

---

# Equivalence of robust VaR and CVaR optimization

Somayyeh Lotfi · Stavros A. Zenios

Working Paper 16-03  
The Wharton Financial Institutions Center  
The Wharton School, University of Pennsylvania, PA.  
Date of first version: April 4, 2016.

**Abstract** We show that robust optimization of the VaR and CVaR risk measures with a minimum return constraint under distribution ambiguity reduce to the same second order cone program. We use this result to formulate models for robust risk optimization under joint ambiguity in distribution, mean returns and covariance matrices, under ellipsoidal ambiguity sets. We also obtain models for robust VaR and CVaR optimization for polytopic and interval ambiguity sets of the means and covariance. The models unify and/or extend several existing models. We also propose an algorithm and a heuristic for constructing an ellipsoidal ambiguity set from point estimates given by multiple securities analysts, and show how to overcome the well-known conservatism of robust optimization models. Using CDS spread return data from eurozone crisis countries we illustrate that investment strategies using robust optimization models perform well even out-of-sample. Finally, using a controlled experiment we show how the well-known sensitivity of CVaR to mis-specifications of the first four moments of the distribution is alleviated with the robust models.

**Keywords** Data ambiguity · Coherent risk measures · Robust optimization · Value-at-risk · Conditional value-at-risk · Portfolio strategies · Scenarios · Eurozone crisis.

---

Somayyeh Lotfi  
slotfi@phd.guilan.ac.ir  
Department of Applied Mathematics, Faculty of Mathematical Sciences,  
University of Guilan, P.O. Box 1914, Rasht, Iran.

Stavros A. Zenios (corresponding author)  
zenios.stavros@ucy.ac.cy  
Department of Accounting and Finance, Faculty of Economics and Management,  
University of Cyprus, P.O. Box 20537, Nicosia, Cyprus; Wharton Financial Institutions Center, University of Pennsylvania, Philadelphia, USA; Norwegian School of Economics.

---

**Contents**

1	Introduction . . . . .	3
1.1	Review of robust VaR and CVaR models . . . . .	4
2	Optimization of VaR and CVaR risk measures and their stability . . . . .	6
3	Robust VaR and CVaR under distribution and moment ambiguity . . . . .	9
3.1	Explicit formulation of RVaR and RCVaR optimization models . . . . .	12
3.2	Constructing the ambiguity set . . . . .	13
3.2.1	Algorithm for constructing a joint ellipsoidal ambiguity set . . . . .	15
3.2.2	Heuristic for constructing a joint ellipsoidal ambiguity set . . . . .	15
3.2.3	Comments on the choice of method . . . . .	15
3.3	Unifying and extending some results on RVaR and RCVaR optimization . . . . .	16
3.4	Extensions to polytopic and interval ambiguity sets . . . . .	17
4	Numerical tests . . . . .	17
4.1	Robustness of different investment strategies . . . . .	17
4.1.1	Buy-and-hold . . . . .	18
4.1.2	Active management . . . . .	18
4.2	Robustness under distribution ambiguity: moment mis-specification . . . . .	20
4.2.1	Mean and variance mis-specification . . . . .	23
4.2.2	Skewness and kurtosis mis-specification . . . . .	24
5	Conclusions . . . . .	27
	References . . . . .	28
A	Appendix: Proofs . . . . .	30
A.1	Proof of Theorem 2 . . . . .	30
A.2	Proof of Theorem 3 . . . . .	32
B	Appendix: Extension to polytopic ambiguity sets . . . . .	34
C	Appendix: Extension to interval ambiguity sets . . . . .	35
D	Appendix: Data . . . . .	38

## 1 Introduction

Mean-variance portfolio optimization from the seminal thesis of Harry Markowitz provided the basis for a descriptive theory of portfolio choice: how investors make decisions. This led to further research in financial economics, with the development of a theory on price formation for financial assets (by William Sharpe) and on corporate finance taxation, bankruptcy and dividend policies (by Merton Miller). These descriptive contributions of the behavior of financial agents was recognized by a joint Nobel Prize in 1990. The prescriptive part of the theory—how investors should make decisions—was also acclaimed by practitioners and mean-variance models proliferated. Here, however, problems surfaced: mean-variance portfolio optimization is sensitive to perturbations of input data (Best and Grauer, 1991; Chopra and Ziemba, 1993). Since the estimation of market parameters is error prone, the models are severely handicapped. In theory they produce well diversified portfolios but in practice they generate portfolios biased towards estimation errors.

With advances in financial engineering, variance was replaced by more sophisticated risk measures. We have seen value-at-risk (VaR) becoming an industry standard and written into the Basel II and III accords to calculate capital adequacy or calculate insurance premia or set margin requirements. However, value-at-risk is criticized for being non-convex and it is also computationally intractable to optimize. In a seminal paper Artzner et al. (1999) provided an axiomatic characterization of risk measures, which they call *coherent*, and conditional value-at-risk (CVaR) emerged as one such risk measure. CVaR rose to prominence with the work of Rockafellar and Uryasev (2000) who showed that it can be minimized as a linear program. CVaR optimization emerged as a credible successor to mean-variance models: it is coherent, computationally tractable and found numerous applications (Zenios, 2007, ch. 5).

How come then, that VaR and not CVaR became the industry standard? Serendipity played a role, of course. VaR was introduced first and promoted vigorously by J.P. Morgan, an industry leader. Furthermore, optimization of the risk measure is not required by the regulators and, hence, the computational challenge of VaR optimization passed unnoticed. But there is another, more fundamental reason, for preferring VaR over CVaR. VaR estimated from a set of sampled scenarios is a robust statistic, i.e., it is insensitive to small deviations of the underlying distribution from the observed distribution, whereas CVaR is not. Kou et al. (2013) argue that risk measures should be robust but coherent risk measures are not, so that CVaR lacks a key property.

In this paper we eliminate the sensitivity of CVaR by incorporating data ambiguity in the optimization model (Section 3). Kaut et al. (2007) exemplified that CVaR optimization models are sensitive to mis-specifications in means, standard deviation, skewness and, to a lesser extent, kurtosis. We show that these problems are alleviated (Section 4.2).

### 1.1 Review of robust VaR and CVaR models

Early suggestions in dealing with the sensitivity of portfolio optimization models to data estimation errors use Bayesian or James-Stein estimators, resampling, or restricting portfolio choices with ad hoc constraints. We do not review this literature as it is outside the scope of our work.

The 2008 global crisis revived the work of Chicago economist Frank H. Knight (1921) that considers financial and economic data as ambiguous instead of uncertain, whereby under uncertainty a probability model is known but the random variables are observed with some measurement error, whereas under ambiguity the probability model is unknown. Hence, data mis-specification is not only due to measurement error that can be reduced with improved estimation techniques. Data ambiguity is an integral part of financial decision making and deserves attention as an issue to be modeled, not a problem to be eliminated. It is from this perspective that we develop this study.

We build on recent research that brings developments in *robust optimization* to bear on portfolio selection under data ambiguity. Robust optimization models require constraints to be satisfied even with ambiguous data, and the objective value to be insensitive to the ambiguity. Concepts of robustness in optimization have been developed independently in the fields of operations research and engineering design. Mulvey et al. (1995) proposed the robust optimization of large scale systems when data take values from a discrete scenario set, using a regularization of the objective function to control its sensitivity and penalty functions to control constraint violation. This approach spurred numerous applications in facility location, power capacity planning, disaster response, supply chain management, production and process planning and so on. Robust convex optimization was developed in Ben-Tal and Nemirovski (1998) for optimization problems with data ambiguity described by an ellipsoid, and show that important convex optimization problems admit a tractable robust counterpart. The foundational papers spurred extensive theoretical and applied research (Ben-Tal et al., 2009; Bertsimas et al., 2011).

In a way robust portfolio optimization brings ideas from Taguchi robust engineering design to the design of portfolios. Authors usually adopt the robust convex optimization framework over an appropriate ambiguity set, and it is in this domain that our paper makes a contribution. Fabozzi et al. (2010) review robust portfolio optimization using mean, VaR and CVaR risk measures.

The first robust counterpart to mean-variance optimization was proposed by Goldfarb and Iyengar (2003). Using a linear factor model for asset returns they introduce “uncertainty structures” —the confidence regions associated with parameter estimation— and formulate robust portfolio selection models corresponding to these uncertainty structures as second order cone programs (SOCP). They also develop robust counterparts for VaR and CVaR optimization under the normality assumption of mean-variance models. Schottle and Werner (2009); Tütüncü and Koenig (2004) develop further robust mean-standard deviation and mean-variance models.

Our paper develops a robust counterpart of CVaR (RCVaR) optimization and finds it identical to robust VaR (RVaR) optimization. Hence, we give a detailed review of previous works on RVaR and RCVaR optimization so we can place our own contribution. But first, note an important distinction in terminology. VaR and CVaR *minimization* models lack a minimum return constraint, and ambiguity is restricted to the objective function. VaR and CVaR *optimization* trade off the risk measure against a minimum return target. These follow Markowitz's mean-variance tradition but are more difficult to analyze as ambiguity appears also in the constraints.

Current literature addresses the following problems relating to model parameters: (i) Ambiguity in mean return estimates, (ii) ambiguity in covariance matrix estimates, and (iii) ambiguity in the distribution of the data. Ambiguity can be independent for each parameter or joint for multiple parameters. If ambiguity is independent for each parameter we have simple sets constraining the parameters (e.g., a (sub)vector of parameters restricted to some intervals). For joint ambiguity we have sets such as ellipsoids or convex polytopes constraining the parameters. Models based on discrete scenarios may have distribution ambiguity in the scenario values or the scenario probabilities or both. For models with continuous distributions, ambiguity is in the moments.

Ghaoui et al. (2003) address RVaR minimization without the normality assumption. They consider partially known distributions of returns, whereby means and covariance lie within a known uncertainty set, such as an interval, a polytope (polytopic uncertainty), or a convex subset (convex moment uncertainty). Given this information on return distributions they cast RVaR minimization for interval uncertainty as semidefinite program (SDP), and for polytopic uncertainty as SOCP. They also give a general, but potentially intractable, model for convex moment uncertainty.

The first RCVaR optimization model is by Quaranta and Zaffaroni (2008) for interval uncertainty of the means. Zhu and Fukushima (2009) consider RCVaR minimization for box and ellipsoidal uncertainty in distribution, as well as distribution mixtures of convex combination of predetermined distributions. By "distribution" the authors mean the probabilities of the discretized data. By considering distribution ambiguity with known means, their RCVaR minimization model extends to RCVaR optimization simply by adding a non-ambiguous minimum return constrain.

Delage and Ye (2010) show (as a special case of their work) that RCVaR minimization for ambiguity in the probabilities, mean and second moment, can be solved in polynomial time. The authors provide bounds and generate confidence regions on the mean and covariance matrix in case of moment uncertainty but do not develop RCVaR models using this information. Instead, their robust model is for expected utility maximization under moment uncertainty.

Chen et al. (2011) point out that robust solutions come with a computational price: robust optimization models can be infinite dimensional and, without proper choice of uncertainty sets, the model may be intractable. They obtain bounds on worst case value of lower partial moments and use them to develop RVaR and RCVaR minimization under

distribution ambiguity with closed form solution under a normalization constraint. Paç and Pinar (2014) extend further RVaR and RCVaR optimization under distribution and mean returns ambiguity, but fixed covariance matrix.

A contribution that filled several gaps is Gotoh et al. (2013). Scenario based VaR and CVaR minimization models use discrete data observations (i.e., scenarios) and their probabilities to determine the empirical distribution. There are three possible ways to introduce ambiguity and formulate RVaR and RCVaR counterparts. The first approach (Zhu and Fukushima, 2009), keeps the scenarios fixed and considers ambiguous probabilities from a box or an ellipsoid. Gotoh et al. (2013) consider a second approach with uncertainty in scenarios but fixed probabilities, and a third approach, where both scenarios and probabilities are ambiguous.

Our work considers ambiguity in the distribution as well as mean returns, covariance matrix, and joint ambiguity in combinations of the above. These are, to the best of our knowledge, the most general ambiguity sets considered in the literature. Joint ambiguity provides a modeling capability not available in previous RVaR and RCVaR models (Schottle and Werner (2009) consider joint uncertainty in means and covariance matrix for mean-standard deviation models). We use an ellipsoidal ambiguity set which is quite general and obtain tractable optimization models as SOCP. We use the term ambiguity sets in the Knightean sense, instead of uncertainty sets in discussing robust models. Robust optimization literature typically refers to uncertainty sets although usually ambiguity is meant.

The paper is organized as follows. Section 2 defines VaR, CVaR, RVaR and RCVaR models and illustrates the instability of VaR and CVaR optimal portfolios. Section 3 is the main one. It formulates RVaR and RCVaR under distribution and ellipsoidal ambiguity in means and covariance, and establishes their equivalence (sec. 3.1), discusses the construction of ambiguity sets (sec. 3.2), extends or unifies existing results (sec. 3.3) and develops models for polytopic and interval ambiguity sets (sec. 3.4). We also identify some implicit assumptions made in previous works that limit their applicability to special cases and explain how we overcome the limitations. Section 4 reports on two distinct numerical tests. First, using historical CDS spread returns from eurozone crisis countries we investigate the robustness of alternative investment strategies. Second, using simulations we test the robustness of optimal portfolios under mis-specification of mean, variance, skewness and kurtosis. Proofs are gathered in Appendices.

## 2 Optimization of VaR and CVaR risk measures and their stability

The mean of  $\alpha$ -tail<sup>1</sup> distribution of portfolio loss  $X$ ,  $\text{CVaR}_\alpha(X)$ , and its minimization formula were developed in Rockafellar and Uryasev (2000):

---

<sup>1</sup> We use  $\alpha = 0.95$  whenever we do numerical experiments throughout the paper.

**Theorem 1 Fundamental minimization formula.**

As a function of  $\gamma \in \mathbb{R}$ , the auxiliary function

$$F_\alpha(X, \gamma) = \gamma + \frac{1}{1 - \alpha} \mathbb{E}\{[X - \gamma]^+\},$$

where  $\alpha \in (0, 1]$  is the confidence level and  $[t]^+ = \max\{0, t\}$ , is finite and convex, with

$$\text{CVaR}_\alpha(X) = \min_{\gamma \in \mathbb{R}} F_\alpha(X, \gamma).$$

Moreover, the set  $M_\alpha$  of minimizers to  $F_\alpha(X, \gamma)$  is a compact interval,  $M_\alpha = [x_\alpha, x^\alpha]$ , where  $x_\alpha = \inf\{x \in \mathbb{R} : P[X \leq x] \geq \alpha\}$  and  $x^\alpha = \inf\{x \in \mathbb{R} : P[X \leq x] > \alpha\}$ .

*Remark 1* Note that  $x_\alpha$ , the left end-point of the set  $M_\alpha$ , and not every minimizer of  $F_\alpha(X, \gamma)$ , is equal to  $\text{VaR}_\alpha(X)$ . Hence, the statement  $\text{VaR}_\alpha(X) = \underset{\gamma \in \mathbb{R}}{\text{argmin}} F_\alpha(X, \gamma)$  is true only when the minimum is unique and the interval reduces to a point.

Consider an investor operating in a market with  $n$  risky assets, a riskless asset and no short-selling. The riskless asset has rate of return  $r_f$  and the  $n$  risky assets have rates of return denoted by random vector  $\xi$ . The loss function associated with decision variable  $x \in \mathbb{R}^n$  of proportionate allocations to the risky assets is given by

$$f(x, \xi) = -(x^\top \xi + r_f(1 - x^\top e)),$$

where  $e$  is an  $n$ -vector of ones. (When dealing with portfolio optimization models, loss is a function of the portfolio  $x$  and we write the auxiliary function, and CVaR, as functions of  $x$ .) According to Theorem 1 the conditional value-at-risk of the loss function is the solution of

$$\text{CVaR}_\alpha(x) = \min_{\gamma \in \mathbb{R}} F_\alpha(x, \gamma), \quad (1)$$

where

$$F_\alpha(x, \gamma) = \gamma + \frac{1}{1 - \alpha} \mathbb{E}\{[f(x, \xi) - \gamma]^+\}.$$

If  $\bar{\gamma}$  denotes  $\underset{\gamma}{\text{argmin}} F_\alpha(x, \gamma)$ , then, by Theorem 1,  $\text{VaR}_\alpha(x)$  is obtained from

$$\begin{aligned} \text{VaR}_\alpha(x) &= \min_{\gamma \in \mathbb{R}} \gamma \\ &\text{s.t. } F_\alpha(x, \gamma) \leq F_\alpha(x, \bar{\gamma}). \end{aligned} \quad (2)$$

Generally, this problem is non-convex. But if  $F_\alpha(x, \gamma)$  has a unique minimum —e.g., for normally distributed returns with strictly increasing cumulative distribution function— it is convex.

The definition of VaR in (2) uses the auxiliary function  $F_\alpha(x, \gamma)$ , whereas the original formulation of VaR is

$$\begin{aligned} \text{VaR}_\alpha(x) &= \min_{\gamma \in \mathbb{R}} \gamma & (3) \\ \text{s.t. } & \text{Prob}\{\gamma \leq f(x, \xi)\} \leq 1 - \alpha. \end{aligned}$$

Hence, models for selecting a portfolio with minimal VaR or CVaR and a minimum return constraint can be posed as follows:

### I. VaR optimization

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \quad & \gamma & (4) \\ \text{s.t. } \quad & F_\alpha(x, \gamma) \leq F_\alpha(x, \bar{\gamma}), \\ & (\bar{\mu} - r_f e)^\top x \geq d - r_f, \end{aligned}$$

or

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \quad & \gamma & (5) \\ \text{s.t. } \quad & \text{Prob}\{\gamma \leq f(x, \xi)\} \leq 1 - \alpha, \\ & (\bar{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned}$$

### II. CVaR optimization

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \quad & F_\alpha(x, \gamma) & (6) \\ \text{s.t. } \quad & (\bar{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned}$$

$\bar{\mu}$  is a vector of the means of risky assets and  $d \in \mathbb{R}^+$  is the minimum return satisfying  $d \geq r_f$ .

It is well-known that scenario based CVaR is not a robust estimator whereas VaR is, but less is known about the stability of the portfolio weights obtained from minimizing either measure. We illustrate with a simple example the instability of minimum VaR and CVaR portfolios. We consider VaR and CVaR minimization without short-selling, a budget constraint and no risk-free asset. We assume confidence level  $\alpha \in (0, 1]$ , denote by  $\mathbb{X}$  the set  $\{x \in \mathbb{R}^n \mid x \geq 0, \sum_{i=1}^n x_i = 1\}$ , and use Monte Carlo simulation to generate an  $S \times n$  matrix  $R$  of return scenarios for  $n$  risky assets. VaR model (5) is solved as the mixed integer linear program

$$\begin{aligned} \min_{x \in \mathbb{X}, \gamma \in \mathbb{R}, y \in \{0,1\}^S} \quad & \gamma & (7) \\ \text{s.t. } \quad & & \\ & -Rx - My - e\gamma \leq 0, \\ & e^\top y \leq S(1 - \alpha), \end{aligned}$$



where  $M$  is a large positive number. CVaR model (6) is formulated as the linear program

$$\begin{aligned} \min_{x \in \mathbb{X}, u \in \mathbb{R}^S, \gamma \in \mathbb{R}} \quad & \gamma + \frac{1}{S(1-\alpha)} e^\top u \\ \text{s.t.} \quad & -Rx - e\gamma \leq u, \\ & u \geq 0. \end{aligned} \tag{8}$$

We consider two risky assets with independently distributed returns, with mean 10%, variance 21%, kurtosis 3, and skewness -0.06 and -0.26, respectively, and perform a rolling horizon simulation. We generate a time series of 240 asset returns by sampling from the respective distributions, see Figure 1(a)–(b), and use the first 120 returns to compute the corresponding optimal portfolio weights as well as the optimal VaR and CVaR. Then we repeat the procedure by rolling forward the estimation window by one period, and repeat 120 times until we reach the end of the time series. For benchmark comparison we run VaR and CVaR models over the whole time series and plot the optimal weights and the risk values. The benchmark experiment is “in-sample” and the 120 rolling horizon experiments are “out-of-sample”. Figure 1(c)–(d) depicts the instability of optimal portfolio weights, with both VaR and CVaR optimal weights fluctuating significantly, and Figure 1(e)–(f) shows that the instability of optimal values. This instability is exemplified further later, in Figure 7, when we use market data from the eurozone crisis.

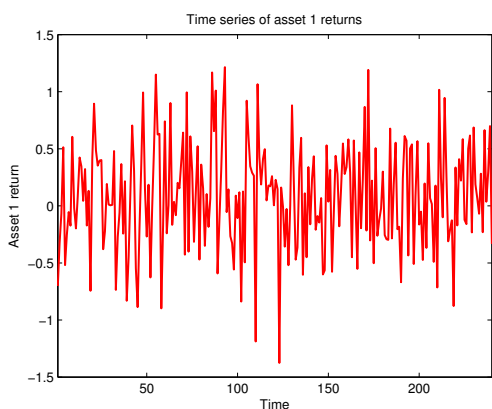
### 3 Robust VaR and CVaR under distribution and moment ambiguity

We introduce now ambiguity in the VaR and CVaR optimization models. The robust counterparts for both VaR and CVaR are formulated as SOCPs and we will observe that they are the same. We consider a joint ellipsoidal structure for the ambiguity set of mean returns and covariance matrix. Ellipsoidal sets can be viewed as generalizations of polytopic sets (Ben-Tal et al., 2009), and therefore our model generalizes Ghaoui et al. (2003) who developed RVaR minimization models under polytopic uncertainty. Extending the models from sets with independence between means and covariance used in earlier works (Ghaoui et al., 2003; Goldfarb and Iyengar, 2003; Tütüncü and Koenig, 2004) to sets that capture these dependencies, we generate better diversified and less conservative portfolios as argued by Lu (2011) for mean-variance models.

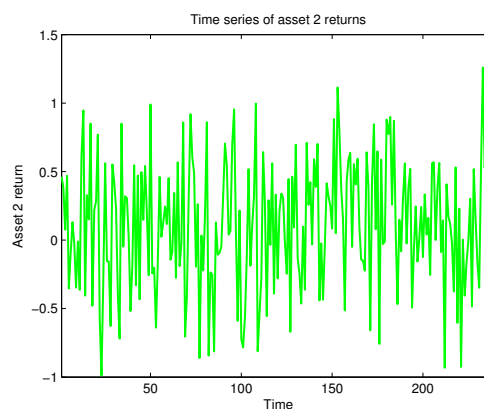
**Definition 1 (Ambiguity in distribution)** The random variable  $\xi$  assumes a distribution from

$$\mathbb{D} = \{\pi \mid \mathbb{E}_\pi[X] = \bar{\mu}, \text{Cov}_\pi[X] = \bar{\Gamma} \succ 0\},$$

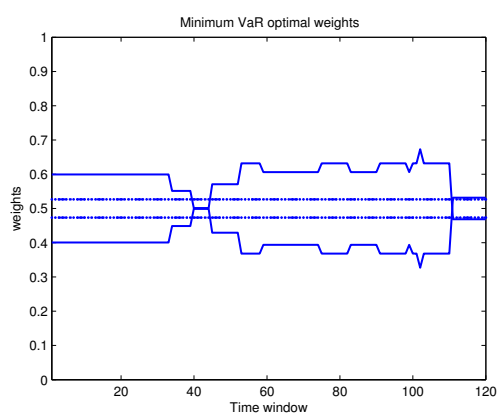
where  $\bar{\mu}$  and  $\bar{\Gamma}$  are given.



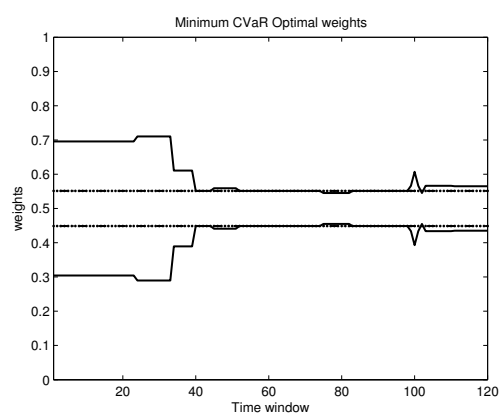
(a) Time series of asset 1 returns.



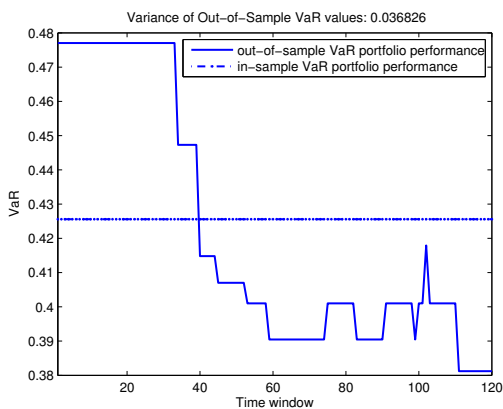
(b) Time series of asset 2 returns.



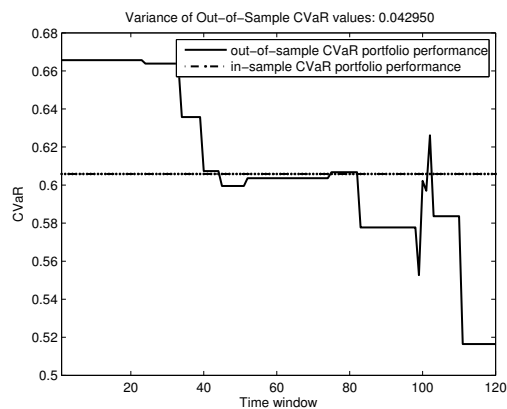
(c) Minimum VaR portfolio weights.



(d) Minimum CVaR portfolio weights.



(e) Minimum VaR values.



(f) Minimum CVaR values.

Fig. 1: A simple example of asset returns, portfolio weights and corresponding risk values.

**Definition 2 (Ellipsoidal ambiguity for mean returns and covariance matrix)**

Mean returns and covariance matrix belong to the joint ellipsoidal set:

$$U_\delta(\hat{\mu}, \hat{\Gamma}) = \{(\bar{\mu}, \bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid S(\bar{\mu} - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu} - \hat{\mu}) + \frac{S-1}{2} \|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma} - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2 \leq \delta^2\},$$

where  $\|A\|_{tr}^2 = tr(AA^\top)$ . (We use the definition of (Schottle and Werner, 2009, Proposition 3.3).)

*Remark 2* The set  $U_\delta(\hat{\mu}, \hat{\Gamma})$  is a generalization of ambiguity sets that have been used in the literature. Setting  $\bar{\Gamma} = \hat{\Gamma}$ , i.e., certainty about the estimate of the covariance matrix, we obtain the ellipsoidal set for mean returns used in Ceria and Stubbs (2006); Chen et al. (2011); Schottle and Werner (2009); Zhu et al. (2008). Similarly, if we fix  $\bar{\mu} = \hat{\mu}$ , i.e., certainty about the estimate of the means, we obtain the ellipsoidal set for covariance matrix. Goldfarb and Iyengar (2003) use this structure of uncertainty set for the factor loading matrix of a factor model of returns.

*Remark 3*  $U_\delta(\hat{\mu}, \hat{\Gamma})$  can be decomposed to  $U_{\sqrt{\kappa}\delta}(\hat{\mu})$  and  $U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma})$  using a parameter  $\kappa \in [0, 1]$ , with  $U_{\sqrt{\kappa}\delta}(\hat{\mu}) = \{\bar{\mu} \in \mathbb{R}^n \mid S(\bar{\mu} - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu} - \hat{\mu}) \leq \kappa\delta^2\}$  and  $U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma}) = \{\bar{\Gamma} \in \mathbb{S}^n \mid \frac{S-1}{2} \|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma} - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2 \leq (1-\kappa)\delta^2\}$ . This representation is useful later.

To develop RVaR and RCVaR models we start from one of the following:

1. RVaR<sub>I</sub> (robust counterpart of model (4))

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \quad & \gamma & (9) \\ \text{s.t.} \quad & \max_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} [F_\alpha(x, \gamma) - F_\alpha(x, \bar{\gamma})] \leq 0, \\ & \min_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} (\bar{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned}$$

2. RVaR<sub>II</sub> (robust counterpart of model (5))

$$\begin{aligned} \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \quad & \gamma & (10) \\ \text{s.t.} \quad & \max_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} \text{Prob}\{\gamma \leq f(x, \xi)\} \leq 1 - \alpha, \\ & \min_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} (\bar{\mu} - r_f e)^\top x \geq d - r_f, \end{aligned}$$

where  $\bar{\gamma} = \text{argmin}_\gamma F_\alpha(x, \gamma)$ .

3. RCVaR (robust counterpart of model (6))

$$\begin{aligned} \min_{x \in \mathbb{R}^n, \gamma \in \mathbb{R}} \quad & \max_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} F_\alpha(x, \gamma) & (11) \\ \text{s.t.} \quad & \min_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} (\bar{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned}$$

*Remark 4* The maximization problem in the first constraint of (9) can not, in general, be solved explicitly. Existing papers for RVaR minimization (Chen et al., 2011) and RVaR optimization (Paç and Pinar, 2014) solve a special case by assuming  $F_\alpha(x, \gamma)$  has a unique minimum<sup>2</sup>, thereby obtaining an explicit solution to the maximization problem in the constraint and simplifying the RVaR formulation.

*Remark 5* Unique minimum of  $F_\alpha(x, \gamma)$  implies unique solution of  $\text{Prob}\{f(x, \xi) \leq \gamma\} = \alpha$ , which is then  $\text{VaR}_\alpha(x)$ . This occurs when the distribution function of portfolio loss is strictly increasing. However, the loss distribution function is often a (non-decreasing) continuous step function, and Rockafellar and Uryasev (2002) extended their original contribution to derive the fundamental properties of CVaR for general loss distributions. We work with (10) to deal with the inner maximization for general loss distributions.

### 3.1 Explicit formulation of RVaR and RCVaR optimization models

We obtain now explicit formulations for models (10) and (11). First we prove an essential proposition and then the main theorem.

**Proposition 1** *If random variable  $\xi$  has a distribution from the set  $\mathbb{D}$  with fixed  $\bar{\mu}$  and  $\bar{\Gamma}$ , then*

$$\begin{aligned} \min_{\gamma \in \mathbb{R}} \max_{\pi \in \mathbb{D}} F_\alpha(x, \gamma) &= \min_{\gamma \in \mathbb{R}, x \in \mathbb{R}^n} \gamma \\ \text{s.t.} \quad \max_{\pi \in \mathbb{D}} \text{Prob}\{\gamma \leq f(x, \xi)\} &\leq 1 - \alpha, \end{aligned} \quad (12)$$

and both are equal to  $-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}$ .

**Proof.** From equation (10) in Paç and Pinar (2014) we know that

$$\min_{\gamma \in \mathbb{R}} \max_{\pi \in \mathbb{D}} F_\alpha(x, \gamma) = -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}. \quad (13)$$

For the constraint of the optimization problem in the right-hand side of (12), we use Theorem 1 of Ghaoui et al. (2003) (set  $1 - \alpha$  and  $f(x, \xi)$  instead of  $\epsilon$  and  $-r(w, x)$ , respectively), which means

$$\max_{\pi \in \mathbb{D}} \text{Prob}\{\gamma \leq f(x, \xi)\} \leq 1 - \alpha$$

---

<sup>2</sup> This implicit assumption is made in the proofs of Theorems 2.9 and 1, respectively, when the authors invoke the equality  $\text{VaR}_\alpha(X) = \underset{\gamma \in \mathbb{R}}{\text{argmin}} F_\alpha(X, \gamma)$  which holds true only when  $F_\alpha(x, \gamma)$  has a unique minimum, see Remark 1.

is equivalent to

$$-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \leq \gamma.$$

Hence, the optimization problem in the right-hand side is equivalent to

$$\begin{aligned} \min_{\gamma \in \mathbb{R}} \quad & \gamma \\ \text{s.t.} \quad & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \leq \gamma, \end{aligned} \tag{14}$$

which has the minimum value  $-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}$ . This completes the proof.

*Remark 6* The left and right-hand sides of (12) are RCVaR and RVaR, respectively, associated with the ambiguity set of Definition 1. Hence, the robust counterpart to VaR and CVaR optimization under distribution ambiguity is the same optimization model.

We obtain now RVaR and RCVaR optimization models for ambiguity in distributions, mean returns and covariance matrix.

**Theorem 2** *If random variable  $\xi$  has a distribution from the set  $\mathbb{D}$  and  $(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma})$ . Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) are both represented by the following SOCP:*

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & -r_f - (\hat{\mu} - r_f e)^\top x + \left( \max_{\kappa \in [0,1]} f(\kappa) \right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \\ \text{s.t.} \quad & -\frac{\delta}{\sqrt{S}} \|\hat{\Gamma}^{\frac{1}{2}} x\| + (\hat{\mu} - r_f e)^\top x \geq d - r_f, \end{aligned} \tag{15}$$

where  $f(k) = \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{(1 + \delta \sqrt{\frac{2(1-\kappa)}{S-1}}) + \delta \sqrt{\frac{\kappa}{S}}}$ .

**Proof.** See Appendix A.1.

*Remark 7*  $f(\kappa)$  is a strictly concave function with  $\lim_{\kappa \rightarrow 0} f'(\kappa) = \infty$  and  $\lim_{\kappa \rightarrow 1} f'(\kappa) = -\infty$ . Hence,  $f(\kappa)$  has a unique maximum in the interval  $(0, 1)$ .

This result is new and generalizes the result of Čerbáková (2006) for symmetric distributions identified by the first two moments. It was anticipated by Bertsimas et al. (2004) who obtain identical bounds on VaR and CVaR under distribution ambiguity.

### 3.2 Constructing the ambiguity set

An appropriate ambiguity set is typically taken as input in the robust optimization literature. Some times this is the uncertainty set corresponding to the confidence regions of

the statistical estimators of the model parameters (Goldfarb and Iyengar, 2003; Schottle and Werner, 2009). Other times —such as in the examples we solve later— we may be given multiple estimates of model parameters and then the question is raised as to what is the appropriate ambiguity set. In finance it is not uncommon to be given estimates by multiple securities analysts. In such cases we need a method to construct an ambiguity set, including its center. We propose and solve analytically a nonlinear SDP for finding the center of a joint ellipsoidal set.

Assume  $K$  experts provide estimates for mean returns and covariance matrices  $(\bar{\mu}_k, \bar{\Gamma}_k)$ ,  $k = 1, 2, \dots, K$ . (For convenience we assume they were all estimated using the same number of scenarios  $S$ .) To construct their joint ellipsoidal ambiguity set we need to fix the center  $(\hat{\mu}, \hat{\Gamma})$ . This is obtained as the solution of a nonlinear convex program for minimizing the  $l_2$ -norm of the parameters  $\delta_k$ , for  $k = 1, 2, \dots, K$ , where each parameter corresponds to the ellipsoid with center  $(\hat{\mu}, \hat{\Gamma})$  containing observation  $(\bar{\mu}_k, \bar{\Gamma}_k)$ . Referring to Definition 2 the optimization problem is given by

$$\min_{\hat{\mu} \in \mathbb{R}^n, \hat{\Gamma} \in \mathbb{S}_{++}^n} \sqrt{\sum_{k=1}^K S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2} \|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2}, \quad (16)$$

where  $\mathbb{S}_{++}^n$  is the set of all  $n$ -dimensional, symmetric, positive definite matrices. This problem is equivalent to

$$\min_{\hat{\mu} \in \mathbb{R}^n, \hat{\Gamma} \in \mathbb{S}_{++}^n} \sum_{k=1}^K S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2} \|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2. \quad (17)$$

The next theorem give the solution of this problem, if a solution exists.

**Theorem 3** *If (17) is solvable, then it admits the following solution:*

1.  $\hat{\mu} = \frac{1}{K} \sum_{k=1}^K \bar{\mu}_k$ .
2.  $\Gamma^-$ , the inverse of optimal  $\hat{\Gamma}$ , is obtained from the linear system of equations

$$\left[ \sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k \right] \text{vec}(\Gamma^-) = \sum_{k=1}^K \text{vec}(\bar{\Gamma}_k) - \frac{S}{(S-1)} \sum_{k=1}^K \text{vec} \left( (\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top \right), \quad (18)$$

where  $\otimes$  is the Kronecker product and  $\sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k$  is positive semidefinite.

If at least one of  $\bar{\Gamma}_k$ ,  $k = 1, \dots, K$ , is positive definite, then (18) has a unique solution.

**Proof.** See Appendix A.2.

We now state a simple algorithm for constructing ellipsoidal ambiguity sets.

### 3.2.1 Algorithm for constructing a joint ellipsoidal ambiguity set

1. Compute  $\hat{\mu} = \frac{1}{K} \sum_{k=1}^K \bar{\mu}_k$ , solve the system of linear equations (18) for  $\text{vec}(\Gamma^-)$ , form matrix  $\Gamma^-$ , and calculate its inverse  $\hat{\Gamma}$  to obtain the center  $(\hat{\mu}, \hat{\Gamma})$ .
2. Choose  $\delta$  such that the resulting ellipsoidal set inscribes  $(\bar{\mu}_k, \bar{\Gamma}_k)$ ,  $k = 1, \dots, K$ , i.e., compute the distance of each estimate from the center,  $\delta_1, \dots, \delta_K$ , and let  $\delta$  be the maximum value.

System (18) is of dimension  $n^2 \times n^2$  which depends only on the number of assets  $n$ . Methods for solving systems of equations based on LU or Cholesky (when applicable) factorizations are polynomial of cubic order. Hence, the computational complexity of the algorithm is  $O(n^6)$ , and for medium portfolio sizes this is tractable.

### 3.2.2 Heuristic for constructing a joint ellipsoidal ambiguity set

It is also possible to construct the ambiguity set with a simple heuristic. In some cases the heuristic gives tighter ellipsoids than the algorithm and this results to less conservative robust solutions. The heuristic needs  $K$  inversions of a matrix of dimension  $n \times n$  and elementary matrix operations, and its computational complexity is  $O(Kn^3)$ .

1. For each estimate  $(\bar{\mu}_k, \bar{\Gamma}_k)$  we compute the sum of its distances from all others

$$\text{dist}_k = \sqrt{\sum_{k'=1}^K S(\bar{\mu}_{k'} - \bar{\mu}_k)^\top \bar{\Gamma}_k^{-1} (\bar{\mu}_{k'} - \bar{\mu}_k) + \frac{S-1}{2} \|\bar{\Gamma}_k^{-\frac{1}{2}} (\bar{\Gamma}_{k'} - \bar{\Gamma}_k) \bar{\Gamma}_k^{-\frac{1}{2}}\|_{tr}^2}.$$

2. The estimate with the minimum value of  $\text{dist}_k$  is the center, and we choose  $\delta$  such that the constructed ellipsoidal set inscribes all  $(\bar{\mu}_k, \bar{\Gamma}_k)$ ,  $k = 1, \dots, K$ . Specifically, we compute  $\delta_1, \dots, \delta_K$  as the distance of each point in the ellipsoid from the center and let  $\delta$  be the maximum value.

### 3.2.3 Comments on the choice of method

It remains an open question how to construct an ellipsoidal ambiguity set that is big enough to guarantee robustness but tight enough to avoid conservative solutions. One may wish to try both the algorithm and the heuristic and pick the ellipsoid with the smaller  $\delta$ , knowing that both ellipsoidal sets ensure robust solutions and the one with the smallest  $\delta$  is the less conservative. Furthermore, both methods provide an intuitive way to choose smaller values of  $\delta$  by choosing a suitable quantile of  $\delta_k$ ,  $k = 1, \dots, K$ . It is not clear on the outset which method generates the tighter ellipsoid. Figure 2 illustrates two situations when one method dominates the other. When observations are evolving slowly, or differ slightly from each other, the heuristic performs better since one of the observations provides a good center. When observations change significantly then the algorithm is better in finding a center of the diverse observations.

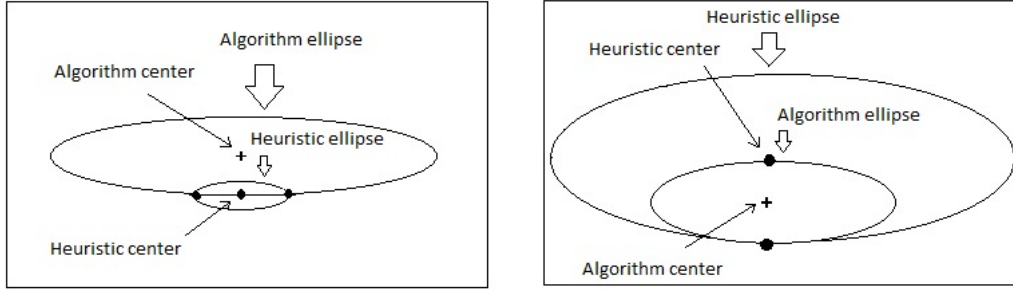


Fig. 2: Illustrating algorithm- and heuristic-constructed ellipsoids. For the observations (solid bullets) on the left, with small changes from each other, the heuristic estimates the tighter ellipsoid. For the observations on the right, with larger differences, the algorithm ellipsoid is tighter.

### 3.3 Unifying and extending some results on RVaR and RCVaR optimization

From our model we obtain, as special cases, known results from the literature.

1. For distribution ambiguity with known mean returns and covariance matrix, set  $\bar{\mu} = \hat{\mu}$ ,  $\bar{\Gamma} = \hat{\Gamma}$  in Proposition 1, to get the results of Chen et al. (2011); Paç and Pinar (2014) for RCVaR, and of Ghaoui et al. (2003) for RVaR.
2. For ambiguity in distribution and means, and known covariance, set  $\kappa = 1$  in Theorem 2 to get

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & -r_f - (\hat{\mu} - r_f e)^\top x + \left( \frac{\delta}{\sqrt{S}} + \sqrt{\frac{\alpha}{1-\alpha}} \right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \\ \text{s.t.} & -\frac{\delta}{\sqrt{S}} \|\hat{\Gamma}^{\frac{1}{2}} x\| + (\hat{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned} \quad (19)$$

This is Paç and Pinar (2014) RCVaR model, with their constant  $\epsilon = \frac{\delta}{\sqrt{S}}$ .

3. For ambiguity in distribution and covariance, and known means, set  $\kappa = 0$  in Theorem 2 to get

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & -r_f - (\hat{\mu} - r_f e)^\top x + \left( \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{1 + \delta \sqrt{\frac{2}{S-1}}} \right) \|\hat{\Gamma}^{\frac{1}{2}} x\| \\ \text{s.t.} & (\hat{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned} \quad (20)$$

We are not aware of any studies of this case, which is of interest in risk minimization models or when there is special knowledge on the mean return. For instance, in index



tracking (Zenios, 2007, ch. 7) the mean excess return of a portfolio over the index is zero in efficient markets.

Using Theorem 2 we can relax the assumption of Chen et al. (2011); Paç and Pinar (2014), see Remark 4. Their RVaR model is

$$-r_f - (\bar{\mu} - r_f e)^\top x + \frac{2\alpha - 1}{2\sqrt{\alpha}\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x},$$

while we and Ghaoui et al. (2003) have

$$-r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}.$$

From Remark 5 we know that their RVaR model is computed over a subset of the original ambiguity set, i.e., the set of all *strictly increasing* distribution functions with fixed means and covariance, and their parameter differs from ours. Obviously  $\frac{2\alpha-1}{2\sqrt{\alpha}\sqrt{1-\alpha}} < \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}}$  and their robust counterpart is less conservative but is valid only under the assumption.

### 3.4 Extensions to polytopic and interval ambiguity sets

It is possible to extend our work to models under distribution ambiguity and polytopic and interval ambiguity sets in the mean returns and covariance matrix. Polytopic and interval uncertainty was studied by Ghaoui et al. (2003) for RVaR minimization. The extensions contribute models for RVaR optimization, and, by Proposition 1, new RCVaR optimization models for these two ambiguity sets. The extensions are given in Appendices B and C.

## 4 Numerical tests

We illustrate the performance of the robust models and compare the robust models vis-a-vis the non-robust (inomial) models. First, we use historical CDS spread returns from eurozone crisis countries to test the robustness of buy-and-hold and active management investment strategies obtained using the models. Second, we use simulations to test the robustness of solutions under mis-specification of mean, variance, skewness and kurtosis of the return distributions. All computations were performed using MATLAB 7.14.0 on a Core i7 CPU 2.5GHz laptop with 8GB of RAM. SOCPs are solved using CVX and (mixed integer) linear programs with CPLEX.

### 4.1 Robustness of different investment strategies

We consider portfolios trading in CDS of Portugal, Slovenia, Italy, Spain, Ireland, Germany, Cyprus and Greece using daily spread returns from 2 Feb. 2009 to 16 Sept. 2011.

This period covers the eurozone crisis. Analyzing Greece CDS spreads using Bai-Perron tests we note regime switching at 20 April 2010 and 15 April 2011 (see Appendix D). Up to April 2010 there is a tranquil period (days 1–317), until April 2011 a turbulent period (days 318–575), and post April 2011 the crisis (days 576–685). This classification is convenient to stress-test the robustness of the model as the market changes from tranquil to turbulent and into a crisis. We consider RVaR and RCVaR minimization without a risk free asset, no short-selling and a budget constraint, following different investment management strategies.

#### 4.1.1 Buy-and-hold

Buy-and-hold investors use the available information to set up the model and obtain an asset allocation which is held throughout the investment horizon. Consider an investor who develops robust models based on the scenarios observed during the tranquil period. Subsequently, as the markets move into turbulence and new information is observed, it is used to compute out-of-sample portfolio performance, but the portfolio is not re-optimized. For each new observation we drop the oldest observation, so that VaR and CVaR are computed on a constant size window of recent data.

The ellipsoidal ambiguity set is constructed as follows. First we estimate  $(\bar{\mu}, \bar{\Gamma})$  using the first 150 (out of 316) return observations in the tranquil period. Then we discard the first observation, add the 151st, and compute a new estimate of  $(\bar{\mu}, \bar{\Gamma})$ . This procedure is repeated by “rolling” the estimation window forward one period at a time until the end of the tranquil period. At the end of this procedure, we have 166 estimates of  $(\bar{\mu}, \bar{\Gamma})$ , and compute the center and  $\delta$  of the ambiguity set using the algorithm of subsection 3.2.1. Reference (non robust) models use observed data over the tranquil period to obtain minimum VaR and CVaR portfolios, which are held throughout the turbulent period. We repeat the process using scenarios in the turbulent period and evaluate out-of-sample performance into the crisis period.

Results are shown in Figures 3–4. We observe that out-of-sample risk measure for the reference portfolios may be larger than the in-sample value, but not so for the the robust portfolios. Also, as a result of incorporating distribution ambiguity in the model, robust portfolios remain robust even when there is a regime switch in market data. A shortcoming of robust portfolios is that they are too conservative, as observed in the big gap between the in-sample and out-of-sample values.

#### 4.1.2 Active management

Active portfolio managers use the available information to set up the model and optimize asset allocation for one time period, but as new information arrives the data estimates are updated and the portfolio is re-optimized. We consider an investor who starts calibrating reference (non-robust) and robust models starting with the scenarios from the first 150 observations of the tranquil period, with  $\delta = 0$  for the starting robust model. Subsequently,

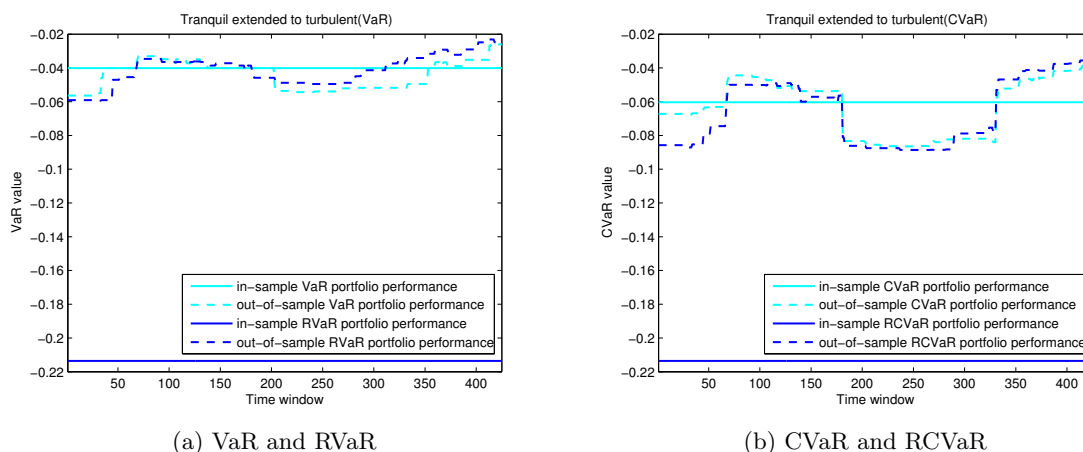


Fig. 3: Out-of-sample performance of buy-and-hold for tranquil-to-turbulent. ( $\delta = 20.24$ )

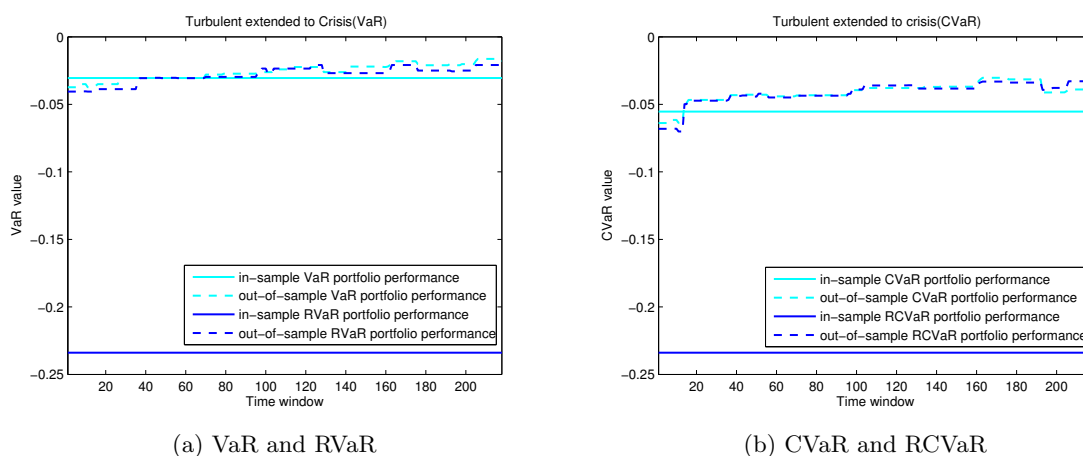


Fig. 4: Out-of-sample performance of buy-and-hold for turbulent-to-crisis. ( $\delta = 21.63$ )

a new data point is observed, the oldest observation is dropped and we compute the risk measures with the shifted window and the portfolio ex post return for the new data point. After the time window is shifted we re-optimize the asset allocation with the new information. For the robust model we use the new information to update  $\delta$  and construct an ellipsoid using the algorithm of subsection 3.2.1. This procedure is repeated until the end of the turbulent period. The same experiment is carried out starting with the first 150 observations of the turbulent period and finishing at the end of the crisis.

Results are reported in Figures 5–6. Panels (a) and (b) show the difference between in- and out-of-sample risk measures. The investor is on the safe side when the difference is positive, but suffers losses beyond expectation for negative differences. Figure 5(a)-(b)

shows out-of-sample performance occasionally deviating from the in-sample estimate. As the time window rolls forward the robust model registers few and minor downside violations, as a result of enlarged ambiguity sets with increasing  $\delta$  (Figure 8). This improvement is less pronounced in Figure 6(a)-(b), since spreads change substantially during the crisis and learning is insufficient to build an ellipsoid containing crisis movements. Robust models can not be better than the data defining the ambiguity sets.

Figures 5(c) and 6(c) plot the ex-post cumulative growth of a 100 unit investment using both the robust and non-robust models. We calculate the Sharpe ratios for the returns of portfolios developed using VaR, CVaR and their robust counterpart. We take the German 3-month treasury bill rate as the risk free in Sharpe ratio calculations, and the results are reported in the figure. We tested the hypothesis that the Sharpe ratios are identical between the robust and non robust strategies using the test of Wright et al. (2014), and could not reject it at the 0.95 level, for both the tranquil-to-turbulent and turbulent-to-crisis periods. Robust solutions do not pay a price in portfolio performance.

We also revisit the instability issue demonstrated in Section 2. Figure 7 illustrates the portfolio composition for the tranquil-to-turbulent period. The robust portfolios change gradually but not so the non-robust counterparts. Portfolio turnover of the robust model is 0.004, an order of magnitude smaller than that of VaR models (0.09) and CVaR (0.03).

Figure 8 shows the values of  $\delta$  obtained with the algorithm and the heuristic. The algorithm generates tighter ellipsoids for tranquil-to-turbulent period, while there is no clear advantage of one method over the other in turbulent-to-crisis period. In all experiments reported above we use ellipsoids constructed by the algorithm. We also performed experiments using heuristic-constructed ellipsoids, without any significant differences.

#### 4.2 Robustness under distribution ambiguity: moment mis-specification

We demonstrate the robustness of RCVaR optimal portfolios to mis-specification in the first four marginal moments. Mis-specification of higher moments is a form of distribution ambiguity and these tests illustrate robustness with respect to distribution ambiguity. A more interesting interpretation of our results is in conjunction with the work of Kaut et al. (2007), where it was established that CVaR optimization models are sensitive to mis-specification of means, covariance and skewness, and less so to kurtosis. Our results show that these sensitivities are eliminated from RCVaR and, by Proposition 1, from RVaR too.

We consider CVaR and RCVaR optimization with a minimum return constraint, no short-selling and a budget constraint. We perturb one moment at a time while keeping all other moments fixed to their original (assumed “true”) value, and repeat the perturbation 100 times. Data are from Kaut et al. (2007), see Appendix D, Table 1, for an international investment portfolio. These data are assumed to be the true values of the moments.

We test as follows the impact of moment mis-specification on the models:

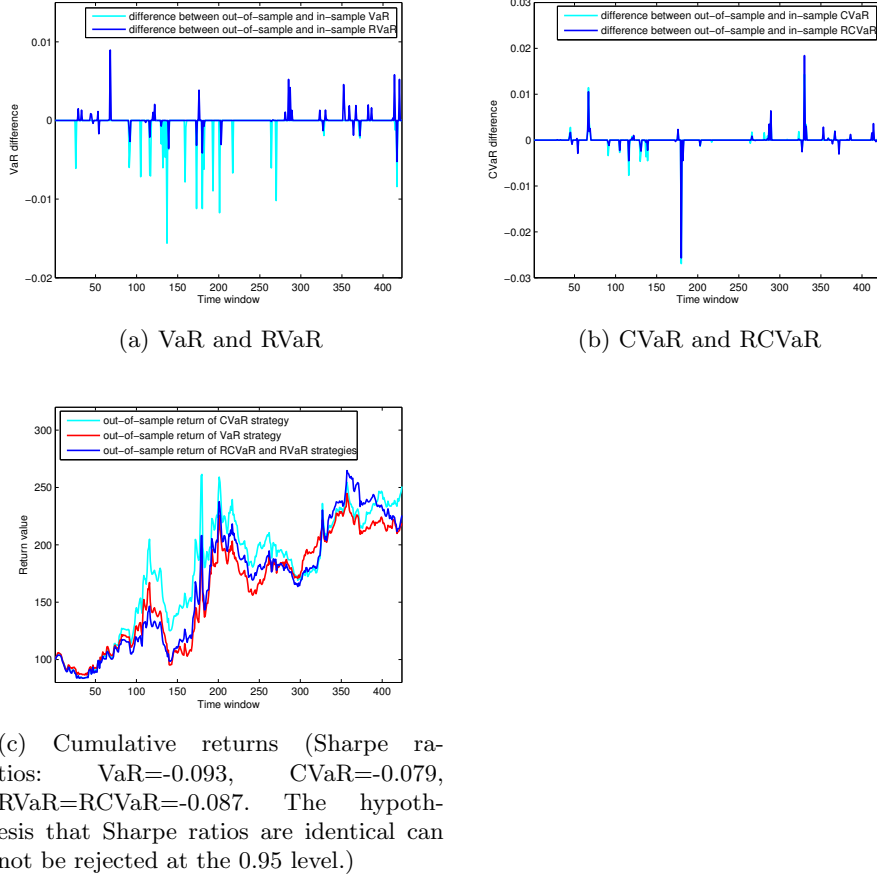


Fig. 5: Out-of-sample performance with active management for tranquil-to-turbulent.

**Step 0:** Fix parameter  $\theta$  and define  $\theta\%$  error on a moment as<sup>3</sup>

$$\text{true value} \left(1 + \epsilon \frac{\theta}{100}\right), \epsilon \in U[-1, 1].$$

Generate 100 perturbations for one moment by randomly generating  $\epsilon$ , while all other moments are fixed to their true values.

**Step 1:** Generate 2000 scenarios using Pearson random numbers with the specified mean, standard deviation, skewness or kurtosis for each one of the 100 perturbations from Step 0. Record the scenario sets  $\{\bar{R}_k\}_{k=1}^{100}$  and their means and covariance  $\{(\bar{\mu}_k, \bar{\Gamma}_m)\}_{k=1}^{100}$ .

<sup>3</sup> We follow Chopra and Ziemba (1993), except that they use normally distributed  $\epsilon \in N[0, 1]$ , while we use uniformly distributed  $\epsilon \in U[-1, 1]$ .

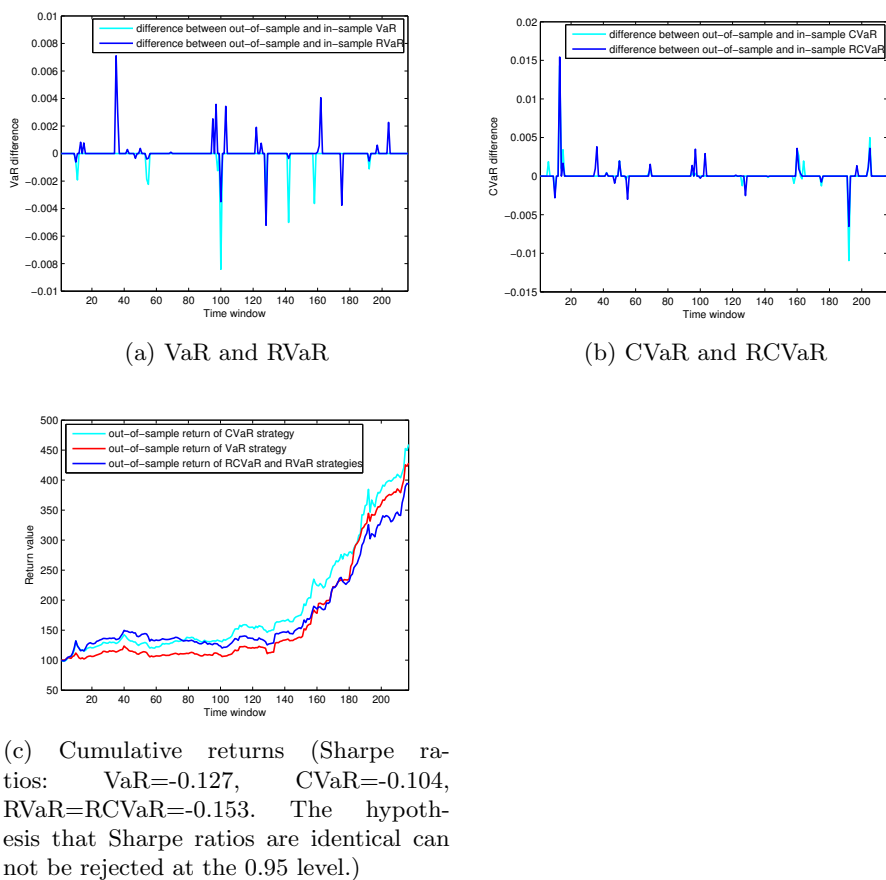


Fig. 6: Out-of-sample performance with active management for turbulent-to-crisis.

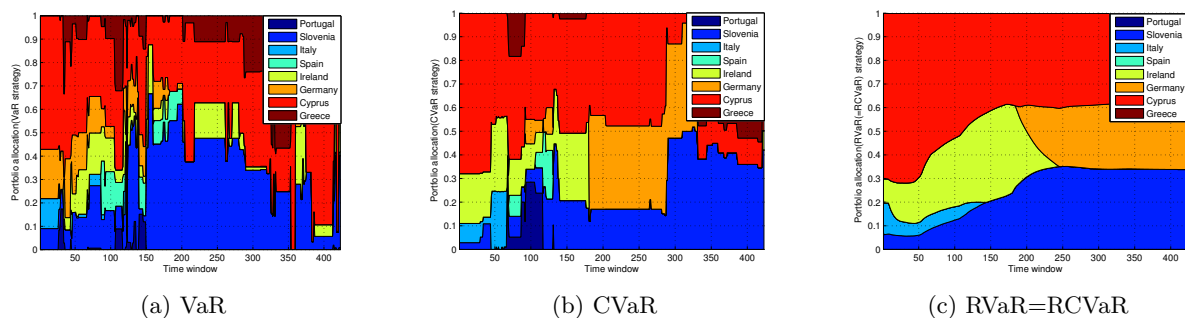


Fig. 7: Portfolio composition with different models for tranquil-to-turbulent. Portfolio turnover for VaR=0.09, CVaR=0.03 and RVaR=RCVaR=0.004.

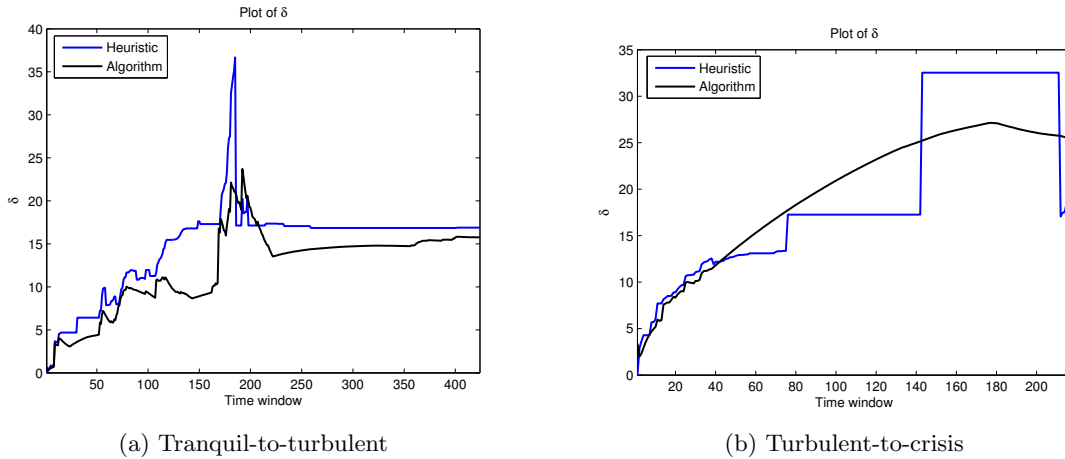


Fig. 8: Values of  $\delta$  estimated with the algorithm and the heuristic.

**Step 2:** Apply the algorithm of subsection 3.2.1 to  $\{(\bar{\mu}_k, \bar{\Gamma}_k)\}_{k=1}^{100}$ , to find the center  $(\hat{\mu}, \hat{\Gamma})$  and the parameter  $\delta$ . Chose the point with the smallest distance  $\delta_k$  from the center, with its scenario set  $\hat{R}$ , as the *reference* scenario set.

**Step 3:** Solve the model on the reference scenario set and the robust counterpart, and record the optimal portfolios.

**Step 4:** Compute return and risk measure of the optimal portfolios over  $\{\bar{R}_k\}_{k=1}^{100}$ .

The model on the reference set is a proxy for the nominal model since we do not have a scenario set corresponding to the center of the ellipsoid generated by the algorithm. When using the heuristic to compute the ellipsoid we have the scenario set for the center and hence we have exactly the nominal model. The performances of the proxy and the nominal models do not differ significantly and in the experiments we compare the robust model with the proxy.

#### 4.2.1 Mean and variance mis-specification

We have already illustrated using CDS data the robustness of the model with respect to ambiguity in distribution, means and covariance. Applying the simulation procedure outlined above for the first two moments we affirm the findings from the previous testing. However, the reason we perform simulations on the first two moments is to illustrate another feature of the model. Robust models are conservative and we illustrate how to control conservatism by adjusting  $\delta$ . We pick  $\delta$  so that the ambiguity set is large enough to contain all 100 perturbations, or only 90 or 80 and so on, by choosing in Step 2 of the algorithm the appropriate quantile of  $\delta_k$ ,  $k = 1, \dots, K$ .

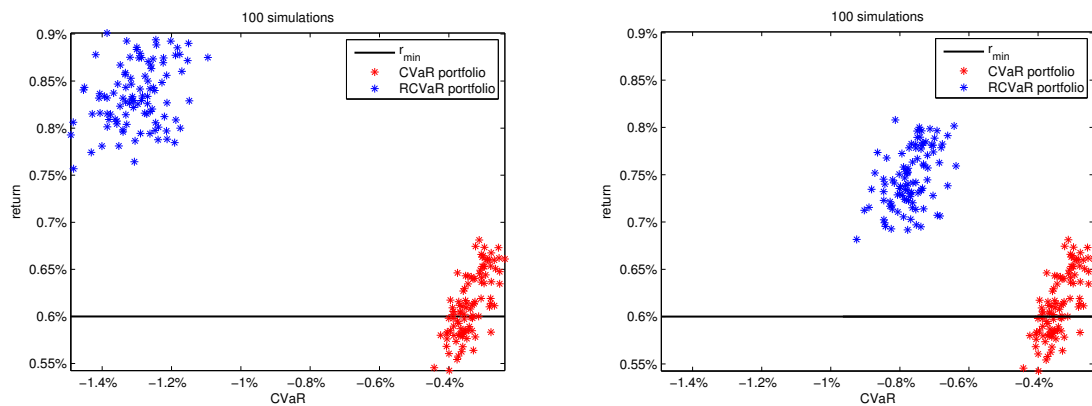
The results of simulating perturbations in the mean and variance by  $\theta = 10\%$  and  $20\%$ , respectively, and different values of  $\delta$ , are illustrated in Figures 9–10. We observe from panel (a) that the robust portfolio never violates the minimum return constraint when simulated using out-of-sample data, while the nominal portfolio violates the minimum return for about half of the perturbations. On the other hand, the robust optimal portfolio is conservative and mean return is significantly higher than the target for all perturbations. We can remedy this situation by reducing  $\delta$  using a suitable quantile of  $\delta_k, k = 1, \dots, K$ . Panel (b) illustrates the effect of this modification in obtaining less conservative portfolios that still satisfy the minimum return constraint for all perturbations. Running the models on the heuristic generated ellipsoids we find them more conservative but we do not report the results as they do not provide any additional insights.

#### 4.2.2 Skewness and kurtosis mis-specification

Kaut et al. (2007) established that CVaR portfolios are sensitive to mis-specification in the first four moments and correlations. In particular, they solve CVaR optimization models for each of 100 generated scenario sets and then evaluate all these optimal portfolios on the benchmark scenario set. Our experiment studies the same issue for CVaR and RCVaR optimization. We solve the CVaR optimization with reference scenario  $\hat{R}$  and its robust counterpart, and evaluate the optimal portfolio on 100 scenario sets of skewness and kurtosis perturbations.

The results are illustrated in Figures 11–12. We examine the performance of optimal portfolios with increasing perturbation parameter  $\theta$ , and  $\delta$  chosen by the algorithm of subsection 3.2.1. Our observations are consistent with Kaut et al. (2007) on the sensitivity of CVaR strategy with respect to higher moments mis-specification in the sense that the optimal CVaR portfolio violates the minimum return constraint for the perturbed scenario sets. The sensitivities of CVaR portfolios to errors in skewness and kurtosis are in agreement with the findings of Kaut et al. (2007). The optimal RCVaR portfolios, however, satisfy the minimum return constraint even with perturbations in the higher moments. Higher moment perturbation is a form of distribution ambiguity and, hence, the results are expected since the models are robust with respect to distribution ambiguity.

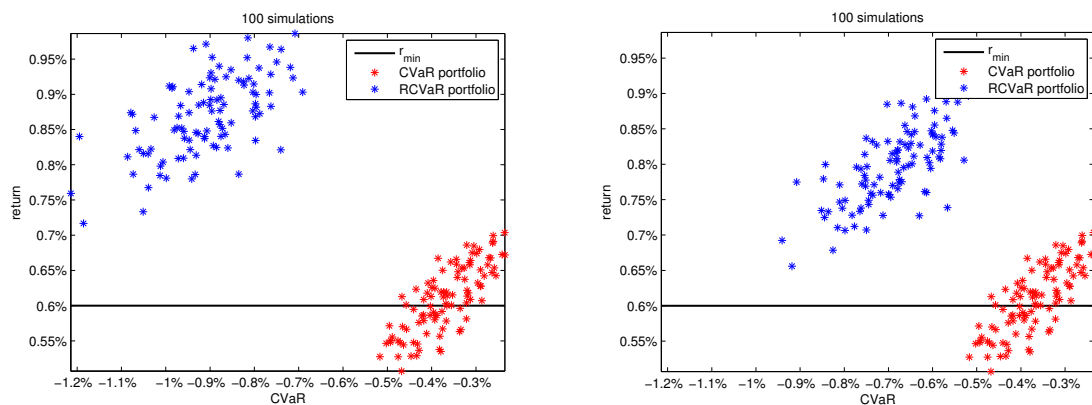




(a)  $\delta = 12.24$ , solution from subsection 3.2.1 algorithm

(b)  $\delta = 10.61$ , 80th quantile of  $\delta_k, k = 1, \dots, K$

Fig. 9: Out-of-sample performance of RCVaR optimization for perturbations in means and variances with  $\theta = 10\%$ . Portfolios are less conservative for smaller delta but satisfy the minimum return.



(a)  $\delta = 21.13$ , solution from subsection 3.2.1 algorithm

(b)  $\delta = 16.94$ , 80th quantile of  $\delta_k, k = 1, \dots, K$

Fig. 10: Out-of-sample performance of RCVaR optimization for perturbations in means and variances with  $\theta = 20\%$ . Portfolios are less conservative for smaller delta but satisfy the minimum return.

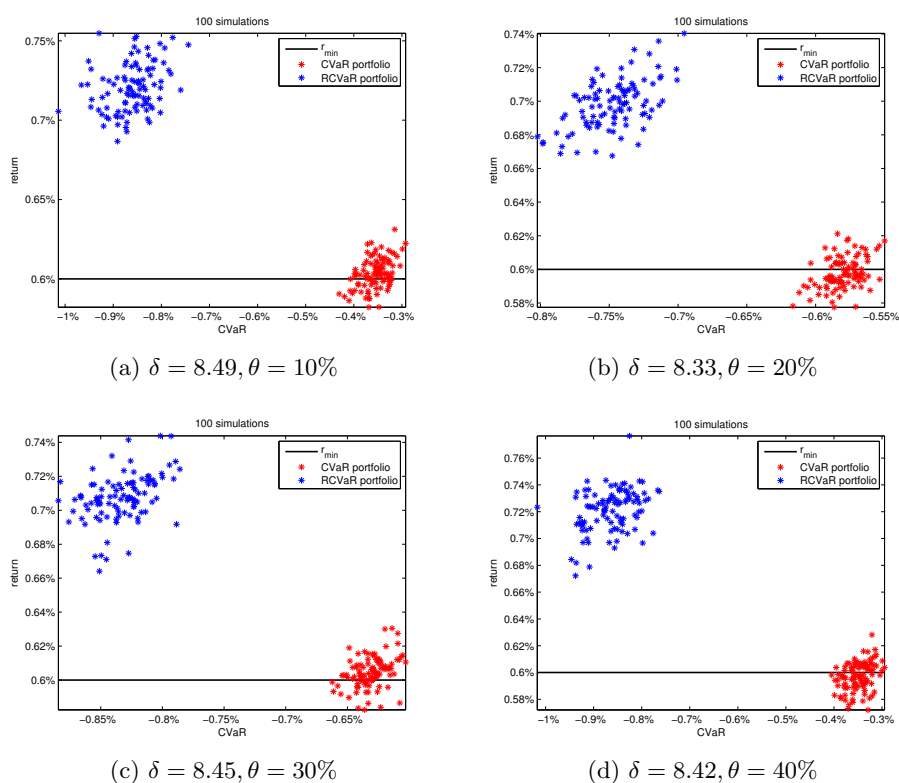


Fig. 11: Out-of-sample performance for perturbations in skewness for increasing  $\theta$ . RCVaR portfolios meet the minimum return whereas CVaR portfolios do not.

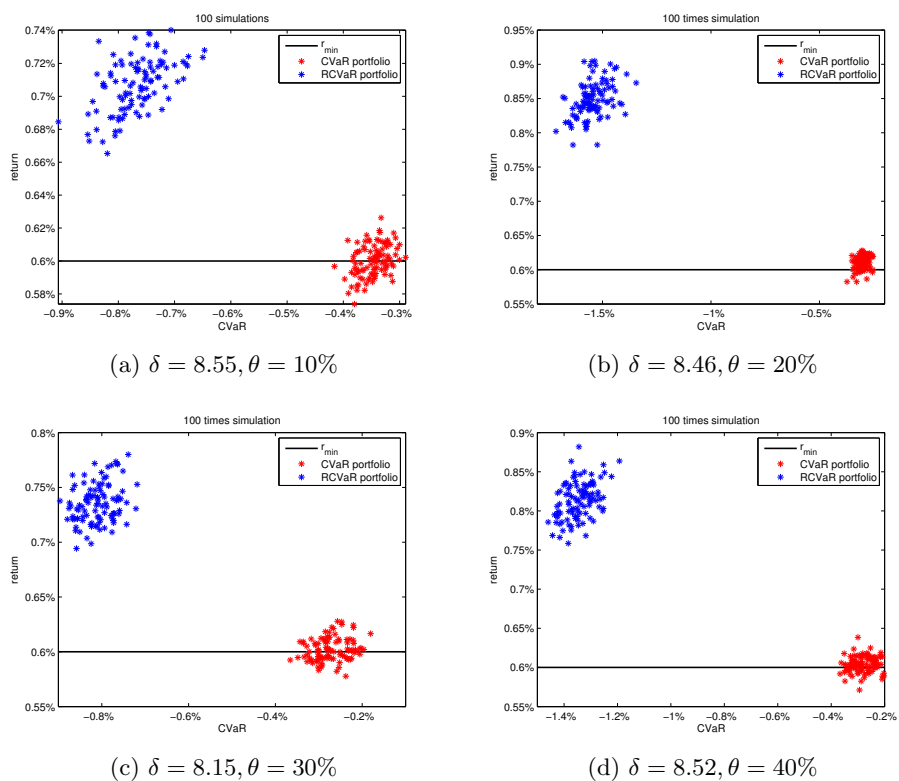


Fig. 12: Out-of-sample performance for perturbations in kurtosis for increasing  $\theta$ . RCVaR portfolios meet the minimum return whereas CVaR portfolios do not.

## 5 Conclusions

This paper develops models for robust optimization of VaR and CVaR under the most general ambiguity sets known so far, namely joint ambiguity in the distribution, mean returns and covariance matrix. RVaR and RCVaR optimization under distribution ambiguity reduce to the same second order cone program. This result allows us to develop several tractable models using ellipsoidal, polytopic and interval ambiguity sets for mean returns and covariance matrix. These models expand the arsenal of robust optimization tools for risk management

The paper also suggests an algorithm and a heuristic to construct ellipsoidal ambiguity sets from a set of point estimates. We also show how to control the size of the ellipsoid, thus limiting the well known conservatism of robust optimization models.

Numerical results support the following conclusions:

1. Buy-and-hold investment strategies based on robust optimization models perform well even out-of-sample and under extreme market movements of a financial crisis. Active investment strategies based on robust models perform better out-of-sample than the non-robust counterparts, but they are still sensitive to extreme market movements. This finding is tentative since it is based on limited experimentation. The claim supported from this experiment is that strategies based on robust models perform better than strategies based on non-robust models when applied out-of-sample. However, there is no assurance that they remain robust out-of-sample.
2. RVaR and RCVaR models produce conservative solutions. However, both the algorithm and the heuristic for constructing an ellipsoidal ambiguity set provide a way to select the ellipsoidal parameter in order to control conservatism.
3. RCVaR optimal portfolios are robust with respect to mis-specifications in the first four moments.

Significant progress has been made to establish RCVaR optimization as a viable —i.e., computationally tractable and theoretically sound— successor to the classical Markowitz mean-variance model. However, a problem now emerges. In particular while CVaR is a typical example of coherent risk measures, neither RVaR nor RCVaR are coherent. RVaR and RCVaR satisfy the sub-additivity axiom (an important improvement of RVaR over VaR), however, both RVaR and RCVaR violate the monotonicity axiom. Zhu and Fukushima (2009) argued that RCVaR is coherent, but their robust counterpart is for the risk measure associated with random variables  $X$  and  $Y$  over the *same* uncertainty set. We envision a more general scheme whereby the robust counterpart of a risk measure associated with  $X$  is to be taken over an ambiguity set associated with this random variable, which might differ from the ambiguity set of  $Y$ . It would be an unusual financial application that has the ambiguity set of a vector of random variables be the same for all elements of the vector and it is easy to construct an example showing that monotonicity is violated. Non-coherence of RCVaR poses a challenge for further research.

---

**References**

- P. Artzner, F. Delbaen, J. M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.
- A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
- A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, Princeton, NJ, 2009.
- D. Bertsimas, G. J. Lauprete, and A. Samarov. Shortfall as a risk measure: Properties, optimization and applications. *Journal of Economic Dynamics and Control*, 28(7):1353–1381, 2004.
- D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM review*, 53(3):464–501, 2011.
- M. J. Best and R. R. Grauer. Sensitivity analysis for mean-variance portfolio problems. *Management Science*, 37(8):980–989, 1991.
- J. Čerbáková. Worst-case VaR and CVaR. In *Operations Research Proceedings 2005*, pages 817–822. Springer Berlin Heidelberg, 2006.
- S. Ceria and R. A. Stubbs. Incorporating estimation errors into portfolio selection: Robust portfolio construction. *Journal of Asset Management*, 7(2):109–127, 2006.
- L. Chen, S. He, and S. Zhang. Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research*, 59(4):847–865, 2011.
- V. K. Chopra and W. T. Ziemba. The effect of errors in means, variances, and covariances on optimal portfolio choice. *The Journal of Portfolio Management*, 19(2):6–11, 1993.
- E. Delage and Y. Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.
- F. J. Fabozzi, D. Huang, and G. Zhou. Robust portfolios: contributions from operations research and finance. *Annals of Operations Research*, 176(1):191–220, 2010.
- L. E. Ghaoui, M. Oks, and F. Oustry. Worst-case Value-at-Risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51(4):543–556, 2003.
- D. Goldfarb and G. Iyengar. Robust portfolio selection problems. *Mathematics of Operations Research*, 28(1):1–38, 2003.

- J.-Y. Gotoh, K. Shinozaki, and A. Takeda. Robust portfolio techniques for mitigating the fragility of CVaR minimization and generalization to coherent risk measures. *Quantitative Finance*, 13(10):1621–1635, 2013.
- M. Kaut, H. Vladimirov, S. W. Wallace, and S. A. Zenios. Stability analysis of portfolio management with conditional value-at-risk. *Quantitative Finance*, 7(4):397–409, 2007.
- F. H. Knight. *Risk, Uncertainty and Profit*. Houghton Mifflin Company, Boston, MA, 1921.
- S. Kou, X. Peng, and C. C. Heyde. External risk measures and Basel accords. *Mathematics of Operations Research*, 38(3):393–417, 2013.
- Z. Lu. A computational study on robust portfolio selection based on a joint ellipsoidal uncertainty set. *Mathematical Programming*, 126(1):193–201, 2011.
- J. M. Mulvey, R. J. Vanderbei, and S. A. Zenios. Robust optimization of large-scale systems. *Operations Research*, 43(2):264–281, 1995.
- A. B. Paç and M. Ç. Pinar. Robust portfolio choice with CVaR and VaR under distribution and mean return ambiguity. *Top*, 22(3):875–891, 2014.
- A. G. Quaranta and A. Zaffaroni. Robust optimization of conditional value at risk and portfolio selection. *Journal of Banking and Finance*, 32(10):2046–2056, 2008.
- R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.
- R. T. Rockafellar and S. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002.
- K. Schottle and R. Werner. Robustness properties of mean-variance portfolios. *Optimization*, 58(6):641–663, 2009.
- R. H. Tütüncü and M. Koenig. Robust asset allocation. *Annals of Operations Research*, 132(1-4):157–187, 2004.
- J. A. Wright, S.-C. P. Yam, and S. Yung. A test for the equality of multiple Sharpe ratios. *Journal of Risk*, 16(4):3–21, 2014.
- S. Zenios. *Practical Financial Optimization. Decision making for financial engineers*. Blackwell, Malden, MA, 2007.
- L. Zhu, Y. Li, and T. F. Coleman. Min-max robust and CVaR robust mean-variance portfolios. *Journal of Risk*, 11(3), 2008.
- S. Zhu and M. Fukushima. Worst-case conditional value-at-risk with application to robust portfolio management. *Operations Research*, 57(5):1155–1168, 2009.

## A Appendix: Proofs

### A.1 Proof of Theorem 2

To formulate the robust counterpart of (5) and (6) we need an explicit formulation of (10) and (11), respectively. Using Proposition 1 we write both RVaR and RCVaR models (10) and (11) as

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \max_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma})} & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} & \min_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma}), \pi \in \mathbb{D}} (\bar{\mu} - r_f e)^\top x \geq d - r_f. \end{aligned} \quad (21)$$

To find an explicit formulation we need the optimal value of the inner problem

$$\max_{(\bar{\mu}, \bar{\Gamma}) \in U_\delta(\hat{\mu}, \hat{\Gamma})} -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}. \quad (22)$$

We decompose the robustification first in  $\bar{\mu}$  and then in  $\bar{\Gamma}$  via an additional parameter  $\kappa \in (0, 1)$ , where  $U_{\sqrt{\kappa}\delta}(\hat{\mu})$ ,  $U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma})$  are defined as in Remark 3. It is easy to see that (22) is equivalent to:

$$\max_{\kappa \in [0, 1]} \max_{\bar{\Gamma} \in U_{\sqrt{1-\kappa}\delta}(\hat{\Gamma})} \max_{\bar{\mu} \in U_{\sqrt{\kappa}\delta}(\hat{\mu})} -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x}. \quad (23)$$

We start with the innermost maximization problem

$$\max_{\bar{\mu} \in U_{\sqrt{\kappa}\delta}(\hat{\mu})} -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x},$$

or, equivalently,

$$\begin{aligned} \max_{\bar{\mu} \in \mathbb{R}^n} & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} & S(\bar{\mu} - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu} - \hat{\mu}) \leq \kappa \delta^2. \end{aligned} \quad (24)$$

This is maximized for  $\bar{\mu}^* = \hat{\mu} - \delta \sqrt{\frac{\kappa}{S}} \frac{\hat{\Gamma}^{\frac{1}{2}} x}{\|\hat{\Gamma}^{\frac{1}{2}} x\|}$ . Plugging in this solution we obtain the middle maximization problem as

$$\begin{aligned} \max_{\bar{\Gamma} \in \mathbb{S}^n} & -r_f - (\hat{\mu} - r_f e)^\top x + \delta \sqrt{\frac{\kappa}{S}} \|\hat{\Gamma}^{\frac{1}{2}} x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} & \frac{S-1}{2} \|\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma} - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}\|_{tr}^2 \leq (1-\kappa)\delta^2. \end{aligned}$$

Using the variable transformation  $\bar{\bar{\Gamma}} = \hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma} - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}$ , this becomes

$$\begin{aligned} \max_{\bar{\bar{\Gamma}} \in \mathbb{S}^n} & -r_f - (\hat{\mu} - r_f e)^\top x + \delta \sqrt{\frac{\kappa}{S}} \|\hat{\Gamma}^{\frac{1}{2}} x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \hat{\Gamma} x + x^\top \hat{\Gamma}^{\frac{1}{2}} \bar{\bar{\Gamma}} \hat{\Gamma}^{\frac{1}{2}} x} \\ \text{s.t.} & \|\bar{\bar{\Gamma}}\|_{tr}^2 \leq \frac{2}{S-1} (1-\kappa)\delta^2. \end{aligned} \quad (25)$$

Since the square root function is monotonically increasing, the objective function of (25) is maximized if and only if  $x^\top \hat{T}^{\frac{1}{2}} \bar{\bar{T}} \hat{T}^{\frac{1}{2}} x$  is maximized. Thus, we let  $y := \hat{T}^{\frac{1}{2}} x$  and solve

$$\max_{\bar{\bar{T}} \in \mathbb{S}^n} y^\top \bar{\bar{T}} y \quad (26)$$

$$s.t. \quad \|\bar{\bar{T}}\|_{tr}^2 \leq \frac{2}{S-1} (1-\kappa) \delta^2. \quad (27)$$

The optimal solution  $\bar{\bar{T}}^*$  is given by

$$\bar{\bar{T}}^* = \delta \sqrt{\frac{2}{S-1} (1-\kappa)} \frac{y}{\|y\|} \cdot \frac{y^\top}{\|y\|}.$$

Plugging everything back we get the optimal value of (25) as

$$-r_f - (\hat{\mu} - r_f e)^\top x + \delta \sqrt{\frac{\kappa}{S}} \|\hat{T}^{\frac{1}{2}} x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{(1 + \delta \sqrt{\frac{2(1-\kappa)}{S-1}}) \|\hat{T}^{\frac{1}{2}} x\|},$$

which is substituted back in problem (23) to get

$$\begin{aligned} & \max_{\kappa \in [0,1]} -r_f - (\hat{\mu} - r_f e)^\top x + \delta \sqrt{\frac{\kappa}{S}} \|\hat{T}^{\frac{1}{2}} x\| + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{(1 + \delta \sqrt{\frac{2(1-\kappa)}{S-1}}) \|\hat{T}^{\frac{1}{2}} x\|} \\ &= -r_f - (\hat{\mu} - r_f e)^\top x + \max_{\kappa \in [0,1]} \left( \delta \sqrt{\frac{\kappa}{S}} + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{(1 + \delta \sqrt{\frac{2(1-\kappa)}{S-1}}) \|\hat{T}^{\frac{1}{2}} x\|} \right) \|\hat{T}^{\frac{1}{2}} x\| \\ &= -r_f - (\hat{\mu} - r_f e)^\top x + \left( \max_{\kappa \in [0,1]} f(\kappa) \right) \|\hat{T}^{\frac{1}{2}} x\|, \end{aligned}$$

which is the objective function of (15).

Now, the robust counterpart of minimum return constraint is equivalent to:

$$\min_{\kappa \in [0,1]} \min_{\bar{\mu} \in U_{\sqrt{\kappa}\delta}(\hat{\mu})} (\bar{\mu} - r_f e)^\top x \geq d - r_f.$$

Following the same course as in solving the inner problem (22), we get:

$$\min_{\kappa \in [0,1]} (\hat{\mu} - r_f e)^\top x - \delta \sqrt{\frac{\kappa}{S}} \|\hat{T}^{\frac{1}{2}} x\| \geq d - r_f.$$

The solution is  $\kappa = 1$  and the robust counterpart of minimum return constraint is:

$$(\hat{\mu} - r_f e)^\top x - \frac{\delta}{\sqrt{S}} \|\hat{T}^{\frac{1}{2}} x\| \geq d - r_f,$$

which is the constraint in (15). This completes the proof.

## A.2 Proof of Theorem 3

First we state some well-known properties of Kronecker product  $\otimes$ .

**Proposition 2** *Assume  $A, B, C, D$  and  $X$  are given matrices of conformable sizes.*

$$(i) \operatorname{tr}(AB) = \operatorname{tr}(BA)$$

$$(ii) \operatorname{tr}(A^\top B) = \operatorname{vec}(A)^\top \operatorname{vec}(B)$$

$$(iii) \operatorname{vec}(AXB) = (B^\top \otimes A)\operatorname{vec}(X)$$

$$(iv) (B \otimes A)(C \otimes D) = BC \otimes AD$$

$$(v) (B \otimes A)(C \otimes D) = BC \otimes AD$$

where  $\operatorname{vec}(A)$  denotes the vector obtained by stacking the columns of  $A \in \mathbb{R}^{m \times n}$  successively underneath each other.

We transform the problem using new variables  $(\hat{\mu}, \hat{\Gamma}^-)$  where  $\hat{\Gamma}^- = \hat{\Gamma}^{-1}$  is also positive definite, and develop the analysis on the transformed equivalent problem. We show that the transformed problem is convex and Slater condition hold. Therefore, we use the KKT optimality conditions to derive the optimal solution of the transformed problem and compute the inverse of  $\hat{\Gamma}^-$  to obtain  $\hat{\Gamma}$ .

Let  $H = H(\hat{\mu}, \hat{\Gamma})$  denote the objective function in (17). Then

$$H = \sum_{k=1}^K S(\bar{\mu}_k - \hat{\mu})^\top \hat{\Gamma}^{-1}(\bar{\mu}_k - \hat{\mu}) + \frac{S-1}{2} \operatorname{tr}(\hat{\Gamma}^{-\frac{1}{2}}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-1}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-\frac{1}{2}}). \quad (28)$$

Using property (i) we get

$$\begin{aligned} H &= \sum_{k=1}^K S \operatorname{tr}(\hat{\Gamma}^{-1}(\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) + \frac{S-1}{2} \operatorname{tr}(\hat{\Gamma}^{-1}(\bar{\Gamma}_k - \hat{\Gamma})\hat{\Gamma}^{-1}(\bar{\Gamma}_k - \hat{\Gamma})) \\ &= \sum_{k=1}^K S \operatorname{tr}(\hat{\Gamma}^{-1}(\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) + \frac{S-1}{2} \operatorname{tr}((\hat{\Gamma}^{-1}\bar{\Gamma}_k - I)^2). \end{aligned}$$

Applying properties (ii)–(iii) we get

$$\begin{aligned} H &= \sum_{k=1}^K S \operatorname{vec}((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top)^\top \operatorname{vec}(\hat{\Gamma}^{-1}) + \frac{S-1}{2} \operatorname{vec}(\bar{\Gamma}_k \hat{\Gamma}^{-1} - I)^\top \operatorname{vec}(\hat{\Gamma}^{-1} \bar{\Gamma}_k - I) \\ &= \sum_{k=1}^K S \operatorname{vec}((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top)^\top \operatorname{vec}(\hat{\Gamma}^{-1}) + \\ &\quad \sum_{k=1}^K \frac{S-1}{2} ((I \otimes \bar{\Gamma}_k) \operatorname{vec}(\hat{\Gamma}^{-1}) - \operatorname{vec}(I))^\top ((\bar{\Gamma}_k \otimes I) \operatorname{vec}(\hat{\Gamma}^{-1}) - \operatorname{vec}(I)). \end{aligned}$$



Finally, by replacing  $\hat{\Gamma}^{-1}$  by  $\Gamma^{-}$ , doing some straightforward calculations, and using properties (iv)–(v) we get the new formulation of  $H(\hat{\mu}, \hat{\Gamma})$  in terms of  $(\hat{\mu}, \Gamma^{-})$ , which we call  $G(\hat{\mu}, \Gamma^{-})$ ,

$$G(\hat{\mu}, \Gamma^{-}) = \text{vec}(\Gamma^{-})^\top \left[ \frac{S-1}{2} \sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k \right] \text{vec}(\Gamma^{-}) + \frac{nK(S-1)}{2} + \left[ S \sum_{k=1}^K \text{vec}((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) - (S-1) \sum_{k=1}^K \text{vec}(\bar{\Gamma}_k) \right]^\top \text{vec}(\Gamma^{-}). \quad (29)$$

Now that we have  $H$  as a function of  $(\hat{\mu}, \Gamma^{-})$ , we write the transformed problem as

$$\min_{\hat{\mu} \in \mathbb{R}^n, \Gamma^{-} \in \mathbb{S}_{++}^n} G(\hat{\mu}, \Gamma^{-}).$$

One can easily check that the Hessian matrix of function  $G$  is positive semidefinite and Slater condition holds, hence KKT conditions give us the optimal solution. There are no constraints on mean returns, and  $\Gamma^{-}$ , being the inverse of a positive definite matrix, is positive definite and in the interior of the positive semidefinite cone. Hence, the KKT optimality conditions reduce to:

$$\begin{aligned} \nabla G_{\hat{\mu}} &= 0, \\ \nabla G_{\text{vec}(\Gamma^{-})} &= 0. \end{aligned} \quad (30)$$

To obtain  $\nabla G_{\hat{\mu}}$  we take the differential with respect to  $\hat{\mu}$ :

$$\begin{aligned} dG &= d \sum_{k=1}^K S \text{vec}((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top)^\top \text{vec}(\Gamma^{-}) \\ &= \sum_{k=1}^K S \text{vec}(d((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top))^\top \text{vec}(\Gamma^{-}) \\ &= \sum_{k=1}^K S (\text{vec}(d\hat{\mu}(\hat{\mu} - \bar{\mu}_k)^\top)^\top \text{vec}(\Gamma^{-}) + \text{vec}((\hat{\mu} - \bar{\mu}_k)(d\hat{\mu})^\top)^\top \text{vec}(\Gamma^{-})). \end{aligned}$$

Using properties (ii)–(iii) we get:

$$\begin{aligned} dG &= \sum_{k=1}^K S (\text{tr}((\hat{\mu} - \bar{\mu}_k)(d\hat{\mu})^\top \Gamma^{-}) + \text{tr}(d\hat{\mu}(\hat{\mu} - \bar{\mu}_k)^\top \Gamma^{-})) \\ &= \sum_{k=1}^K S (\text{tr}((\hat{\mu} - \bar{\mu}_k)(d\hat{\mu})^\top \Gamma^{-}) + \text{tr}(d\hat{\mu}(\hat{\mu} - \bar{\mu}_k)^\top \Gamma^{-})) \\ &= \sum_{k=1}^K 2S (\hat{\mu} - \bar{\mu}_k)^\top \Gamma^{-} d\hat{\mu}. \end{aligned}$$

Hence,

$$\nabla G_{\hat{\mu}} = \left( \frac{dG}{d\hat{\mu}} \right)^\top = \sum_{k=1}^K 2S \Gamma^{-} (\hat{\mu} - \bar{\mu}_k).$$

Therefore the solution of  $\nabla G_{\hat{\mu}} = 0$  is  $\hat{\mu} = \frac{1}{K} \sum_{k=1}^K \bar{\mu}_k$ .

Calculating  $\nabla G_{vec(\Gamma^-)}$  we obtain the second equation in (30) as:

$$\left[ (S-1) \sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k \right] vec(\Gamma^-) + \left[ S \sum_{k=1}^K vec((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) - (S-1) \sum_{k=1}^K vec(\bar{\Gamma}_k) \right] = 0,$$

which suggests the following system of linear equations in terms of  $vec(\Gamma^-)$ :

$$\left[ \sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k \right] vec(\Gamma^-) = \left[ \sum_{k=1}^K vec(\bar{\Gamma}_k) - \frac{S}{(S-1)} \sum_{k=1}^K vec((\hat{\mu} - \bar{\mu}_k)(\hat{\mu} - \bar{\mu}_k)^\top) \right] = 0.$$

All  $\bar{\Gamma}_k$ ,  $k \in \{1, \dots, K\}$  are positive semidefinite matrices. The sum and Kronecker product of two positive semidefinite matrices are positive semidefinite matrices, thus  $\sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k$  is positive semidefinite. To prove uniqueness of solution, assume  $\bar{\Gamma}_l$ , for some  $l \in \{1, \dots, K\}$  is a positive definite matrix, then so is  $\bar{\Gamma}_l \otimes \bar{\Gamma}_l$ . Also,  $\bar{\Gamma}_k \otimes \bar{\Gamma}_k$ , for all  $k \in \{1, \dots, K\} \setminus \{l\}$  are positive semidefinite matrices. These all together imply that  $\sum_{k=1}^K \bar{\Gamma}_k \otimes \bar{\Gamma}_k$  is a positive definite matrix, that is the coefficient matrix is a full rank matrix and thus (18) has a unique solution.

## B Appendix: Extension to polytopic ambiguity sets

We give now the formal definition, the relevant theorem, and its proof, to specify RVaR and RCVaR optimization models for polytopic ambiguity sets on the mean returns and covariance matrix.

**Definition 3 (Polytopic ambiguity for mean returns and covariance matrix)** Mean returns and covariance matrix belong to the following polytopic set:

$$U_P = \{(\bar{\mu}, \bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid \bar{\mu} = \sum_{j=1}^J \rho_j \bar{\mu}_j, \bar{\Gamma} = \sum_{j=1}^J \rho_j \bar{\Gamma}_j, \sum_{j=1}^J \rho_j = 1, \rho_j \geq 0, j = 1, \dots, J\},$$

where  $(\bar{\mu}_j, \bar{\Gamma}_j) \in \mathbb{R}^n \times \mathbb{S}_+^n$ ,  $j = 1, \dots, J$ , are the polytope vertices.

*Remark 8* The polytopic ambiguity set can be written as  $U_{P_1} \times U_{P_2}$ , where

$$U_{P_1} = \{\bar{\mu} \in \mathbb{R}^n \mid \bar{\mu} = \sum_{j=1}^J \rho_j \bar{\mu}_j, \sum_{j=1}^J \rho_j = 1, \rho_j \geq 0, j = 1, \dots, J\}$$

$$U_{P_2} = \{\bar{\Gamma} \in \mathbb{S}^n \mid \bar{\Gamma} = \sum_{j=1}^J \rho_j \bar{\Gamma}_j, \sum_{j=1}^J \rho_j = 1, \rho_j \geq 0, j = 1, \dots, J\}$$

are polytopic ambiguity sets for means and covariance matrix, respectively.

**Theorem 4** *If random variable  $\xi$  has a distribution from the set  $\mathbb{D}$  and  $(\bar{\mu}, \bar{\Gamma}) \in U_P$ . Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) under polytopic ambiguity are represented by the following SOCP:*

$$\begin{aligned} \min_{x \in \mathbb{R}^n, \omega \in \mathbb{R}, \beta \in \mathbb{R}} \quad & \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \omega - \beta \\ \text{s.t.} \quad & r_f + (\bar{\mu}_j - r_f e)^\top x \geq d, \quad j = 1, \dots, J, \\ & r_f + (\bar{\mu}_j - r_f e)^\top x \geq \beta, \quad j = 1, \dots, J, \\ & \sqrt{x^\top \bar{\Gamma}_j x} \leq \omega, \quad j = 1, \dots, J. \end{aligned} \quad (31)$$

**Proof.**

By Proposition (1), RVaR and RCVaR optimization models can be written as:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \max_{(\bar{\mu}, \bar{\Gamma}) \in U_P} \quad & -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} \quad & \min_{(\bar{\mu}, \bar{\Gamma}) \in U_P} r_f + (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \quad (32)$$

Using the representation  $U_{P_1} \times U_{P_2}$  of  $U_P$ , one can easily see that the inner optimization problems appeared in objective function and constraint can be decomposed into easier subproblems and thus we get:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & -r_f - \min_{\bar{\mu} \in U_{P_1}} (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \max_{\bar{\Gamma} \in U_{P_2}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} \quad & r_f + \min_{\bar{\mu} \in U_{P_1}} (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \quad (33)$$

Obviously,  $\min_{\bar{\mu} \in U_{P_1}} (\bar{\mu} - r_f e)^\top x = \min_{1 \leq j \leq J} (\bar{\mu}_j - r_f e)^\top x$  and

$$\max_{\bar{\Gamma} \in U_{P_2}} \sqrt{x^\top \bar{\Gamma} x} = \max_{1 \leq j \leq J} \sqrt{x^\top \bar{\Gamma}_j x}.$$

Letting  $\beta = r_f + \min_{1 \leq j \leq J} (\bar{\mu}_j - r_f e)^\top x$  and  $\omega = \max_{1 \leq j \leq J} \sqrt{x^\top \bar{\Gamma}_j x}$ , we get the result.

## C Appendix: Extension to interval ambiguity sets

We give now the formal definition, the relevant theorem, and its proof, to specify RVaR and RCVaR optimization models for interval ambiguity sets on the mean returns and covariance matrix.

**Definition 4 (Interval ambiguity for mean returns and covariance matrix)** Mean returns and covariance matrix belong to the following interval set:

$$U_I = \{(\bar{\mu}, \bar{\Gamma}) \in \mathbb{R}^n \times \mathbb{S}^n \mid \bar{\mu}_- \leq \bar{\mu} \leq \bar{\mu}_+, \bar{\Gamma}_- \leq \bar{\Gamma} \leq \bar{\Gamma}_+\},$$

where  $\bar{\mu}_-, \bar{\mu}_+, \bar{\Gamma}_-, \bar{\Gamma}_+$  are given vectors and matrices and the inequalities are component-wise. We assume there is at least one  $(\bar{\mu}, \bar{\Gamma}) \in U_I$  for which  $\bar{\Gamma} \succeq 0$ .

*Remark 9* The interval ambiguity set can be written as  $U_{I_1} \times U_{I_2}$ , where

$$U_{I_1} = \{\bar{\mu} \in \mathbb{R}^n \mid \bar{\mu}_- \leq \bar{\mu} \leq \bar{\mu}_+\}$$

$$U_{I_2} = \{\bar{\Gamma} \in \mathbb{S}^n \mid \bar{\Gamma}_- \leq \bar{\Gamma} \leq \bar{\Gamma}_+\}$$

are interval ambiguity sets for means and covariance matrix, respectively.

**Theorem 5** *If random variable  $\xi$  has a distribution from the set  $\mathbb{D}$  and  $(\bar{\mu}, \bar{\Gamma}) \in U_I$ . Then, the robust counterpart to VaR portfolio optimization model (5) and the robust counterpart to CVaR model (6) under interval ambiguity are represented by the following SDP:*

$$\begin{aligned} \min_{v \in \mathbb{R}, x_+, x_- \in \mathbb{R}^n, \Lambda, \Lambda_+, \Lambda_- \in \mathbb{S}^n} & \quad tr(\Lambda_+ \bar{\Gamma}_+) - tr(\Lambda_- \bar{\Gamma}_-) + \\ & \quad \frac{\alpha}{1-\alpha} v + (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_- \\ \text{s.t.} & \quad \begin{bmatrix} \Lambda & \frac{x_- - x_+}{2} \\ \frac{(x_- - x_+)^\top}{2} & v \end{bmatrix} \succeq 0, \\ & \quad \Lambda \preceq \Lambda_+ - \Lambda_-, \\ & \quad (\bar{\mu}_- - r_f e)^\top x_- - (\bar{\mu}_+ - r_f e)^\top x_+ \geq d - r_f, \\ & \quad x_+, x_- \geq 0, \Lambda, \Lambda_+, \Lambda_- \succeq 0. \end{aligned} \tag{34}$$

**Proof.**

By Proposition (1), RVaR and RCVaR optimization models can be written as:

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \max_{(\bar{\mu}, \bar{\Gamma}) \in U_I} & \quad -r_f - (\bar{\mu} - r_f e)^\top x + \frac{\sqrt{\alpha}}{\sqrt{1-\alpha}} \sqrt{x^\top \bar{\Gamma} x} \\ \text{s.t.} & \quad \min_{(\bar{\mu}, \bar{\Gamma}) \in U_I} r_f + (\bar{\mu} - r_f e)^\top x \geq d. \end{aligned} \tag{35}$$

We start with the objective function. Using Theorem 1 of Ghaoui et al. (2003) we replace the inner maximization objective function by the following:

$$\begin{aligned} \min_{v \in \mathbb{R}, \Lambda \in \mathbb{S}_+^n} & \quad tr(\Lambda \bar{\Gamma}) + \frac{\alpha}{1-\alpha} v - (r_f + (\bar{\mu} - r_f e)^\top x) \\ \text{s.t.} & \quad \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x^\top}{2} & v \end{bmatrix} \succeq 0, \end{aligned} \tag{36}$$

Hence, the inner maximization in the objective function of model (35) is equivalent to:

$$\begin{aligned} \max_{(\bar{\mu}, \bar{\Gamma}) \in U_I} \min_{v \in \mathbb{R}, \Lambda \in \mathbb{S}_+^n} & \quad tr(\Lambda \bar{\Gamma}) + \frac{\alpha}{1-\alpha} v - (r_f + (\bar{\mu} - r_f e)^\top x) \\ \text{s.t.} & \quad \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x^\top}{2} & v \end{bmatrix} \succeq 0. \end{aligned} \tag{37}$$

Convexity and compactness of feasible region and linearity of objective function with respect to  $\bar{\mu}$  and  $\bar{\Gamma}$  for fixed  $\Lambda$  and  $v$  (and conversely) imply that we can exchange “min” and “max” to obtain

$$\begin{aligned} \min_{v \in \mathbb{R}, \Lambda \in \mathbb{S}_+^n} \max_{(\bar{\mu}, \bar{\Gamma}) \in U_I} \operatorname{tr}(\Lambda \bar{\Gamma}) + \frac{\alpha}{1-\alpha} v - (r_f + (\bar{\mu} - r_f e)^\top x) \quad (38) \\ \text{s.t.} \\ \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & v \end{bmatrix} \succeq 0. \end{aligned}$$

Decompose now the inner maximizations into easier subproblems by applying the  $U_{I_1} \times U_{I_2}$  representation of  $U_I$  to derive the following formulation of (38):

$$\begin{aligned} \min_{v \in \mathbb{R}, \Lambda \in \mathbb{S}_+^n} -r_f + \frac{\alpha}{1-\alpha} v + \max_{\bar{\Gamma} \in U_{I_2}} \operatorname{tr}(\Lambda \bar{\Gamma}) + \max_{\bar{\mu} \in U_{I_1}} -(\bar{\mu} - r_f e)^\top x \quad (39) \\ \text{s.t.} \\ \begin{bmatrix} \Lambda & \frac{x}{2} \\ \frac{x}{2}^\top & v \end{bmatrix} \succeq 0. \end{aligned}$$

The dual formulations of the maximization problems in (39) are

$$\max_{\bar{\Gamma} \in U_{I_2}} \operatorname{tr}(\Lambda \bar{\Gamma}) = \min_{\Lambda_+, \Lambda_- \succeq 0, \Lambda \preceq \Lambda_+ - \Lambda_-} \operatorname{tr}(\Lambda_+ \bar{\Gamma}_+) - \operatorname{tr}(\Lambda_- \bar{\Gamma}_-),$$

and

$$\max_{\bar{\mu} \in U_{I_1}} -(\bar{\mu} - r_f e)^\top x = \min_{x_+, x_- \geq 0, x = x_- - x_+} (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_-.$$

Under suitable conditions —primal and dual strict feasibility— the duality gap in the first optimization problem above is zero, and we obtain the objective function of (35) as

$$\begin{aligned} \min_{v \in \mathbb{R}, x_+, x_- \in \mathbb{R}^n, \Lambda, \Lambda_+, \Lambda_- \in \mathbb{S}^n} \frac{\alpha}{1-\alpha} v + \operatorname{tr}(\Lambda_+ \bar{\Gamma}_+) - \operatorname{tr}(\Lambda_- \bar{\Gamma}_-) + \quad (40) \\ (\bar{\mu}_+ - r_f e)^\top x_+ - (\bar{\mu}_- - r_f e)^\top x_- \\ \text{s.t.} \\ \begin{bmatrix} \Lambda & \frac{x_- - x_+}{2} \\ \frac{(x_- - x_+)}{2}^\top & v \end{bmatrix} \succeq 0, \\ \Lambda \preceq \Lambda_+ - \Lambda_-, \\ x_+, x_- \geq 0, \Lambda, \Lambda_+, \Lambda_- \succeq 0. \end{aligned}$$

To complete the robust counterpart (35), we need an explicit formulation of the robust counterpart of minimum return constraint:

$$\min_{(\bar{\mu}, \bar{\Gamma}) \in U_I} r_f + (\bar{\mu} - r_f e)^\top x \geq d.$$

To do this, we write the minimization problem as  $-r_f + \max_{\bar{\mu} \in U_{I_1}} -(\bar{\mu} - r_f e)^\top x$  and use the result on dual form discussed above. Hence, the robust counterpart of minimum return constraint is:

$$r_f - (\bar{\mu}_+ - r_f e)^\top x_+ + (\bar{\mu}_- - r_f e)^\top x_- \geq d.$$

This completes the proof.

## D Appendix: Data

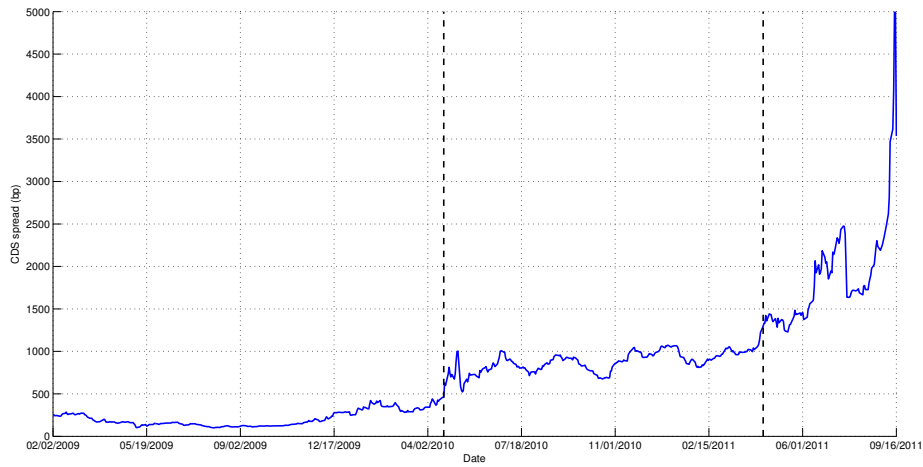


Fig. 13: CDS spreads of the Greek sovereign with identified regimes: Leftmost is the tranquil period, center is the turbulent, and the rightmost is the crisis.

Table 1: Moments of monthly differentials of the historical market data.

	Stock.USA	Stock.UK	Stock.DE	Stock.JP	Bnd1.USA	Bnd2.USA
Mean	0.01296	0.01047	0.01057	-0.00189	0.00553	0.00702
SD	0.04101	0.04150	0.05796	0.06184	0.00467	0.01620
Skewness	-0.47903	-0.19051	-0.47281	0.04768	-0.18341	-0.07482
Kurtosis	3.76519	3.11399	4.11970	3.62119	2.77801	3.23974