the distribution of portfolio losses (Hellmich and Kassberger 2011, Xiao and Valdez 2015) is assumed. In most real-life applications, however, the distribution of those losses is unknown, and therefore it is either estimated from historical data or constructed from expert knowledge (or a combination of the two); see, e.g., Yao *et al.* (2013). Moreover, even when the loss distribution is precisely known, the computation of CVaR typically reduces to the evaluation of a high-dimensional integral, which is itself an important and challenging problem. Furthermore, in the words of Scarf *et al.* (1958), "we may have reason to suspect that the future demand will come from a distribution which differs from that governing past history in an unpredictable way". Such unpredictability provides a strong incentive for the decision-maker to adopt (distributionally) robust optimization approaches, which have recently become increasingly popular in portfolio optimization under ambiguity. [1] An attempt to the robust optimization of CVaR is proposed in Quaranta and Zaffaroni (2008), who implement in a robust manner of the bi-criteria model proposed by Rockafellar and Uryasev (2000), thereby obtaining a robust linear reformulation of the problem. The idea of optimal decisions in ambiguous stochastic models is initiated by Scarf *et al.* (1958), who deal with a distribution-robust inventory control problem in which only mean and variance of demand are known. This ambiguous distribution set generally takes the following form

$$\mathbb{D}_0 = \left\{ \mathrm{P} \in \mathcal{M}_+ : \mathrm{P}(\xi \in \Omega) = 1,\ \mathrm{E}_{\mathrm{P}}(\xi) = \hat{\mu},\ \mathrm{Cov}_{\mathrm{P}}(\xi) = \hat{\Sigma} \succ 0 \right\},$$

where $\mathcal{M}_+$ is the set of all probability measures on the measurable space $(\mathbb{R}^n, \mathfrak{B})$ with the σ-algebra $\mathfrak{B}$ on $\mathbb{R}^n$, $\Omega \subseteq \mathbb{R}^n$ is a closed convex set known to contain the support of random vector ξ, and $\hat{\mu}$ and $\hat{\Sigma}$ are the inferred mean and covariance matrix, respectively. Natarajan *et al.* (2009) present a model for worst-case CVaR based on partial moment information when the underlying distributions of random variables are not precisely known. Also, they discuss how those sets in robust optimization map to risk measures, and propose a specific approach to generate coherent risk measures in this context. Chen *et al.* (2011) consider the worst-case lower partial moments and worst-case CVaR with respect to reliable data consisting of only fixed first and second moments, and derive a tight bound for these two risk measures. The same type of ambiguity set is adopted by Calafiore and El Ghaoui (2006) and Popescu *et al.* (2007), who examine a linear chance-constrained problem and investigate the problem of maximizing a portfolio's expected utility.

However, the assumption that mean and covariance of the data are known exactly is less reasonable from a practical viewpoint. Natarajan *et al.* (2010) derive exact and approximate optimal strategies for the worst-case expected utility model in the portfolio selection problem under distribution ambiguity using a piecewise-linear concave utility function. Recently, Paç and Pinar (2014) consider the problem of optimal portfolio choice using CVaR risk measures in which the mean return is subject to an ellipsoidal uncertainty set and distribution ambiguity. Lotfi and Zenios (2016) develop models of robust VaR and CVaR optimization for joint ambiguity in distribution, mean return and covariance matrix. They present an algorithm and heuristic for constructing an ellipsoidal ambiguity set from a set of point estimates. In addition to moment-based robust approaches, several other facets of robust portfolio construction based on limited statistical information can be found in Zhu and Fukushima (2009), Huang *et al.* (2010), Hellmich and Kassberger (2011), Gotoh *et al.* (2013), Wozabal (2014) and Zhu *et al.* (2015). We do not limit the discussion in our paper to

[1] In general, decision-makers are exposed not only to risk (refers to events for which the probabilities of the future outcomes are known) but also to ambiguity (refers to events for which the probabilities of the future outcomes are unknown) when making investment decisions. The distinction between risk and ambiguity was first made by Knight (1921) and latter supported by the empirical experiments of Ellsberg (1961), whose findings have shown that agents are not always able to derive a unique probability distribution over the reference state space. After Ellsberg's seminal paper, uncertain environment has become better known as ambiguity and the general dislike for it as ambiguity aversion. Ellsberg (1961) argues that most people are ambiguity-averse, that is, they prefer a lottery with known probabilities to a similar lottery with unknown probabilities. In recent years, a rapidly growing literature on ambiguity aversion is emerging; see, among others, Garlappi *et al.* (2007) and Ma *et al.* (2008) for optimal portfolio choice, Cao *et al.* (2005) and Ui (2011) for non-participation or selective participation in markets, and Faria and Correia-da-Silva (2014) for European call option pricing.

the minimization of a risk as in Natarajan *et al.* (2009), Chen *et al.* (2011), Gotoh *et al.* (2013) and Wozabal (2014), where only CVaR minimization is examined. By adding an ambiguous minimum return constraint, we introduce robust portfolio optimization using worst-case CVaR. For details of the historical development and use of robust optimization in portfolio selection, we refer readers to Kim *et al.* (2014), and references therein.

Data-driven approaches have been proposed, in which investors have no information beyond the evolving history of asset return observations. Sample average approximation (see, e.g., Bertsimas et al. 2016) is a popular approach to data-driven decision making under uncertainty. Delage and Ye (2010) novelly construct a confidence set for the mean vector and covariance matrix from historical data. Ben-Tal et al. (2013) study how to construct uncertainty sets from statistical confidence region based on phi-divergences, and generalize the result of Wang et al. (2016). Jiang and Guan (2016) propose a data-driven approach to chance-constrainted problems where the ambiguity of the return distribution is described by phi-divergence. Bertsimas et al. (2017) study a distributionally robust optimization with various choices of distribution sets. In particular, they focus on data analysis and hypothesis-testing tools such as Kolmogorov-Smirnov, χ^2, Anderson-Darling, Watson and likelihood ratio to construct ambiguity sets. Under certain convexity and compactness assumptions, Esfahani and Kuhn (2017) show that data-driven distributionally robust problems with Wasserstein distance are tractable, by transforming the inner maximization problem into a finite dimensional problem. In these existing methods, researchers assume that the ambiguity sets are constructed with certain structures and sizes according to a finite number of data points.

Our work is somewhat related to the paper by Salahi *et al.* (2013). Both are extensions of the traditional mean-CVaR model which uses a pre-given single probability distribution. Salahi *et al.* (2013) only introduce the uncertainty of mean and use the CVaR of mean to replace the mean, where the distribution is the pre-given single and the CVaR is single. While, we incorporate the uncertainty of mean, covariance and distribution, use a set of distributions to replace the pre-given single distribution, and take the worst case on the set. The worst-case approach adopted in our paper is fitter for the spirit of the max-min expected utility (MEU) framework axiomatized by Gilboa and Schmeidler (1989), which can be viewed as reflecting both beliefs (information) and subjective attitude toward ambiguity.[1]

To facilitate the practice, we adopt an ambiguous distribution of the returns where only known mean vector and covariance matrix belong to some ambiguity sets. This is a more simple, clear and flexible way to start real applications. At the same time, we can easily achieve a tractable reformulation of the distributionally robust portfolio optimization problem as a second-order cone program (SOCP), which can be efficiently solved in polynomial time. Furthermore, we integrate the zero net adjustment on mean returns to overcome the drawback of excessive conservativeness in robust portfolio optimization. While there is a large body of existing literature concerned with the formulation of robust portfolio selection, the discussion focusing on how to specify the size of the ambiguity set in an attractive manner is relatively rare. We also focus on providing guidelines for calibrating two levels of ambiguity from available realization of the uncertain data in robust CVaR optimization application, and by which investors can specify the suitable bounds of the ambiguity set. Computational experiments on real market data reveal that, in most cases, our approach

[1] Unlike classical approaches to decision making such as von Neumann-Morgenstern paradigm of expected utility maximization that neglect an agent's preference on the choice among multiple probability models, Gilboa and Schmeidler (1989) provide a system of axioms under which an agent's preference on the choice of the models can be characterized by the worst-case approach. However, this approach does not distinguish between ambiguity and aversion to ambiguity, and hence is sometimes criticized because it apparently implies extreme ambiguity aversion. A few studies overcoming this issue have been proposed. For example, Klibanoff *et al.* (2005) provide an axiomatic foundation for the smooth ambiguity model. This model allows us to separate ambiguity from ambiguity attitudes and allows for smooth indifference curves, avoiding the infinite ambiguity aversion implied in the MEU approach. A recent paper by Izhakian (2017) provides an axiomatic foundation for a model of decision making under ambiguity, where the preference representation is referred to as sign-dependent expected utility under uncertain probabilities (EUUP). However, there is still a debate in the literature about the axiomatic foundations of this line of models (see Epstein 2010, Klibanoff *et al.* 2012). Because of this, the approach of Gilboa and Schmeidler (1989) is still to be the main reference in the literature.

outperforms the one without considering robustness and results in more efficient portfolios. Our work suggests that a data-driven approach to portfolio optimization, which calibrates two levels of ambiguity, is a valuable choice. This work further highlights the importance of taking moment uncertainty into account in portfolio selection. Moreover, an interesting finding in experiments is that model extension with zero net adjustment can truly avoid the excessive conservativeness in robust portfolio optimization and thus provide a more promising performance.

The remainder of the paper is organized as follows. Section 2 constructs a robust portfolio selection optimization model using CVaR as the risk measure. In this section, we derive the tractable conic reformulations for the robust mean-CVaR model, and consider a model extension with zero net adjustment on mean vector. Data-driven technique for calibrating the parameters and simulation analysis of the proposed approaches are derived in Section 3. In Section 4, we report the results of a variety of numerical experiments using real market data. Finally, conclusions are drawn in Section 5.

Notation. The space of the symmetric matrices of dimension n is denoted by S^n. S^n_+ denotes the cone of positive semidefinite matrices. Let I_n and e be the identity matrix of dimension n and the vector in $\mathbb{R}^n$ with unit elements, that is, $e = (1, \cdots, 1)^{\mathrm{T}}$, respectively. 'T' denotes the transpose of a matrix or vector and $||z||_2$ represents the Euclidean norm of the vector z. For any two matrices $A, B \in S^n$, we let $\langle A, B\rangle = tr(AB)$ be the trace scalar product, whereas relation $A \succeq B (A \succ B)$ implies that $A - B$ is positive semidefinite (positive definite). For $X \in S^n_+$, we denote its symmetric square root by $\sqrt{X}$ or $X^{\frac{1}{2}}$. For a given reversible matrix $G \in S^n$, $\|z\|_G$ represents the ellipsoidal norm of a vector z, i.e., $\|z\|_G = \sqrt{z^{\mathrm{T}} G^{-1} z}$. For a matrix $X \in S^n$, $||X||_F$ represents the Frobenius norm, i.e., $||X||_F = (X \bullet X)^{\frac{1}{2}} = \sqrt{tr(XX^{\mathrm{T}})}$.

2. Robust Mean-CVaR Portfolio Optimization Model

This section introduces the portfolio optimization problem, which accounts for ambiguity in the underlying probability distribution. We start from the traditional mean-CVaR portfolio selection model to its tractable distributionally robust counterpart.

2.1. *Worst-case CVaR risk measure*

Let $l(x, \xi)$ be the loss for portfolio vector x, and assume that the random return vector ξ has a continuous probability density function $p(\xi)$ with a finite mean μ and covariance Σ. Given a confidence level of $\beta \in (0, 1)$ and a fixed x, VaR is defined as

$$\mathrm{VaR}_\beta(x) = \min\left\{\eta \in \mathbb{R} : \int_{\{\xi : l(x,\xi) \le \eta\}} p(\xi) d\xi \ge \beta\right\}.$$

The corresponding CVaR at level $1 - \beta$ with respect to the distribution P, which is defined as the expected value of the loss $l(x, \xi)$ exceeding VaR, can be expressed as

$$\begin{aligned}\mathrm{CVaR}_\beta(x, \mathrm{P}) &= \mathrm{E}_{\mathrm{P}}\left[l(x,\xi) | l(x,\xi) \ge \mathrm{VaR}_\beta(x)\right] \\ &= \frac{1}{1-\beta} \int_{\{\xi : l(x,\xi) \ge \mathrm{VaR}_\beta(x)\}} l(x,\xi) p(\xi) d\xi.\end{aligned}$$

Rockafellar and Uryasev (2000) show that the calculation of CVaR can be achieved by minimizing the following function

$$F_\beta(x,\eta) = \eta + \frac{1}{1-\beta}\int_{\xi\in\mathbb{R}^n}[l(x,\xi)-\eta]^+ p(\xi)d\xi,$$

where $[t]^+ = \max\{0,t\}$. That is,

$$\mathrm{CVaR}_\beta(x,\mathrm{P}) = \min_{\eta\in\mathbb{R}} F_\beta(x,\eta). \tag{1}$$

Since risk measures such as VaR and CVaR are defined as functionals of the loss distribution, an implicit starting point is knowledge of that distribution. However, it is difficult to obtain the exact result of $F_\beta(x,\eta)$ if we have no information on the distribution of ξ. Using sampling or simulation methods, the approximation of $F_\beta(x,\eta)$ can be given as

$$\tilde{F}_\beta(x,\eta) = \eta + \frac{1}{S(1-\beta)}\sum_{k=1}^{S}[l(x,\xi^{[k]})-\eta]^+,$$

where S denotes the number of samples, and $\xi^{[k]}$ refers to the k-th sample. According to the law of large numbers in statistics, the empirical mean $\tilde{F}_\beta(x,\eta)$ converges to $F_\beta(x,\eta)$ as the sample size S goes to infinity.

In practice, owing to limited historical data, the estimated CVaR may contain considerable estimation error. Therefore, it may be more difficult for the investor to hedge the worst-case scenario over the set of probability measures, which is defined by the limited information available. The following definition gives the worst-case CVaR risk measure with distribution ambiguity.

Definition 1 Given a probability threshold $\beta > 0$, the worst-case CVaR (WCVaR) of portfolio x, where random vector ξ may assume a distribution from ambiguity set $\mathbb{D}$, is defined by

$$\mathrm{WCVaR}_\beta(x) = \sup_{\mathrm{P}\in\mathbb{D}} \mathrm{CVaR}_\beta(x,\mathrm{P}). \tag{2}$$

Zhu and Fukushima (2009) demonstrate that WCVaR inherits subadditivity, positive homogeneity, monotonicity, and translation invariance. Therefore, WCVaR, like CVaR, is a coherent risk measure.

2.2. *Robust portfolio optimization model with WCVaR risk measure*

Consider a financial market consisting of n different assets. A portfolio is characterized by a vector of asset weights $x \in \mathbb{R}^n$, whose elements add up to 1. The component x_i denotes the percentage of total wealth that is invested in the ith asset at the beginning of the investment period. The classical mean-CVaR portfolio selection problem (MC), which seeks for an optimal trade-off between risk and return, can be formulated as

$$\begin{aligned}(\mathrm{MC}):\ &\min_x\ \mathrm{CVaR}_\beta(x,\mathrm{P}),\\ &\text{s.t.}\ \ \mathrm{E}_\mathrm{P}(\xi)^\mathrm{T} x \geq \rho,\ x\in\mathcal{X},\end{aligned}$$

where ρ stands for the lower limit on the target expected return. Here, the objective is to minimize CVaR under the condition that the expected return is greater than or equal to ρ, and $\mathcal{X} \subseteq \mathbb{R}^n$ denotes the set of admissible portfolios.

In reality, the investor does not know the distribution P. When only a collection of S historical observations $\{\xi^{[1]}, \cdots, \xi^{[S]}\}$ of ξ is available, the original problem (MC) of finding a portfolio allocation can be recast as the following sample-based mean-CVaR optimization problem (SMC)

$$
\begin{aligned}
(\text{SMC}):\ \min_{(x,\eta)} \quad & \tilde{F}_\beta(x,\eta),\\
\text{s.t.} \quad & \frac{1}{S}\sum_{k=1}^{S}(\xi^{[k]})^T x \geq \rho,\\
& x \in \mathcal{X},\ \ \eta \in \mathbb{R}.
\end{aligned}
$$

SMC is easy to solve if $\mathcal{X}$ is convex and $l(x,\xi)$ is convex in x (Rockafellar and Uryasev 2000). It is easy to see that $\tilde{F}_\beta(x,\eta)$ depends on only a small portion of loss scenarios $l(x,\xi^{[1]}), \cdots, l(x,\xi^{[\tau]})$ with $\tau = \lceil(1-\beta)S\rceil$, where $l(x,\xi^{[i]})$ indicates the ith largest loss scenario, and $\lceil d\rceil$ is the smallest integer not less than d (Gotoh *et al.* 2013). This implication is that a perturbation in those loss scenarios can exert an influence on the estimate of $\text{CVaR}_\beta(x,\text{P})$, and simultaneously result in highly unreliable solutions. However, this framework can be validated by the law of large numbers. In other words, if the number of observations goes to infinity, the solution of SMC approaches the optimal portfolio.

Lim *et al.* (2011) show that CVaR is sensitive to the misspecification of the underlying loss distribution, and they also demonstrate empirically the fragility associated with CVaR minimization. A remedy for such fragility is to adopt a distributionally robust approach and embrace the fact that P is known to belong to an ambiguity set $\mathbb{D}$. Mathematically, the distributionally robust counterpart to MC is given by [1]

$$
\begin{aligned}
(\text{RMC}):\ \min_{x} \quad & \sup_{\text{P}\in\mathbb{D}} \text{CVaR}_\beta(x,\text{P}),\\
\text{s.t.} \quad & \inf_{\text{P}\in\mathbb{D}} \text{E}_\text{P}(\xi)^\text{T} x \geq \rho,\ \ x \in \mathcal{X},
\end{aligned}
$$

where $\mathbb{D}$ is the ambiguity set for uncertain underlying distribution P, and ρ denotes the required value of the worst-case expected return specified by the investor.

In RMC, we choose decision variable x in such a way that the CVaR risk measure is minimized under the worst-case distribution. Here, the worst-case distribution is taken over ambiguity set $\mathbb{D}$, that is, a family of distributions characterized through certain known properties of the unknown data-generating distribution P. Historical data and/or expert estimates serve as guiding tools in the estimation of P. Such characterizations of ambiguity should be convenient in terms of estimation and optimization.

Remark 1 Generally, $\mathcal{X}$ includes additional convex constraints on the portfolio structure, such as necessary finite budget, short-selling, diversification bound and cardinality constraints. Because the return on one portfolio can be expressed as $\xi^\text{T} x$, in what follows only a loss of the form $l(x,\xi) = -\xi^\text{T} x$ is considered.

[1] Ghirardato *et al.* (2004) axiomatize a model termed α-maxmin expected utility (α-MEU) wherein it is possible in a certain sense to distinguish ambiguity attitude from ambiguity. The α-MEU model is

$$
\alpha \inf_{\text{P}\in\mathbb{D}} \text{E}_\text{P}[U(X)] + (1-\alpha)\sup_{\text{P}\in\mathbb{D}} \text{E}_\text{P}[U(X)],
$$

where $\alpha \in [0,1]$ is a parameter, X is a random payoff, U is a general utility function, and $\mathbb{D}$ is a set of prior probability measures. A key feature of α-MEU is that it differentiates the level of ambiguity aversion, specified by α, and the level of ambiguity, specified by the range of $\mathbb{D}$. There is more flexibility with α-MEU in capturing the ambiguity attitude (parameterized by α) of the decision maker (α=1, 0 represents, respectively, extremely ambiguity-averse and extremely ambiguity-loving attitudes). Similar to the α-MEU criterion, we can develop a so-called α-maxmin mean-CVaR criterion. This paper just considers the extreme case of this criterion when $\alpha = 1$. The more general cases are left for our future research.

In this paper, we assume that ξ is governed by an ambiguous distribution that only partial information on the moments is known. More specifically, similar to Delage and Ye (2010), but partially different, we consider the ambiguity sets $\mathbb{D}_F(\gamma_1,\gamma_2)$ for the distribution defined as

$$\mathbb{D}_F(\gamma_1,\gamma_2)=\left\{ \mathrm{P}\in\mathcal{M}_+ : \begin{array}{c} \mathrm{P}(\xi\in\Omega)=1, \\ (\mathrm{E_P}(\xi)-\hat{\mu})^{\mathrm{T}}\hat{\Sigma}^{-1}(\mathrm{E_P}(\xi)-\hat{\mu})\le\gamma_1, \\ \|\mathrm{Cov_P}(\xi)-\hat{\Sigma}\|_F\le\gamma_2,\ \mathrm{Cov_P}(\xi)\succ 0 \end{array} \right\},$$

where we assume that $\Omega=\mathbb{R}^n$ holds in this paper. Parameters $\hat{\Sigma}$ and $\hat{\mu}$ can be estimated by the observations $\{\xi^{[i]}\}_{i=1}^S$ from ambiguous distribution P. If $\{\xi^{[i]}\}_{i=1}^S$ are independent and identically distributed (i.i.d.), an unbiased and consistent estimate of Σ is obtained by the sample covariance matrix $\hat{\Sigma}=\frac{1}{S-1}\sum\limits_{i=1}^{S}(\xi^{[i]}-\hat{\mu})(\xi^{[i]}-\hat{\mu})^{\mathrm{T}}$ with $\hat{\mu}=\frac{1}{S}\sum\limits_{i=1}^{S}\xi^{[i]}$. Parameters γ_1 and γ_2 determine the size of the ambiguity set and can be interpreted as a measure of the degree of ambiguity about the estimates of expected return and covariance. They provide means of quantifying one's confidence in $\hat{\mu}$ and $\hat{\Sigma}$, respectively. Hence, we refer to the positive parameters γ_1 and γ_2 as the levels of ambiguity throughout this paper.

The set $\mathbb{D}_F(\gamma_1,\gamma_2)$ can be seen as a generalization of the ambiguity sets considered in the literature. For instance, $\mathbb{D}_F(0,0)$ imposes an exact mean and covariance matrix, as in Chen *et al.* (2011), and $\mathbb{D}_F(\gamma_1,0)$ is related to the exact covariance matrix and ellipsoidal ambiguity set for the mean return considered in Paç and Pinar (2014). Similarly, $\mathbb{D}_F(0,\gamma_2)$ is related to the exact mean and F-norm ball set for the covariance matrix, which is of interest in risk minimization model or when there is special knowledge on the mean return. For practical computation in the subsequent experiments, one important question in solving RMC is how to choose the levels of ambiguity γ_1 and γ_2. If they are set too high, the optimization procedure may be overly conservative. The idea is to choose the ambiguity set that reflects the perceived information indicated by the data, which is particularly important in data-driven setting. In what follows we further investigate the issue of determination of γ_1 and γ_2 and provide a data-driven rule to guide investors with regard to the appropriate choice of them.

Remark 2 The ambiguity set $\mathbb{D}_F(\gamma_1,\gamma_2)$ is different form that of Delage and Ye (2010), where the uncertainty of covariance matrix is modeled as a conic set. $\mathbb{D}_F(\gamma_1,\gamma_2)$ offers a very intuitive way to calibrate the levels of ambiguity γ_1 and γ_2 endogenously (see Subsection 3.1), and provides a simple modeling guidance to practitioners.

In the sequel, we present a rigorous treatment for taking moment ambiguity defined by $\mathbb{D}_F(\gamma_1,\gamma_2)$ into account in the mean-CVaR portfolio optimization framework and develop tractable reformulations. And we extend our discussion to the case of zero net adjustment in Subsection 2.3. Before proceeding to the proofs of our theoretical results, we state a lemma that will be used in our proofs.

LEMMA 2.1 *(Chen et al. 2011). Assume that* $l(x,\xi)=-\xi^{\mathrm{T}}x$ *and random vector* $\xi\in\mathbb{R}^n$*, with mean* $\bar{\mu}$ *and covariance* $\bar{\Sigma}\succ 0$*, follows a family of distributions* $\mathcal{F}$*, which is defined by* $\mathcal{F}=\{\mathrm{P}\in\mathcal{M}_+ \mid \mathrm{P}(\xi\in\Omega)=1, \mathrm{E_P}(\xi)=\bar{\mu}, \mathrm{Cov_P}(\xi)=\bar{\Sigma}\}$*. If the support set of* ξ *covers the whole space, i.e.,* $\Omega=\mathbb{R}^n$*, then we have*

$$\max_{\mathrm{P}\in\mathcal{F}}\mathrm{CVaR}_\beta(x,\mathrm{P})=-\bar{\mu}^{\mathrm{T}}x+\kappa\sqrt{x^{\mathrm{T}}\bar{\Sigma}x}, \tag{3}$$

where $\kappa=\sqrt{\frac{\beta}{1-\beta}}$.

The details of the proof of Lemma 2.1 are referred to Chen *et al.* (2011).

Replacing $\mathbb{D}$ by $\mathbb{D}_F(\gamma_1,\gamma_2)$, we consider the following robust mean-CVaR model under distribution ambiguity,

$$
\begin{aligned}
(\text{RMC-}\mathbb{D}_F(\gamma_1,\gamma_2)):\ \ &\min_{x} \max_{\mathrm{P}\in\mathbb{D}_F(\gamma_1,\gamma_2)} \mathrm{CVaR}_\beta(x,\mathrm{P}),\\
&\text{s.t.} \min_{\mathrm{P}\in\mathbb{D}_F(\gamma_1,\gamma_2)} \mathrm{E}_{\mathrm{P}}(\xi)^{\mathrm{T}}x \ge \rho,\\
&\qquad x\in\mathcal{X}.
\end{aligned}
$$

The following proposition gives the main result in this subsection.

Proposition 2.2 *Suppose that $l(x,\xi) = -\xi^{\mathrm{T}}x$ and $\mathcal{X}$ constitutes a polyhedral set, then both the optimal value and a minimizer of the robust mean-CVaR model* (RMC-$\mathbb{D}_F(\gamma_1,\gamma_2)$) *under distribution ambiguity, can be obtained by solving the SOCP*

$$
\begin{aligned}
\min_{x,s,t}\ \ & \kappa s - \hat{\mu}^{\mathrm{T}}x + \sqrt{\gamma_1}t,\\
\text{s.t.}\ \ & \sqrt{\gamma_1}\|\hat{\Sigma}^{\frac{1}{2}}x\|_2 \le \hat{\mu}^{\mathrm{T}}x - \rho,\\
& \|(\hat{\Sigma}+\gamma_2 I_n)^{\frac{1}{2}}x\|_2 \le s,\\
& \|\hat{\Sigma}^{\frac{1}{2}}x\|_2 \le t,\\
& x\in\mathcal{X},
\end{aligned}
\tag{4}
$$

where the variables are $x\in\mathbb{R}^n, s,t\in\mathbb{R}$.

Proof. For convenience of analysis, we define

$$
\mathcal{U}_{(\hat{\mu},\hat{\Sigma})_F} = \left\{(\mu,\Sigma)\in\mathbb{R}^n\times S_+^n \mid (\mu-\hat{\mu})^{\mathrm{T}}\hat{\Sigma}^{-1}(\mu-\hat{\mu})\le\gamma_1, \|\Sigma-\hat{\Sigma}\|_F\le\gamma_2\right\}.
$$

We first maximize with respect to $\mathrm{P}\in\mathcal{F}$, and then with respect to $(\bar{\mu},\bar{\Sigma})$ within the uncertainty set. From the definition of $\mathbb{D}_F(\gamma_1,\gamma_2)$ and Lemma 2.1, we have

$$
\begin{aligned}
\max_{\mathrm{P}\in\mathbb{D}_F}\mathrm{CVaR}_\beta(x,\mathrm{P}) &= \max_{(\bar{\mu},\bar{\Sigma})\in\mathcal{U}_{(\hat{\mu},\hat{\Sigma})_F}}\ \max_{\mathrm{P}\in\mathcal{F}}\mathrm{CVaR}_\beta(x,\mathrm{P})\\
&= \max_{(\bar{\mu},\bar{\Sigma})\in\mathcal{U}_{(\hat{\mu},\hat{\Sigma})_F}} \{-x^{\mathrm{T}}\bar{\mu}+\kappa\sqrt{x^{\mathrm{T}}\bar{\Sigma}x}\}\\
&= -\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}} x^{\mathrm{T}}\bar{\mu} + \kappa\max_{\bar{\Sigma}\in\mathcal{U}_{\hat{\Sigma}}^F}\sqrt{x^{\mathrm{T}}\bar{\Sigma}x},
\end{aligned}
\tag{5}
$$

where $\mathcal{U}_{\hat{\mu}} = \{\mu\in\mathbb{R}^n \mid (\mu-\hat{\mu})^{\mathrm{T}}\hat{\Sigma}^{-1}(\mu-\hat{\mu})\le\gamma_1\}$ and $\mathcal{U}_{\hat{\Sigma}}^F = \{\Sigma\in S_+^n \mid \|\Sigma-\hat{\Sigma}\|_F\le\gamma_2\}$. Here, the last equality follows directly from the objective function and set of constraints being independent of each other.

Therefore, the two optimization problems in (5) can be solved independently. Clearly, the optimal solution to $\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}} x^{\mathrm{T}}\bar{\mu}$ can be shown to be

$$
\bar{\mu}_{wc} = \hat{\mu} - \frac{\sqrt{\gamma_1}\hat{\Sigma}x}{\sqrt{x^{\mathrm{T}}\hat{\Sigma}x}},
\tag{6}
$$

and the associated optimal value is

$$\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}} x^{\mathrm{T}}\bar{\mu} = \hat{\mu}^{\mathrm{T}}x - \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\hat{\Sigma}x}. \tag{7}$$

Let $\tilde{\Sigma} := \bar{\Sigma} - \hat{\Sigma}$. Then the problem $\max_{\bar{\Sigma}\in\mathcal{U}^F_{\hat{\Sigma}}} \sqrt{x^{\mathrm{T}}\bar{\Sigma}x}$ can be formalized as

$$\begin{aligned}
&\max_{\tilde{\Sigma}\in S^n} \sqrt{x^{\mathrm{T}}\tilde{\Sigma}x + x^{\mathrm{T}}\hat{\Sigma}x},\\
&\text{s.t.} \quad \|\tilde{\Sigma}\|_F \le \gamma_2.
\end{aligned}$$

For this problem, we have

$$\begin{aligned}
\max_{\tilde{\Sigma}:\|\tilde{\Sigma}\|_F\le\gamma_2} \sqrt{x^{\mathrm{T}}\tilde{\Sigma}x + x^{\mathrm{T}}\hat{\Sigma}x} &= \max_{\tilde{\Sigma}:\|\tilde{\Sigma}\|_F\le\gamma_2} \sqrt{\tilde{\Sigma}\bullet xx^{\mathrm{T}} + x^{\mathrm{T}}\hat{\Sigma}x}\\
&\le \sqrt{\gamma_2\|xx^{\mathrm{T}}\|_F + x^{\mathrm{T}}\hat{\Sigma}x}\\
&= \sqrt{\gamma_2\|x\|_2^2 + x^{\mathrm{T}}\hat{\Sigma}x}
\end{aligned}$$

from the Cauchy-Schwartz inequality, and the optimal solution for this auxiliary problem is given by

$$\tilde{\Sigma}^* = \frac{\gamma_2 xx^{\mathrm{T}}}{\|x\|_2^2}.$$

Hence,

$$\max_{\bar{\Sigma}\in\mathcal{U}^F_{\hat{\Sigma}}} \sqrt{x^{\mathrm{T}}\bar{\Sigma}x} = \sqrt{x^{\mathrm{T}}(\hat{\Sigma}+\gamma_2 I_n)x}. \tag{8}$$

Plugging (7) and (8) into (5), the objective function of problem (RMC-$\mathbb{D}_F(\gamma_1,\gamma_2)$) becomes

$$\max_{\mathrm{P}\in\mathbb{D}_F(\gamma_1,\gamma_2)} \mathrm{CVaR}_\beta(x,\mathrm{P}) = -\hat{\mu}^{\mathrm{T}}x + \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\hat{\Sigma}x} + \kappa\sqrt{x^{\mathrm{T}}(\hat{\Sigma}+\gamma_2 I_n)x}.$$

Hence, we have

$$\begin{aligned}
\min_x \quad & -\hat{\mu}^{\mathrm{T}}x + \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\hat{\Sigma}x} + \kappa\sqrt{x^{\mathrm{T}}(\hat{\Sigma}+\gamma_2 I_n)x},\\
\text{s.t.} \quad & \hat{\mu}^{\mathrm{T}}x - \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\hat{\Sigma}x} \ge \rho,\\
& x\in\mathcal{X}.
\end{aligned}$$

By introducing auxiliary variables $s, t \in \mathbb{R}$, we can rewrite this optimization problem as

$$\begin{aligned} \min_{x,s,t} \quad & -\hat{\mu}^{\mathrm{T}}x + \sqrt{\gamma_1}t + \kappa s, \\ \text{s.t.} \quad & \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\hat{\Sigma}x} \leq \hat{\mu}^{\mathrm{T}}x - \rho, \\ & \sqrt{x^{\mathrm{T}}(\hat{\Sigma} + \gamma_2 I_n)x} \leq s, \\ & \sqrt{x^{\mathrm{T}}\hat{\Sigma}x} \leq t, \quad x \in \mathcal{X}, \end{aligned}$$

which yields the desired result. □

Remark 3 We see that problem (4) is an SOCP, which is important for computational tractability and can be solved efficiently by interior-point method. Readers are referred to Alizadeh and Goldfarb (2003) for further details of SOCP problems.

Remark 4 Robust CVaR optimization problems with different types of distributional uncertainty are available for several cases. Delage and Ye (2010) consider the case with ambiguity on the first two moments. Bertsimas *et al.* (2017) focus on using data and hypothesis-testing tools to construct the distributional sets. Esfahani and Kuhn (2017) show that under certain convexity and compactness assumptions, data-driven distributionally robust problems with Wasserstein distance is tractable. One of the differences between our model and the model in Esfahani and Kuhn (2017) is the measure of ambiguity. Esfahani and Kuhn (2017) adopt the Wasserstein metric, which is a mathematically sophisticated concept of a distance among probability distributions and is usually hard to compute. In contrast, our measures of ambiguity are defined by certain constraints on distribution's moments, which are widely used in practice and easy to compute. This particular choice of partial information is the key behind the characteristics of our model. It appears that tractable conic reformulations programming are easier to derive for distributionally robust optimization models with moment-based ambiguity sets (see Esfahani and Kuhn (2017)). Moreover, our proof approach is motivated by the insight gained from the existing result of Chen *et al.* (2011) and hence is more direct in the sense that we do not resort to tools from infinite-dimensional convex optimization as in the proofs of Delage and Ye (2010) and Esfahani and Kuhn (2017).

2.3. *Extension with zero net adjustment*

The robust technique is a relatively conservative method. Ceria and Stubbs (2006) introduce the zero net adjustment-robust framework to reduce the conservativeness of robust mean-variance s-trategies for the input parameter under ellipsoidal uncertainty. They show that efficient frontier generated by robust portfolios to be closer to the true frontier, and realized returns of robust portfolios are greater to exhibit better out-of-sample performance than those using traditional mean-variance optimization. If the stock returns are serially independent and identically distributed, we can invoke the Central Limit Theorem to conclude that the sample mean $\hat{\mu}$ is approximately normally distributed. Hence, the estimated expected returns $\hat{\mu}$ are symmetrically distributed around $\mathrm{E}_{\mathrm{P}}(\xi)$, we expect them to approximate as many realized returns above their expected values as below them in Santos (2010). This amounts to saying that the estimation errors cancel out when summed over all assets. Mathematically, it can be expressed as

$$e^{\mathrm{T}}(\mathrm{E}_{\mathrm{P}}(\xi) - \hat{\mu}) = 0, \tag{9}$$

which forces the net adjustment of expected returns to be zero. Moreover, Santos (2010), in comparing the min-max robust and adjusted robust approaches with the Markowitz's mean-variance and min-variance approaches, provides empirical evidence to show that robust optimization is an effective way to treat the problem of estimation error in mean return. Interested readers are also referred to Zymler *et al.* (2011) and Perchet *et al.* (2015) for related topics on zero net adjustment.

Inspired by the work of Ceria and Stubbs (2006), we extend our model (RMC-$\mathbb{D}_F(\gamma_1, \gamma_2)$) to the case of zero net adjustment, which we call "the adjusted robust mean-CVaR optimization model under distribution ambiguity":

$$
\begin{aligned}
(\text{RMC-}\mathbb{D}_{\text{F}}^{\text{adj}}(\gamma_1,\gamma_2)):\ \min_x \quad & \max_{\text{P}\in\mathbb{D}_F^{adj}(\gamma_1,\gamma_2)} \text{CVaR}_\beta(x,\text{P}),\\
\text{s.t.} \quad & \min_{\text{P}\in\mathbb{D}_F^{adj}(\gamma_1,\gamma_2)} \text{E}_{\text{P}}(\xi)^{\text{T}}x \geq \rho,\\
& x \in \mathcal{X},
\end{aligned}
$$

where

$$
\mathbb{D}_F^{adj}(\gamma_1,\gamma_2) = \left\{ \text{P} \in \mathcal{M}_+ : \begin{array}{l} \text{P}(\xi\in\Omega) = 1,\\ (\text{E}_{\text{P}(\xi)} - \hat{\mu})^{\text{T}}\hat{\Sigma}^{-1}(\text{E}_{\text{P}}(\xi) - \hat{\mu}) \leq \gamma_1,\\ \|\text{Cov}_{\text{P}}(\xi) - \hat{\Sigma}\|_F \leq \gamma_2,\ \text{Cov}_{\text{P}}(\xi) \succ 0,\\ e^{\text{T}}(\text{E}_{\text{P}}(\xi) - \hat{\mu}) = 0 \end{array} \right\}.
$$

We follow a similar line of argument for (RMC-$\mathbb{D}_F(\gamma_1, \gamma_2)$), which is developed in Subsection 2.2, and provide an equivalent robust formulation that relies on strong duality to solve the robust problem using standard cone programming solvers. The following result indicates that the problem (RMC-$\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$) over the family $\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$ can also be solved as an SOCP.

PROPOSITION 2.3 *Suppose that $l(x, \xi) = -\xi^{\text{T}}x$ and $\mathcal{X}$ constitutes a polyhedral set, then the adjusted robust mean-CVaR model under distribution ambiguity (*RMC-$\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$*) is equivalent to an SOCP:*

$$
\begin{aligned}
\min_{x,\delta,\omega} \quad & \kappa\delta - \hat{\mu}^{\text{T}}x + \sqrt{\gamma_1}\omega,\\
\text{s.t.} \quad & \sqrt{\gamma_1}\|\Lambda^{\frac{1}{2}}x\|_2 \leq \hat{\mu}^{\text{T}}x - \rho,\\
& \|(\hat{\Sigma} + \gamma_2 I_n)^{\frac{1}{2}}x\|_2 \leq \delta,\\
& \|\Lambda^{\frac{1}{2}}x\|_2 \leq \omega,\\
& x \in \mathcal{X},
\end{aligned}
\tag{10}
$$

where the variables are $x \in \mathbb{R}^n, \delta, \omega \in \mathbb{R}$ and $\Lambda = \hat{\Sigma} - \frac{1}{e^{\text{T}}\hat{\Sigma}e}\hat{\Sigma}ee^{\text{T}}\hat{\Sigma} \succ 0$.

Proof. Define a set as

$$
\mathcal{U}_{\hat{\mu}}^{adj} = \left\{ \mu \in \mathbb{R}^n \mid (\mu - \hat{\mu})^{\text{T}}\hat{\Sigma}^{-1}(\mu - \hat{\mu}) \leq \gamma_1, e^{\text{T}}(\mu - \hat{\mu}) = 0 \right\}.
$$

Applying Lemma 2.1 yields

$$\begin{aligned}
\max_{\mathrm{P}\in\mathbb{D}_F^{adj}(\gamma_1,\gamma_2)} \mathrm{CVaR}_\beta(x,\mathrm{P}) &= \max_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}^{adj},\bar{\Sigma}\in\mathcal{U}_{\hat{\Sigma}}^F} \max_{\mathrm{P}\in\mathcal{F}} \mathrm{CVaR}_\beta(x,\mathrm{P}) \\
&= \max_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}^{adj},\bar{\Sigma}\in\mathcal{U}_{\hat{\Sigma}}^F} \left\{-x^{\mathrm{T}}\bar{\mu} + \kappa\sqrt{x^{\mathrm{T}}\bar{\Sigma}x}\right\} \\
&= -\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}^{adj}} x^{\mathrm{T}}\bar{\mu} + \kappa \max_{\bar{\Sigma}\in\mathcal{U}_{\hat{\Sigma}}^F} \sqrt{x^{\mathrm{T}}\bar{\Sigma}x}. \qquad (11)
\end{aligned}$$

The two optimization problems in (11) can be solved independently. Clearly, the problem $\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}^{adj}} x^{\mathrm{T}}\bar{\mu}$ can be re-expressed as

$$\begin{aligned}
&\min_{\bar{\mu}\in\mathbb{R}^n} \; x^{\mathrm{T}}\bar{\mu}, \\
&\text{s.t.} \;\; \|\hat{\Sigma}^{-\frac{1}{2}}(\bar{\mu}-\hat{\mu})\|_2 \le \sqrt{\gamma_1}, \qquad (12) \\
&\qquad e^{\mathrm{T}}(\bar{\mu}-\hat{\mu}) = 0.
\end{aligned}$$

Its conic dual is

$$\max_{q\in\mathbb{R}} \; x^{\mathrm{T}}\hat{\mu} - \sqrt{\gamma_1}\|\hat{\Sigma}^{\frac{1}{2}}(x-qe)\|_2. \qquad (13)$$

It can be shown that the optimal solution to the dual problem (13) is

$$q^* = \operatorname*{argmin}_{q\in\mathbb{R}} \|\hat{\Sigma}^{\frac{1}{2}}(x-qe)\|_2 = \frac{x^{\mathrm{T}}\hat{\Sigma}e}{e^{\mathrm{T}}\hat{\Sigma}e}.$$

Replacing q in the dual problem (13) with the foregoing equation, we can express the optimal value of problem (13) as

$$x^{\mathrm{T}}\hat{\mu} - \sqrt{\gamma_1}\|\hat{\Sigma}^{\frac{1}{2}}(x - \frac{x^{\mathrm{T}}\hat{\Sigma}e}{e^{\mathrm{T}}\hat{\Sigma}e}e)\|_2 = x^{\mathrm{T}}\hat{\mu} - \sqrt{\gamma_1}\|\Lambda^{\frac{1}{2}}x\|_2,$$

where $\Lambda = \hat{\Sigma} - \frac{1}{e^{\mathrm{T}}\hat{\Sigma}e}\hat{\Sigma}ee^{\mathrm{T}}\hat{\Sigma} \succ 0$ (by Cauchy-Schwartz inequality).

It is not difficult to show that problem (12) is strictly feasible in variable $\bar{\mu}$, and therefore the strong conic duality theorem holds (Shapiro 2001). In other words, the primal problem (12) has the same objective value with the dual problem (13). Thus, it is also easy to show that the optimal solution to the optimization problem $\min_{\bar{\mu}\in\mathcal{U}_{\hat{\mu}}^{adj}} x^{\mathrm{T}}\bar{\mu}$ is

$$\bar{\mu}_{wc} = \hat{\mu} - \frac{\sqrt{\gamma_1}\Lambda x}{\sqrt{x^{\mathrm{T}}\Lambda x}}. \qquad (14)$$

Hence, substituting (8) and (14) into (11), we have

$$\max_{\mathrm{P}\in\mathbb{D}_F^{adj}(\gamma_1,\gamma_2)} \mathrm{CVaR}_\beta(x,\mathrm{P}) = -(x^{\mathrm{T}}\hat{\mu} - \sqrt{\gamma_1}\|\Lambda^{\frac{1}{2}}x\|_2) + \kappa\sqrt{x^{\mathrm{T}}(\hat{\Sigma}+\gamma_2 I_n)x}.$$

Therefore, problem (RMC-$\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$) can be written as

$$\begin{aligned} \min_x \quad & \kappa\sqrt{x^{\mathrm{T}}(\hat{\Sigma} + \gamma_2 I_n)x} - x^{\mathrm{T}}\hat{\mu} + \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\Lambda x}, \\ \text{s.t.} \quad & x^{\mathrm{T}}\hat{\mu} - \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\Lambda x} \geq \rho, \quad x \in \mathcal{X}. \end{aligned}$$

Similarly, by introducing auxiliary variables δ and ω, the problem can be expressed as an SOCP:

$$\begin{aligned} \min_{x,\delta,\omega} \quad & \kappa\delta - \hat{\mu}^{\mathrm{T}}x + \sqrt{\gamma_1}\omega, \\ \text{s.t.} \quad & \sqrt{\gamma_1}\sqrt{x^{\mathrm{T}}\Lambda x} \leq \hat{\mu}^{\mathrm{T}}x - \rho, \\ & \sqrt{x^{\mathrm{T}}(\hat{\Sigma} + \gamma_2 I_n)x} \leq \delta, \\ & \sqrt{x^{\mathrm{T}}\Lambda x} \leq \omega, \quad x \in \mathcal{X}, \end{aligned}$$

where the variables are $x \in \mathbb{R}^n, \delta, \omega \in \mathbb{R}$. We now have the desired conclusion. □

Remark 5 For the case of $\gamma_1 = \gamma_2 = 0$, the robust and adjusted-robust counterparts to CVaR optimization for distribution ambiguity, i.e., (RMC-$\mathbb{D}_F(0, 0)$) and (RMC-$\mathbb{D}_F^{adj}(0, 0)$), can also be formulated as SOCPs, and we can easily observe that they are the same.

3. Simulation Analysis

3.1. *Calibration of parameters γ_1 and γ_2*

Use of the robust or adjusted-robust optimization approach requires choosing values for parameters γ_1 and γ_2 to control the degree of ambiguity. In other words, parameters γ_1 and γ_2 can be specified exogenously. To specify the parameter γ_1 that corresponds to the $100(1 - \alpha)\%$ confidence region, we can apply a similar statistical analysis approach as that used in Garlappi *et al.* (2007) under standard assumptions concerning the time series of the returns. For instance, if the population covariance matrix Σ is unknown, we can employ the $95\%(\alpha = 0.05)$ percentile of the F-distribution with N and $S - N$ degrees of freedom and use a quantile framework to set meaningful values for γ_1 in a practical computation using return data. This approach is reasonable from a statistical viewpoint, and its computational efficiency is attractive. In reality, however, we rarely have complete information on the distribution of asset returns. Therefore, a formal rule to guide an investor in making an appropriate choice of parameters is crucial in decisions based on a few historical samples.

Generally, we can use several nonparametric techniques, such as the bootstrapping technique and cross-validation principle (Efron and Gong 1979, DeMiguel *et al.* 2013, Maillet *et al.* 2015), to estimate parameters γ_1 and γ_2. However, because the cross-validation principle can produce very unstable sequences of parameters over time, we use bootstrapping techniques to calibrate the parameters endogenously. This methodology, which refers to redrawing historical observations with replacement, is highly intuitive. We assume that asset returns are independent and identically distributed, but impose no other assumptions on the distribution. The algorithm for computing the optimal parameters with the bootstrap analogue is as follows.

Bootstrapping procedure:

Step 1. Construct B bootstrap samples $\{Y_1, Y_2, \cdots, Y_B\}$ (for example, $B = 10000$) by drawing random observations with replacement from the available observations.

Step 2. For each bootstrap sample Y_b, compute the corresponding mean $\hat{\mu}_b$ and covariance matrix $\hat{\Sigma}_b$, and then generate a sample

$$\mathcal{C} = \left\{ (\hat{\mu}_b, \hat{\Sigma}_b) : b = 1, \cdots, B \right\}.$$

Step 3. For sample $\mathcal{C}$, define data sets $\mathcal{C}_{\gamma_1}$ and $\mathcal{C}_{\gamma_2}$ as

$$\mathcal{C}_{\gamma_1} = \left\{ \gamma_{1b} : \gamma_{1b} = (\hat{\mu}_b - \hat{\mu})^{\mathrm{T}} \hat{\Sigma}^{-1} (\hat{\mu}_b - \hat{\mu}), b = 1, \cdots, B \right\},$$
$$\mathcal{C}_{\gamma_2} = \left\{ \gamma_{2b} : \gamma_{2b} = \|\hat{\Sigma}_b - \hat{\Sigma}\|_F, b = 1, \cdots, B \right\}$$

to ensure reasonable values of γ_1 and γ_2. The percentiles of the empirical distributions of $\mathcal{C}_{\gamma_1}$ and $\mathcal{C}_{\gamma_1}$ can then be referenced to derive γ_1 and γ_2. Consequently, the calibrated values of γ_1 and γ_2 are

$$\hat{\gamma_1} = q_\zeta(\mathcal{C}_{\gamma_1}), \quad \hat{\gamma_2} = q_\zeta(\mathcal{C}_{\gamma_2}),$$

where $q_\zeta(\cdot)$ is an upper quantile of the corresponding data sets (for example, $\zeta = 95\%$).

In the following, we use an example with four assets considered in Yam *et al.* (2016) to evaluate the bootstrap procedure, where the four risky assets are chosen as the indices of S&P 500, DAX, HSI and FTSE 100, whose true mean vector and variance matrix are given by

$$\mu = \begin{pmatrix} 0.061166 \\ 0.109547 \\ 0.090358 \\ 0.040923 \end{pmatrix} \text{ and } \Sigma = \begin{pmatrix} 0.018632 & 0.020056 & 0.020646 & 0.015213 \\ 0.020056 & 0.034507 & 0.027412 & 0.020652 \\ 0.020646 & 0.027412 & 0.048680 & 0.021663 \\ 0.015213 & 0.020652 & 0.021663 & 0.018791 \end{pmatrix},$$

respectively. Assume that the return vector of underlying risky assets follows a multivariate Gaussian distribution $\mathcal{N}(\mu, \Sigma)$, by which we can easily generate random samples of returns for further analysis. This example will also be used in the Subsections 3.2 and 3.3 for simulation analysis.

In addition to the estimation of ambiguity parameters γ_1 and γ_2, it is also interested to clarify the relationship between the estimated ambiguity parameters $\hat{\gamma}_1$ and $\hat{\gamma}_2$ and the uncertainty of the estimations of the mean vector and the covariance matrix. Denote the sample mean vector and sample covariance matrix by $\hat{\mu}$ and $\hat{\Sigma}$, respectively. The parameter uncertainties due to the estimation error can be measured by the ellipsoidal norm of difference between $\hat{\mu}$ and the true mean vector μ, and the Frobenius norm of difference between $\hat{\Sigma}$ and the true covariance matrix Σ. Mathematically, these two measures are defined as

$$\gamma_1^{PU} = \|\hat{\mu} - \mu\|^2_{\hat{\Sigma}^{-1}}, \quad \gamma_2^{PU} = \|\hat{\Sigma} - \Sigma\|_F.$$

Based on 600 times of simulations, Table 1 displays the average values of the bootstrap-estimated values of the level of ambiguity ($\hat{\gamma}_1$ and $\hat{\gamma}_2$) and the average values of degrees of parameter uncertainty (γ_1^{PU} and γ_2^{PU}) for the different sample length S. As shown in Table 1, when the size of the fund universe is fixed, the average values of $\hat{\gamma}_1$, $\hat{\gamma}_2$, γ_1^{PU} and γ_2^{PU} decrease as the number of samples S increases. The intuition behind this result is that the larger the sample size, the less the uncertainty and ambiguity. Moreover, when parameter uncertainties in the estimated mean and covariance matrix (γ_1^{PU} and γ_2^{PU}) increase, the bootstrap-estimators of the levels of ambiguity ($\hat{\gamma}_1$ and $\hat{\gamma}_2$) also increase. The most interesting thing is that the bootstrap-estimated value of the level of ambiguity $\hat{\gamma}_1$ is relatively more sensitive to sample length S as compared with $\hat{\gamma}_2$.

In summary, the data-driven bootstrap approach allows us to obtain reliable information for the distribution of returns over the investment period. Moreover, to capture the time-series effects in

Table 1. The average values of the bootstrap-estimated values of the level of ambiguity ($\hat{\gamma}_1$ and $\hat{\gamma}_2$) and the average values of degrees of parameter uncertainty (γ_1^{PU} and γ_2^{PU}).

	S=100	S=200	S=500	S=800	S=1000
$\hat{\gamma}_1$	0.3748	0.1812	0.0713	0.0445	0.0355
$\hat{\gamma}_2$	0.0831	0.0793	0.0767	0.0762	0.0758
γ_1^{PU}	0.0410	0.0193	0.0082	0.0050	0.0040
γ_2^{PU}	0.0135	0.0097	0.0062	0.0049	0.0043

Notes. This table reports the average values of the bootstrap-estimated values of the levels of ambiguity $\hat{\gamma}_1$ and $\hat{\gamma}_2$ and the average values of degrees of parameter uncertainty γ_1^{PU} and γ_2^{PU} for different estimation sample length S. For each given estimation sample size, 600 simulations are performed to produce average results. The bootstrap samples B and the significance level ζ are set to 10000 and 95% respectively.

returns, one could adopt the stationary bootstrap method in Politis and Romano (1994). Methods of selecting these parameters remain an interesting topic for further investigation.

3.2. *Impact of γ_1 and γ_2 on portfolio decision*

To further study the effects of the levels of ambiguity on the optimal investment strategies of the two robust optimization problems (RMC-$\mathbb{D}_{\mathrm{F}}(\gamma_1, \gamma_2)$) and (RMC-$\mathbb{D}_{\mathrm{F}}^{\mathrm{adj}}(\gamma_1, \gamma_2)$), we define the following relative difference (RD) in the portfolio decision:

$$\text{RD-}x_i = \frac{x_i^*(\gamma_1, \gamma_2) - x_i^*(\hat{\gamma}_1, \hat{\gamma}_2)}{x_i^*(\hat{\gamma}_1, \hat{\gamma}_2)} \times 100\%, \text{ for } i = 1, 2, 3, 4, \tag{15}$$

where $x_i^*(\hat{\gamma}_1, \hat{\gamma}_2)$ and $x_i^*(\gamma_1, \gamma_2)$ are the weight of the robust portfolios obtained by using the parameters $(\hat{\gamma}_1, \hat{\gamma}_2)$ and (γ_1, γ_2), respectively. The indices $i = 1, 2, 3, 4$ represent four asset indices. Let us consider the problem (RMC) with the following range of parameters (a change of 20%):

$$\frac{|\gamma_i - \hat{\gamma}_i|}{\hat{\gamma}_i} \le 20\%, \text{ for } i = 1, 2.$$

More specifically, the value of parameters γ_1 and γ_2 varies in the range $[0.1449, 0.2174] \times [0.0635, 0.0951]$, respectively.

Figs. 1 and 2 demonstrate the relative differences of the two strategies for a range of values of these parameters. From Figs. 1 and 2 one can observe that the relative differences in the portfolios are less than 10%. In other words, a 20% error in the estimation of the parameters γ_1 and γ_2 leads to at most a 10% change in the relative difference. This indicates that a small change in two input parameters can only give rise to subtle changes in the optimal portfolio weights. For instance, for values of γ_2 in the range $[0.0635, 0.0951]$, the relative difference RD-x_2 in both RMC-$\mathbb{D}_{\mathrm{F}}(\gamma_1, \gamma_2)$ and RMC-$\mathbb{D}_{\mathrm{F}}^{\mathrm{adj}}(\gamma_1, \gamma_2)$ can be smaller than 2%. Thus, values of γ_1 and γ_2 within the range used in our test do not appear to have a great impact on the behaviour of the robust strategies and our robust approach is relatively immune to variations in input values. In particular, the relative difference for the robust approach are relatively insensitive to the value of γ_1.

3.3. *Comparison of efficient frontiers and diversifications*

In this subsection, the (robust) efficient frontiers are drawn to examine the differences among three models. We adopt a similar market setting as described in Subsection 3.1 and compute portfolios that lie in the set, $\mathcal{X} = \{x \in \mathbb{R}^n \mid e^{\mathrm{T}}x = 1, x \ge 0\}$. The estimation sample is set at length $S = 200$.

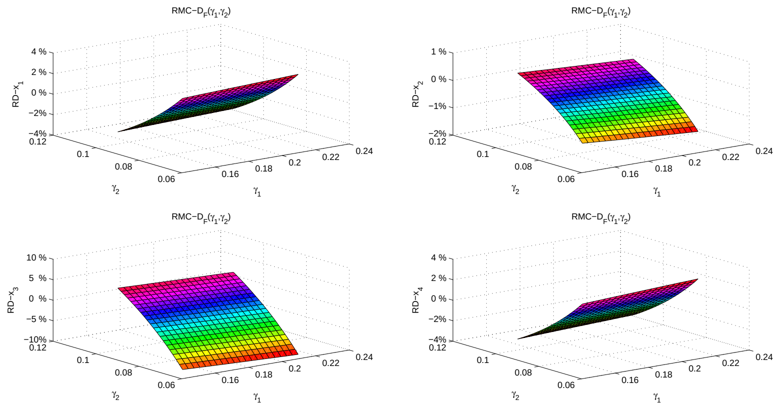

Figure 1. Relative difference RD-x_i ($i = 1, 2, 3, 4$, in percentage) of strategies (RMC-$\mathbb{D}_{\mathrm{F}}(\gamma_1, \gamma_2)$) for varying parameters γ_1 and γ_2. The bootstrap samples B and the significance level ζ used in bootstrapping techniques are set to 10000 and 95%. The estimation sample is set at length $S = 200$. The confidence level used in calculation of CVaR is $\beta = 0.95$ and target return is $\rho = 0.01$.

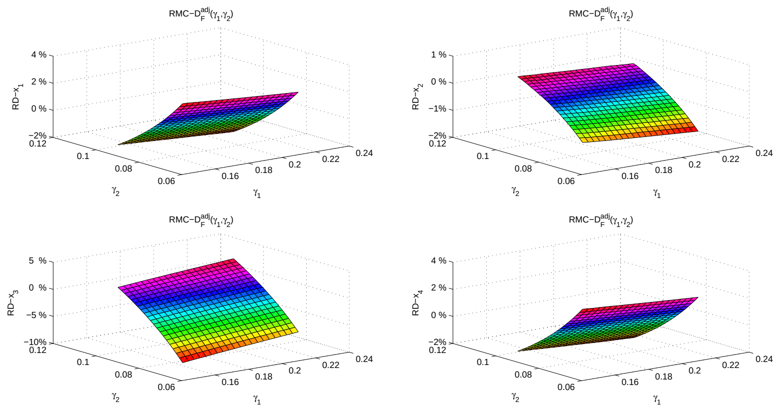

Figure 2. Relative difference RD-x_i ($i = 1, 2, 3, 4$, in percentage) of strategies (RMC-$\mathbb{D}_{\mathrm{F}}^{\mathrm{adj}}(\gamma_1, \gamma_2)$) for varying parameters γ_1 and γ_2. The bootstrap samples B and the significance level ζ used in bootstrapping techniques are set to 10000 and 95%. The estimation sample is set at length $S = 200$. The confidence level used in calculation of CVaR is $\beta = 0.95$ and target return is $\rho = 0.01$.

The bootstrap samples B and the significance level ζ used in bootstrapping techniques are also set to 10000 and 95%. The algorithm used to generate a discrete approximation to the robust efficient frontier, which is formally outlined below, is similar to that in Tütüncü and Koenig (2004) and Ye *et al.* (2012). However, unlike those authors, we use CVaR as the risk measure to determine the efficient set of portfolios from the mean-CVaR model or robust mean-CVaR model (RMC). Moreover, to measure the effect of the zero net adjustment methodology on the efficient frontier, we also run the experiment using adjusted-robust mean-CVaR model.

Procedure for generating robust Mean-CVaR efficient frontier:

Step 1. Solve RMC without the expected return constraint to compute the global minimum-risk portfolio x_{min} and attain the worst-case mean μ_{wc}. Set $\rho_{min} = \mu_{wc}^{\mathrm{T}} x_{min}$.

Step 2. Solve problem $\max_{x \in \mathcal{X}} \{\min_{\mathrm{P} \in \mathbb{D}} \mathrm{E}_{\mathrm{P}}(\xi)^{\mathrm{T}} x\}$ to attain the optimal portfolio x_{max} and corresponding objective value ρ_{max}. Set $\Delta = \rho_{max} - \rho_{min}$.

Step 3. Choose M, the number of desired points on the efficient frontier. For $\rho \in \{\rho_{min}, \rho_{min} +$

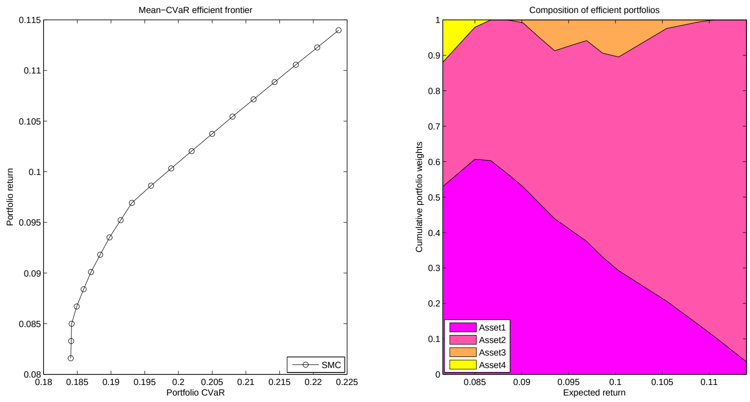

Figure 3. The efficient frontier and the composition of efficient portfolios from the SMC. The percentage allocation of assets 1-4 in the optimal allocation x^* has been illustrated in different colors.

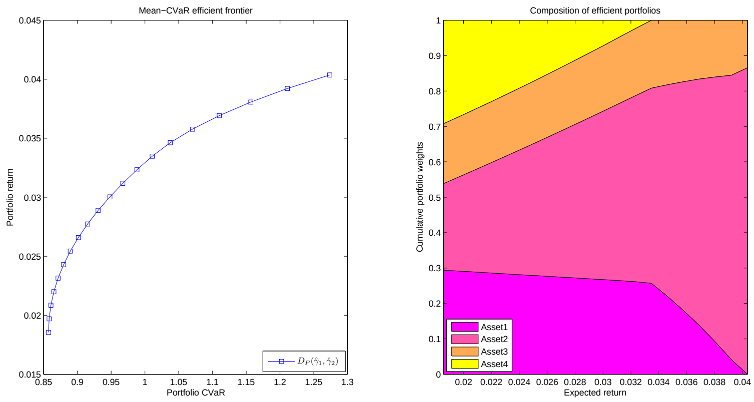

Figure 4. The efficient frontier and the composition of efficient portfolios from $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$. The portfolio CVaR in the x-axis and the portfolio return in the y-axis are the worst case values of CVaR risk measure and expected return, respectively. The percentage allocation of assets 1-4 in the optimal allocation x^* has been illustrated in different colors.

$\frac{\Delta}{M-1}, \rho_{min} + 2 \cdot \frac{\Delta}{M-1}, \cdots, \rho_{min} + (M-1) \cdot \frac{\Delta}{M-1}\}$, solve RMC with expected return constraint.

The RMC model depends on the distributional ambiguity set $\mathbb{D}$: it is either $\mathbb{D}_F(\gamma_1, \gamma_2)$ or $\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$. Corresponding models (RMC-$\mathbb{D}_F(\gamma_1, \gamma_2)$) and (RMC-$\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$) will be denoted, for shortly, as $\mathbb{D}_F(\gamma_1, \gamma_2)$ and $\mathbb{D}_F^{adj}(\gamma_1, \gamma_2)$, respectively.

Fig. 3 presents the sample-based mean-CVaR (SMC) efficient frontier and the composition of efficient portfolios. The values of CVaR range from 0.1840 for the minimum-CVaR portfolio (which at the same time has a minimum expected return of 0.0816) to 0.2237 for the portfolio with a maximum expected return (0.1140). The minimum-CVaR portfolio primarily consists of a position in asset 1, while encompassing only a small position in assets 2 and 4. The maximum-return portfolio is made up of a large position in asset 2, the index with a maximum expected return. In Figs. 4 and 5, we illustrate the robust efficient frontier and composition of robust efficient portfolios identified using the classical robust $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ and adjusted-robust asset allocation approach $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$. Comparing the compositions of the SMC and RMC efficient portfolios, we can see that the weights of the asset 2 increase throughout the full spectrum of the expected return. As Figs. 4 and 5 show, the allocation transition proceeds smoothly from one portfolio to the next. Moreover, the robust

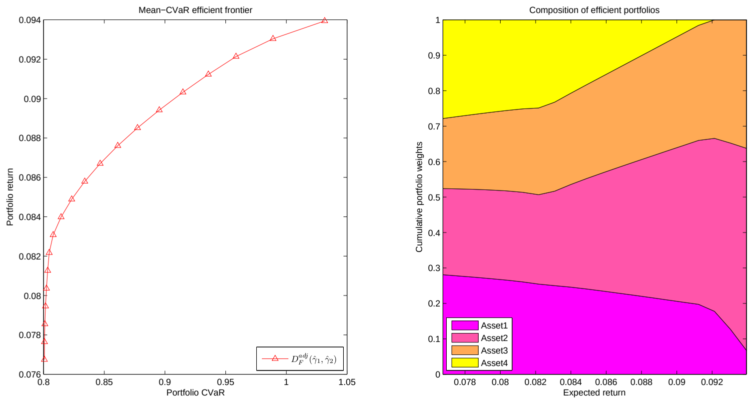

Figure 5. The efficient frontier and the composition of efficient portfolios from $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$. The portfolio CVaR in the x-axis and the portfolio return in the y-axis are the worst case values of CVaR risk measure and expected return, respectively. The percentage allocation of assets 1-4 in the optimal allocation x^* has been illustrated in different colors.

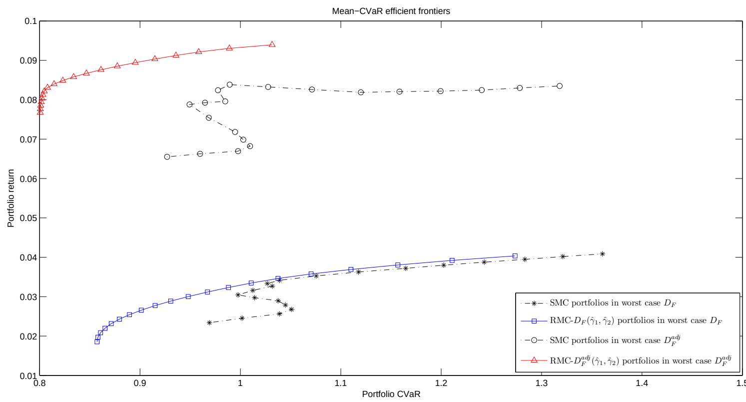

Figure 6. The efficient frontiers of $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ models in the worst case. The blue and red curves are the robust efficient frontiers. The black curves are the trajectories of worst-case risk and return of SMC efficient portfolios.

optimal portfolios deliver greater diversification than those identified using the SMC approach, and the weights of the individual assets change smoothly when moving toward higher returns.

Fig. 6 compares the worst-case performance of SMC and RMC efficient portfolios. The two black curves reflect the performances of SMC portfolios in the worst cases $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$, respectively. Each point on the curves correspond to the worst-case risk and return of SMC efficient portfolios. As evidenced by the figure, the SMC curve is obviously below the robust efficient frontier in each case. In other words, the robust portfolios (blue and red curves) perform substantially better than the SMC portfolios in the worst cases ($\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$).

4. Empirical Analysis

This section reports the results of empirical study for the robust portfolio, based on real market data. As is common practice in the literature, we use rolling horizon analysis to evaluate the out-of-sample performance of different portfolio criteria and choose CVaR with parameter $\beta = 0.95$

Table 2. Dow Jones Credit Suisse Hedge Fund Indexes.

1	Convertible Arbitrage
2	Dedicated Short Bias
3	Emerging Markets
4	Equity Market Neutral
5	Event Driven
6	Distressed
7	Multi-Strategy
8	Risk Arbitrage
9	Fixed Income Arbitrage
10	Global Macro
11	Long/Short Equity
12	Managed Futures

as risk measure in our empirical tests. All computations are performed on a PC with an Intel(R) Core(TM) 2.30 GHz processor and 8 GB RAM employing the simplex code in the CPLEX 12.1 package (IBM ILOG CPLEX 2009).

Here, the out-of-sample performance of the (adjusted) robust mean-CVaR model based on several widely accepted measures is compared with that of the SMC model using real market data. More specifically, we assess the two models' mean portfolio return, portfolio standard deviations (SDs), reward to risk ratios and portfolio turnover. The levels of ambiguity γ_1 and γ_2, are calibrated via a bootstrapping procedure, that is, the data-based simulation method for statistical inference described in the previous section. We acknowledge that there are many other ways one can choose the portfolio and we do not intend to indicate that our approach is the best one for CVaR optimization problem. Instead, we aim to illustrate that the RMC approaches (RMC-$\mathbb{D}_{\mathrm{F}}(\gamma_1, \gamma_2)$ and RMC-$\mathbb{D}_{\mathrm{F}}^{\mathrm{adj}}(\gamma_1, \gamma_2)$) give a decent performance with some desired features.

4.1. *Data and methodology*

Our dataset, extracted free of charge from http://www.hedgeindex.com, comprises monthly excess returns on the 12 Credit Suisse/Tremont Hedge Fund indices (see Table 2). The sample period of these historical data is January 1994 to December 2015 (for 264 observations in total). Because the return data on the Multi-Strategy Hedge Fund indices for the first three months are incomplete, they are omitted in constructing the hedge fund portfolios. Statistical tests show that most of the returns of these indices exhibit negative skewness and a high degree of kurtosis (Zhu *et al.* 2015). For simplicity, assuming that the risk free rate is zero, we shall henceforth interpret the reward to risk ratio as the Sharpe ratio.

Throughout the experiment, portfolio performance is evaluated on the basis of a monthly rebalancing portfolio strategy. A rolling horizon procedure similar to that in DeMiguel and Nogales (2009) is used for portfolio construction. More specifically, given an L-month-long dataset of asset returns, we choose moving windows of S months in length and then compute the corresponding optimal portfolio weights, which are considered to constitute the portfolios for period $S + 1$. This process is continued by removing the first return and adding a return to the dataset for the next period until the end of the dataset is reached. Note that at the end of the procedure, we obtain $L - S$ portfolio weights x_t^*, $t = S, \cdots, L - 1$, with corresponding out-of-sample returns for each portfolio, allowing us to compare the performance of the two models.

For each strategy k, the out-of-sample mean $\tilde{\mu}_k$, SD $\tilde{\sigma}_k$, Sharpe ratio $\widetilde{sr}_k$, average turnover $\widetilde{trn}_k$

and out-of-sample CVaR $\mathrm{CVaR}_{\beta,k}$, respectively, are given by

$$\tilde{\mu}_k = \frac{1}{L-S}\sum_{t=S}^{L-1}(\xi^{[t+1]})^{\mathrm{T}}x^*_{k,t},$$

$$\tilde{\sigma}_k = \sqrt{\frac{1}{L-S-1}\sum_{t=S}^{L-1}[(\xi^{[t+1]})^{\mathrm{T}}x^*_{k,t} - \tilde{\mu}_k]^2},$$

$$\widetilde{sr}_k = \frac{\tilde{\mu}_k}{\tilde{\sigma}_k},$$

$$\widetilde{trn}_k = \frac{1}{L-S-1}\sum_{t=S}^{L-1}\sum_{j=1}^{N}|x^*_{k,j,t+1} - x^*_{k,j,t^+}|,$$

$$\mathrm{CVaR}_{\beta,k} = \min_{\eta}\{\eta + \frac{1}{(L-S)(1-\beta)}\sum_{t=S}^{L-1}\max\{-(\xi^{[t+1]})^{\mathrm{T}}x^*_{k,t} - \eta, 0\}\}$$

where $x^*_{k,j,t}$ denotes the portfolio weight in asset j at time t under strategy k, x^*_{k,j,t^+} the relative portfolio weights after return $\xi^{[t+1]}$ has been realized but before the rebalancing decision in period $t+1$, and $x^*_{k,j,t+1}$ the desired portfolio weight after rebalancing at time $t+1$. Because the asset weights before rebalancing are given by

$$x^*_{k,j,t^+} = \frac{(1+\xi_j^{[t+1]})x^*_{k,j,t}}{\sum_{j=1}^{N}(1+\xi_j^{[t+1]})x^*_{k,j,t}},$$

the turnover of strategy k is

$$\widetilde{trn}_k = \frac{1}{L-S-1}\sum_{t=S}^{L-1}\sum_{j=1}^{N}\left|(x^*_{k,j,t+1} - x^*_{k,j,t}) - \frac{x^*_{k,j,t}(\xi_j^{[t+1]} - \xi_p^k)}{1+\xi_p^k}\right|,$$

where $\xi_p^k = \sum_{j=1}^{N}\xi_j^{[t+1]}x^*_{k,j,t}$ is the portfolio return from strategy k at time $t+1$. Portfolio turnover measures the average percentage of wealth traded for a given strategy, and is used primarily to demonstrate the realistic nature of the practical implementation of our strategies when transaction costs are considered. The smaller the turnover $\widetilde{trn}$, the lower the transaction costs.

There are 264 observations over L months for each hedge fund considered. As we have only a limited number of samples, we are unable to determine with confidence the distribution of asset returns without any ambiguity. Some of the observations are used as a training set to estimate the sample mean and covariance, and the estimation window is then "rolled" one month forward by dropping the earliest return and adding a new return. Further, to cover the period of the recent financial crisis, we used the full January 2008 to December 2015 period in our performance evaluation. The dataset for the January 1994-December 2007 period is used for the initial parameter estimation. For a more comprehensive comparison and understanding, in addition to the initial estimation, we also carry out an additional re-estimation based on data from January 1994 to December 2011 in the middle of the rebalancing period, i.e., January 2012. Thus, in our later analysis, the portfolio performance of the first 4-year period (January 2008 to December 2011) is presented separately from that of the latter 4-year period (January 2012 to December 2015). Obviously, the test dataset cannot be used to calibrate γ_1 and γ_2, and we thus need to calibrate

Table 3. The bootstrap-estimated values of the level of ambiguity $\hat{\gamma}_1$, $\hat{\gamma}_2$.

	$\hat{\gamma}_1$	$\hat{\gamma}_2$
1994/01-2007/12	22.2314	0.0045
1994/01-2011/12	16.7532	0.0042

Notes. This table reports the bootstrap-estimated values of the level of ambiguity $\hat{\gamma}_1$, $\hat{\gamma}_2$. The bootstrap samples B and the significance level ζ are set to 10000 and 95%.

Table 4. Performances of different portfolio strategies.

	$\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$	$\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$	$\mathbb{D}_F^{adj}(\hat{\gamma}_1, 0)$	$\mathbb{D}_F(\hat{\gamma}_1, 0)$	$\mathbb{D}_F(0, 0)$	SMC
Mean	0.0004	-0.0007	-0.0018	-0.0044	-0.0046	-0.0053
Std	0.0179	0.0168	0.0290	0.0394	0.0406	0.0473
Sharpe	0.0237	(-0.0418)	(-0.0633)	(-0.1108)	(-0.1121)	(-0.1114)
Turnover	0.0172	0.0194	0.0445	0.0497	0.0507	0.0647
$\text{CVaR}_{0.95}$	0.0592	0.0674	0.1774	0.2638	0.2729	0.3208

Notes. This table reports the monthly out-of-sample means, standard deviations, Sharpe ratios, turnover and CVaR of the different portfolio strategies. The data consist of monthly returns on twelve Credit Suisse Hedge Fund indices from January 1994 to December 2011 (216 observations). The portfolio weights for each strategy are determined each month using moments estimated from a rolling-window of 168 months. The resulting out-of-sample period spans from January 2008 to December 2011.

them on the training dataset via the bootstrap method. The estimations of γ_1 and γ_2 with respect to the 95% confidence region across the $L - S$ optimizations are shown in Table 3.

In the rolling procedure, some portfolio optimization problems may be infeasible because of the relatively high specified (worst-case) expected returns required. To render our optimization problems simpler to solve, we first set the parameter ρ as -5% (which means that investors are willing to bear a maximum loss for 5% in the worst-case scenarios). To produce less conservative strategies, furthermore, the values of ρ is set as the simple average of all the mean values of asset returns, denoted by ρ_{average}, which is estimated by the historical samples within the current time period. If infeasibility occurs, we reduce the required expected returns by 20% and resolve the problem until it becomes feasible.

4.2. *Discussion of results*

The results of the model comparison in terms of monthly out-of-sample means, SDs, Sharpe ratios, turnover and CVaR for the January 2008-December 2011 and January 2012-December 2015 periods are presented in Tables 4 and 5. It can be seen that the portfolios calculated using the RMC model perform well in terms of the expected return relative to the other models.

Comparing the performance of $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2), \mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2), \mathbb{D}_F^{adj}(\hat{\gamma}_1, 0)$ and $\mathbb{D}_F(\hat{\gamma}_1, 0)$ shows that $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F^{adj}(\hat{\gamma}_1, 0)$ have a clear advantage over $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F(\hat{\gamma}_1, 0)$ owing to the inclusion of a less pessimistic view of the expected return. In terms of monthly turnover, the adjusted-robust approach delivers better results than traditional robust optimization (non-adjusted robust approach). The worst turnover is achieved by the SMC portfolios, which yield a turnover of 0.0647 for the January 2008-December 2011 period and 0.0183 for the January 2012-December 2015 period. We can also see that the Sharpe ratio of the strategy that yields the largest values is the $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ model. Owing to market instability, the other strategies result in a negative Sharpe ratio during the earlier period. In the computation of the CVaR for the optimal portfolio x^*, we solve the problem (1) with x being replaced by x^*. As expected, the adjusted-robust approach does a better job than the other methods. Moreover, the performance of the portfolios obtained by $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$ is superior to that of $\mathbb{D}_F(0, 0)$ in terms of both average return and

Table 5. Performances of different portfolio strategies.

	$\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$	$\mathbb{D}_F(\hat{\gamma}_1, \hat{\gamma}_2)$	$\mathbb{D}_F^{adj}(\hat{\gamma}_1, 0)$	$\mathbb{D}_F(\hat{\gamma}_1, 0)$	$\mathbb{D}_F(0, 0)$	SMC
Mean	0.0024	0.0014	0.0019	0.0007	0.0007	0.0002
Std	0.0068	0.0057	0.0056	0.0052	0.0053	0.0060
Sharpe	0.3525	0.2506	0.3466	0.1319	0.1420	0.0262
Turnover	0.0108	0.0128	0.0158	0.0138	0.0136	0.0183
$\text{CVaR}_{0.95}$	0.0123	0.0146	0.0141	0.0193	0.0193	0.0197

Notes. This table reports the monthly out-of-sample means, standard deviations, Sharpe ratios, turnover and CVaR of the different portfolio strategies. The data consist of monthly returns on twelve Credit Suisse Hedge Fund indices from January 1994 to December 2015 (264 observations). The portfolio weights for each strategy are determined each month using moments estimated from a rolling-window of 216 months. The resulting out-of-sample period spans from January 2012 to December 2015.

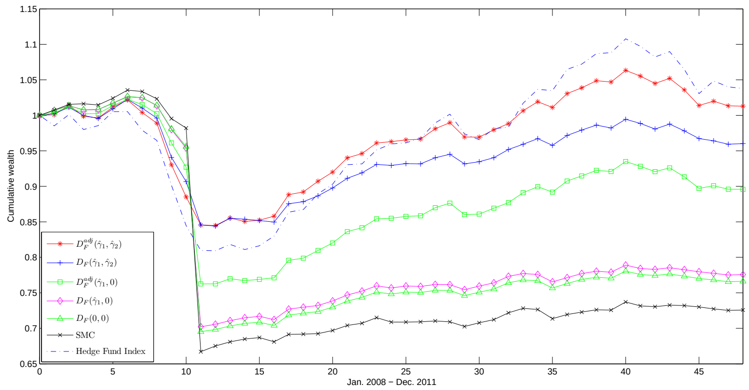

Figure 7. The cumulative wealth of the trading strategies over the period January 2008 to December 2011 for target mean return, $\rho = -5\%$. The evolution of each hedge fund index is also provided for reference purposes.

turnover. This result confirms the importance of taking moment uncertainty into account in real-life portfolio selection, particularly in the presence of uncertainty about the second moments.

Figs. 7-9 show the cumulative wealth of the different dynamic portfolio strategies for each model for the January 2008-December 2011 period under different level of required expected return. They illustrate the performance of the robust strategies for different input parameter values and offer insight into the features of different trading strategies. It is obvious that the cumulative portfolio returns generated by the RMC model are better than those generated by the SMC model at the end of the investment period. Of the six portfolios, that derived by the $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ model typically provides the greatest cumulative wealth and achieves superior performance to the other approaches in terms of the average return. This result highlights the potential benefit and effectiveness of hedging the downside risk by taking moment uncertainty and zero net adjustment into account simultaneously. Furthermore, it's worth noting that the $\mathbb{D}_F^{adj}(\hat{\gamma}_1, \hat{\gamma}_2)$ and $\mathbb{D}_F^{adj}(\hat{\gamma}_1, 0)$ strategies generated by setting ρ as ρ_{average} almost perform similarly.

We can see that the SMC model exhibits vulnerability in a consistently volatile market, with its cumulative wealth declining considerably when the market crashed during the financial crisis. One possible explanation is that the return series is quite skewed and heavy-tailed. The robust optimization approach yields a more stable strategy over time, and the returns of the adjusted-robust portfolio experience relatively little fluctuation. In this regard, the adjusted-robust optimization approach is more robust to the risk of a crash than the other methods, and is thus better able to choose portfolios that exhibit robust performance, which is a desirable property. Note that in Figs. 7-9, the evolution of each hedge fund index is also provided for reference purposes. The empiri-

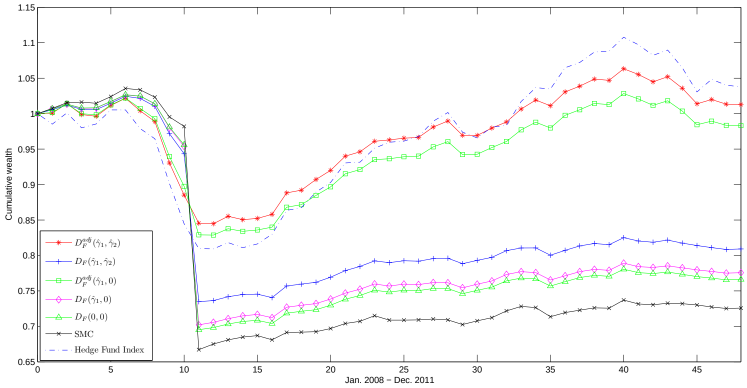

Figure 8. The cumulative wealth of the trading strategies over the period January 2008 to December 2011 for target mean return, $\rho = 0.5 \times \rho_{\text{average}}$. The evolution of each hedge fund index is also provided for reference purposes.

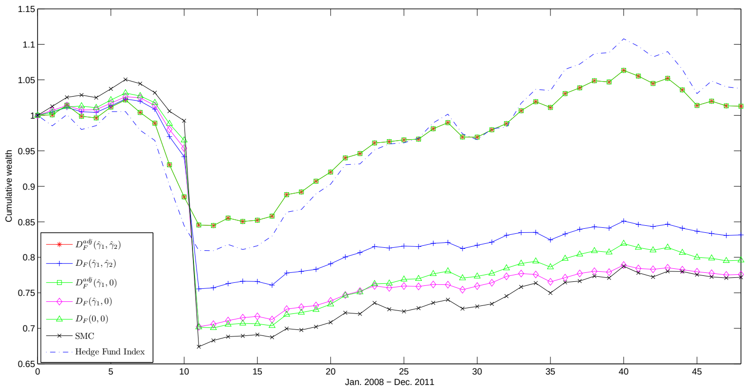

Figure 9. The cumulative wealth of the trading strategies over the period January 2008 to December 2011 for target mean return, $\rho = \rho_{\text{average}}$. The evolution of each hedge fund index is also provided for reference purposes.

cal results do not allow us to draw general conclusions about which model is intrinsically better without more intensive tests. In a downward market, however, the robust strategies are clearly better.

5. Conclusions

In this paper, we show that the worst-case CVaR (WCVaR) risk measure over distributional ambiguity sets can be computed efficiently via conic optimization techniques. In particular, the second-order cone technique for the linear loss function case. We introduce a data-driven criterion for calibrating the levels of ambiguity based on bootstrapping, which provides an important modeling guidance and may be of interest to practitioners. The results of numerical experiments with simulated and real market data demonstrate that our robust methods can construct more diversified portfolios and superior to its non-robust counterpart in terms of portfolio stability, expected returns and turnover. Although a number of ambiguity sets on probability distributions under uncertainty have been proposed, to the best of our knowledge, there seems to be no consensus on whether or not the robust CVaR optimization solution under an ambiguity set is intrinsically better than the one under an alternative one. It remains an interesting topic that deserves to be investigated in future.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 71721001 and 71471180); the Natural Science Foundation of Guangdong Province of China (No. 2014A030312003); the Hong Kong RGC grants (Nos. 15209614, 15224215 and 15255416). The authors also wish to thank the reviewers and the editor for providing feedback which helped improve the final version of this paper.

References

Alizadeh, F. and Goldfarb, D., Second-order cone programming. *Math. Program.,* 2003, **95**, 3–51.

Artzner, P., Delbaen, F., Eber, J.M. and Heath, D., Coherent measures of risk. *Math. Financ.,* 1999, **9**, 203–228.

Ben-Tal, A., Den Hertog, D., Waegenaere, A.D., Melenberg, B. and Rennen, G., Robust solutions of optimization problems affected by uncertain probabilities. *Manag. Sci.* 2013, **59**, 341–357.

Bertsimas, D., Gupta, V. and Kallus, N., Robust sample average approximation. *Math. Program.,* 2016, **3**, 1–66.

Bertsimas, D., Gupta, V. and Kallus, N., Data-driven robust optimization. *Math. Program.,* 2017. Available online at: https://doi.org/10.1007/s10107-017-1125-8

Calafiore, G.C. and El Ghaoui, L., On distributionally robust chance-constrained linear programs. *J. Optimiz. Theory. App.,* 2006, **130**, 1–22.

Cao, H.H., Wang, T. and Zhang, H.H., Model uncertainty, limited market participation, and asset prices. *Rev. Financ. Stud.,* 2005, **18**, 1219–1251.

Ceria, S. and Stubbs, R.A., Incorporating estimation errors into portfolio selection: Robust portfolio construction. *J. Asset. Mgmt.,* 2006, **7**, 109–127.

Chen, L., He, S. and Zhang, S., Tight bounds for some risk measures, with applications to robust portfolio selection. *Oper. Res.,* 2011, **59**, 847–865.

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

DeMiguel, V., Martin-Utrera, A. and Nogales, F.J., Size matters: Optimal calibration of shrinkage estimators for portfolio selection. *J. Bank. Finance.,* 2013, **37**, 3018–3034.

DeMiguel, V. and Nogales, F.J., Portfolio selection with robust estimation. *Oper. Res.,* 2009, **57**, 560–577.

Efron, B. and Gong, G., A leisurely look at the bootstrap, the jackknife, and cross validation. *The American Statistician.,* 1979, **37**, 36–48.

Ellsberg, D., Risk, ambiguity, and the savage axioms. *Q.J. Econ.,* 1961, **75**(4): 643–669.

Epstein, L.G., A paradox for the smooth ambiguity model of preference. *Econometrica* 2010, **78**, 2085–2099.

Esfahani, P.M. and Kuhn, D., Data-driven distributionally robust optimization using the Wasserstein metric: performance guarantees and tractable reformulations. *Math. Program.,* 2017. Available online at: https://doi.org/10.1007/s10107-017-1172-1

Faria, G. and Correia-da-Silva, J., A closed-form solution for options with ambiguity about stochastic volatility. *Review of Derivatives Research,* 2014, **17**, 125–159.

Garlappi, L., Uppal, R. and Wang, T., Portfolio selection with parameter and model uncertainty: A multi-prior approach. *Rev. Financ. Stud.,* 2007, **20**, 41–81.

Ghirardato, P., Maccheroni, F. and Marinacci, M., Differentiating ambiguity and ambiguity attitude. *J Econom Theory.,* 2004, **118**(2): 133–173.

Gilboa, I. and Schmeidler, D., Maxmin expected utility with non-unique prior. *J Math Econ.,* 1989, **18**, 141–153.

Goldfarb, D. and Iyengar, G., Robust portfolio selection problems. *Math. Oper. Res.,* 2003, **28**, 1–38.

Gotoh, J., Shinozaki, K. and Takeda, A., Robust portfolio techniques for mitigating the fragility of CVaR minimization and generalization to coherent risk measures. *Quant. Finance.,* 2013, **13**, 1621–1635.

Hellmich, M. and Kassberger, S., Efficient and robust portfolio optimization in the multivariate generalized hyperbolic framework. *Quant. Finance.,* 2011, **11**, 1503–1516.

Huang, D., Zhu, S., Fabozzi, F. J. and Fukushima, M., Portfolio selection under distributional uncertainty: A relative robust CVaR approach. *Eur. J. Oper. Res.,* 2010, **203**, 185–194.

Izhakian, Y., Expected utility with uncertain probabilities theory. *J Math Econ.,* 2017, **69**, 91–103.

IBM ILOG CPLEX 12.1–Parameters Reference Manual [online]. 2009. Available online at: www.ilog.com. (accessed 23 October 2015).

Jiang, R. and Guan, Y., Data-driven chance constrained stochastic program. *Math. Program.*, 2016, **158**(1-2), 291–327.

Klibanoff, P., Marinacci, M. and Mukerji, S., A smooth model of decision making under ambiguity. *Econometrica.,* 2005, **73**, 1849–1892.

Klibanoff, P., Marinacci, M. and Mukerji, S., On the smooth ambiguity model: A reply. *Econometrica.,* 2012, **80**, 1303–1321.

Kim, J.H., Kim, W.C. and Fabozzi, F.J., Recent developments in robust portfolios with a worst-case approach. *J. Optimiz. Theory. App.,* 2014, **161**, 103–121.

Knight, F.H., Risk, uncertainty and profit, 1921 (Houghton Mifflin Company: Boston, MA).

Lim, A.E., Shanthikumar, J.G. and Vahn, G.Y., Conditional value-at-risk in portfolio optimization: Coherent but fragile. *Oper. Res. Lett.,* 2011, **39**, 163–171.

Lotfi, S. and Zenios, S.A., Equivalence of robust VaR and CVaR optimization. Working Paper, The Wharton Financial Institutions Center WP, 2016.

Ma, X., Zhao, Q. and Qu, J., Robust portfolio optimization with a generalized expected utility model under ambiguity. *Ann Financ,* 2008, **4**(4): 431–444.

Maillet, B., Tokpavi, S. and Vaucher, B., Global minimum variance portfolio optimisation under some model risk: A robust regression-based approach. *Eur. J. Oper. Res.,* 2015, **244**, 289–299.

Natarajan, K., Pachamanova, D. and Sim M., Constructing risk measures from uncertainty sets. *Oper. Res.,* 2009, **57**, 1129–1141.

Natarajan, K., Sim, M. and Uichanco, J., Tractable robust expected utility and risk models for portfolio optimization. *Math. Financ.,* 2010, **20**, 695–731.

Paç, A.B. and Pinar, M.Ç., Robust portfolio choice with CVaR and VaR under distribution and mean return ambiguity. *TOP.,* 2014, **22**, 875–891.

Perchet, R., Xiao, L., Carvalho, R.L.D. and Heckel, T., Insights into robust portfolio optimization: Decomposing robust portfolios into mean-variance and risk-based portfolios. Working Paper, Columbia University, 2015.

Politis, D.N. and Romano, J.P., The stationary bootstrap. *J. Am. Stat. Assoc.,* 1994, **89**, 1303–1313.

Popescu, I., Robust mean-covariance solutions for stochastic optimization. *Oper. Res.,* 2007, **55**, 98–112.

Quaranta, A. G. and Zaffaroni, A., Robust optimization of conditional value at risk and portfolio selection. *J. Bank. Finance.,* 2008, **32**, 2046–2056.

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

Santos, A.A.P., The out-of-sample performance of robust portfolio optimization. *Brazilian. Rev. Financ.,* 2010, **8**, 141–166.

Salahi, M., Mehrdoust, F. and Piri, F., CVaR robust mean-CVaR portfolio optimization. *Applied Maths.,* 2013, doi:10.1155/2013/570950.

Scarf, H., Arrow, K.J. and Karlin, S., A min-max solution of an inventory problem. *Studies in the Mathematical Theory of Inventory and Production.,* 1958, **10**, 201–209.

Shapiro, A., On duality theory of conic linear problems. In *Semi-Infinite Programming: Recent Advancese*, pp. 135–165, 2001 (Springer: US).

Tütüncü, R.H. and Koenig, M., Robust asset allocation. *Ann. Oper. Res.,* 2004, **132**, 157–187.

Ui, T., The ambiguity premium vs the risk premium under limited market participation. *Rev Financ.,* 2011, **15**, 245–275.

Wang, Z., Glynn, P. and Ye, Y., Likelihood robust optimization for data-driven newsvendor problems. *Comput Manag Sci* 2016, **13**, 241–261.

Wozabal, D., Robustifying convex risk measures for linear portfolios: A nonparametric approach. *Oper. Res.,* 2014, **62**, 1302–1315.

Xiao, Y. and Valdez, E., A Black Litterman asset allocation model under elliptical distributions. *Quant. Finance.,* 2015, **15**, 509–519.

Yao, H., Li, Z. and Lai, Y., Mean-CVaR portfolio selection: A non-parametric estimation framework. *Comput.*

Oper. Res., 2013, **40**, 1014–1022.
Yam S.C.P., Yang H. and Yuen F.L., Optimal asset allocation: Risk and information uncertainty. *Eur. J. Oper. Res.,* 2016, **251**, 554–561.
Ye, K., Parpas, P. and Rustem, B., Robust portfolio optimization: a conic programming approach. *Comput. Optim. Appl.,* 2012, **52**, 463–481.
Zhu, S. and Fukushima, M., Worst-case conditional value-at-risk with application to robust portfolio management. *Oper. Res.,* 2009, **57**, 1155–1168.
Zhu, S., Ji, X.D. and Li, D., A robust set-valued scenario approach for handling modeling risk in portfolio optimization. *J. Comput. Financ.,* 2015, **19**, 11–40.
Zymler, S., Rustem, B. and Kuhn, D., Robust portfolio optimization with derivative insurance guarantees. *Eur. J. Oper. Res.,* 2011, **210**, 410–424.