DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 4/11

Construction of Uncertainty Sets for Portfolio Selection Problems

by

Christof Wiechers

March 27, 2011

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

This page is intentionally left blank.

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 4/11

Construction of Uncertainty Sets for Portfolio Selection Problems[1]

by

Christof Wiechers[2]

March 27, 2011

Abstract

While modern portfolio theory grounds on the trade-off between portfolio return and portfolio variance to determine the optimal investment decision, post-modern portfolio theory uses downside risk measures instead of the variance. Prominent examples are given by the risk measures Value-at-Risk and its coherent extension, Conditional Value-at-Risk. When avoiding distributional assumptions on the process that generates the risky assets' returns, historical return data or expert knowledge remain the only data available to the investor. His problem is then to maximize the return of his portfolio given the risk constraint that his portfolio does not fall short of some threshold return. For the Conditional Value-at-Risk, the solution is known to be achievable by a linear program. This paper extends the solution to the investor's problem whenever his risk preferences are given by any coherent distortion risk measure. More precisely, it is shown that whenever the risk constraint is given by a coherent distortion risk measure, a linear program leads to the solution. A geometric interpretation of this solution is immediate, which is related to the non-parametric description of data by so-called weighted-mean trimmed regions. The connections of the solution to robust optimization and decision theory are illustrated.

JEL Classification: C13, C18, C61, G11, G32.

Keywords: Portfolio Optimization, Risk Constraints, Coherent Distortion Risk Measures, Uncertainty Sets.

[1]This paper is part of a doctoral thesis written at Cologne University during 2008-2011. Many thanks belong to Karl Mosler for his important comments on the manuscript. Furthermore, Pavel Bazovkin has largely supported this work by providing the appropriate software and valuable assistance.

[2]Phone: +49 221 470-6334, email: wiechers@wiso.uni-koeln.de, University of Cologne, Graduate School of Risk Management, Albertus-Magnus-Platz, D-50923 Cologne. Wiechers gratefully acknowledges financial support provided the DFG Research Training Group *"Risk Management."* The usual disclaimer applies.

This page is intentionally left blank.

1 Introduction

The one-period portfolio selection problem, as formulated by Markowitz (1952), under certain assumptions - the most common of which is the assumption of normally distributed asset returns, $\bar{R} \sim \mathcal{N}_d(\mu, \Sigma)$ - reads

$$\max_w \ \mu' w - \frac{\lambda}{2} w' \Sigma w \quad \text{s.t. } w \in \mathcal{C}. \tag{1}$$

Here, $\mu = (\mu_1, \ldots, \mu_d)'$ denotes the mean vector of the asset returns, whereas $\Sigma := [\sigma_{ij}],\ i, j = 1, \ldots, d$, is the covariance matrix of asset returns, so that $\sigma_{ij} = \text{cov}(\bar{R}_i, \bar{R}_j)$. The decision variable $w \in \mathbb{R}^d$ represents the fractions of wealth the investor allocates among the d assets, and as above, w might be subject to some constraints given by the constraint set $\mathcal{C}$. λ denotes the investor-specific coefficient of risk aversion.

Besides being called the *Markowitz model*, (1) is often also referred to as the *mean-variance model*. There exist numerous extensions to this model, especially when multi-period portfolio optimization problems are considered. It is noteworthy to emphasize that in this paper, only one-period portfolio optimization is studied.

In the mean-variance model, the portfolio return variance $w'\Sigma w$ is used as the penalizing factor to downsize the expected portfolio return $\mu' w$. This downsizing accounts for the investor's dislike of fluctuation of his portfolio return. In the theoretical case of $\lambda = 0$, the investor fully invests into only one asset, which is the asset with the highest expected mean return. As this typically is not the case whenever $\lambda > 0$, the penalization via the variance term can be seen to enforce diversification among the available assets. This fact is the main finding of Markowitz (1952).

There exists a different strand of literature, called *post-modern portfolio theory*, that criticizes the use of the variance as a means for measuring the dispersion of financial positions, see Ferguson and Rom (1994) or Nawrocki (1999). An approach to measuring multi-period risk is given by Pflug (2006), who uses a value-of-information approach to assess the risk of an investment process.

The variance aggregates upside as well as downside deviations of the portfolio in one number, whereas financial risk measures - as to be defined in Section 3 - only take account of downside deviations of a financial position. As the portfolio return $\mu' w$ of some portfolio w is such an uncertain financial position, a natural approach is to apply downside risk measures. To this end, denote the uncertain mean vector of the d risky asset returns by some random variable $\tilde{\mu}$, the distribution of which is not available to the investor.

For some downside risk measure ρ, a general formulation of the above can be cast in the following way,

$$\max_w \ \left(\mathbb{E}(\tilde{\mu})\right)' w \quad \text{s.t. } \rho(\tilde{\mu}' w - \bar{\mu}) \leq 0, \ \ w \in \mathcal{C}. \tag{2}$$

Here, $\bar{\mu}$ denotes some benchmark level of portfolio return.

In the context of portfolio optimization, (2) shall be called the *mean-risk model* within this paper, as opposed to the mean-variance model (1). The name *mean-risk model* is used in the literature as well whenever *"approaches are based on comparing two scalar characteristics (summary statistics), the first of which* [...] *represents the expected outcome (reward), and the second* [...] *is some measure of risk,"* cp. Ogryczak and Ruszczynski (1999, p. 35).

When the risk measure ρ is given by the Conditional Value-at-Risk at some level α, denoted by CVaR_α, problem (2) can be reformulated as a linear program, as shown by Rockafellar and Uryasev (2000) and Krokhmal et al. (2002). Given a time series $(R_1, \dots, R_T)$ of observed asset returns for the d risky assets, with $R_t \in \mathbb{R}^d$ for $t = 1, \dots, T$, a solution of (2) is thus readily obtained for the risk measure CVaR_α.

In this paper, (2) is solved for the whole class of coherent distortion risk measures, which in turn constitutes a sub-class of the important coherent risk measures. The key observation to the solution is the fact that the mean-risk problem can be reformulated as a linear optimization problem with uncertain data, where the latter is captured by a so-called *uncertainty set* $\mathcal{U}$. An uncertainty set $\mathcal{U}$ is a set in the space of the uncertain parameters, and its shape is determined solely by the data and the risk measure ρ. It turns out that for the class of coherent distortion risk measures, this uncertainty set has a convex shape.

In the literature these sets are sometimes called *confidence regions*. While the concepts of uncertainty sets and confidence regions coincide, the name *uncertainty set* is chosen in this paper to emphasize the absence of any distributional assumption for the construction of these sets. In the same manner, the name *ambiguity set* is used in the literature. Pflug and Wozabal (2007) use the term *ambiguity sets* to describe collections of probability measures rather than collections of parameter values, cp. Section 4 therein. To avoid any confusion, the name *ambiguity set* is thus not used throughout this paper.

The paper by Bertsimas and Brown (2009) illustrates the connection between coherent risk measures and uncertainty sets, so this paper builds heavily on their work. In addition, Dyckerhoff and Mosler (2010b) examine non-parametric descriptions of datasets of arbitrary dimensions by so-called *weighted-mean trimmed regions*. With the advent of an exact algorithm for the calculation of these regions, as given by Bazovkin and Mosler (2010), it is now possible to numerically construct uncertainty sets corresponding to the risk constraint in (2). It is noteworthy that these regions, which coincide with the above mentioned uncertainty sets, are not subject to any distributional assumption, while including only historical data or expert knowledge.

Combining these works, a solution to the mean-risk portfolio problem (2) is given whenever the risk constraint involves a coherent distortion risk measure. This includes the important case of the measure CVaR_α, but furthermore accomodates any measure that is obtained via scenario generation techniques. Furthermore, it is shown that this solution is obtained via solving a finite number of linear programs.

Optimization under input parameter uncertainty is also connected to a recently developed sub-field of robust optimization, the so-called *minimax portfolio optimization*, cp. Halldórsson and Tütüncü (2003) or Garlappi et al. (2007). On the other hand, the minimax criterion is also well known as a paradigm of decision theory, and both connections will be explored in more detail in the sequel.

The outline of this paper is as follows. The minimax methodology is examined in Section 2 in more detail, and its connections to robust portfolio optimization as well as to decision theory are presented. Section 3 provides an overview on coherent risk measures and coherent distortion risk measures. The main Section 4 outlines the generation of uncertainty sets from a given coherent risk measure, as well as its application to the portfolio problem and its solution. Finally, Section 5 applies the methodology to both financial and simulated data, whereas Section 6 concludes this paper.

2 The Minimax Approach

While the minimax approach regarding portfolio optimization is often used in the context of robust optimization, it is important to note that in the first line, the minimax criterion is a concept from decision theory. Whenever a decision that has to be made today is influenced by the uncertain state of nature prevailing in the future, a rational decision maker will apply some methodology to base his decision on. The minimax criterion is such a methodology, and it says, roughly, that the decision is made for *the best among the worst* scenarios. In other words, he identifies the most unfavorable consequences of his decision in all possible future states of the world, and then makes the decision for the best among these alternatives.

2.1 The Minimax Approach in Mean-Variance Analysis

Application of the minimax approach to the mean-variance model (1) means to identify the portfolio w that is optimal for all, thus even the most unfavorable states of nature. These states of nature are represented by all parameters that the investor has no influence on, but which in turn influence the investor's decision and payoff. In the mean-variance model, the states of nature are completely described by the mean of the asset returns, μ, and the covariance structure of these asset returns, Σ.

For the case of the vector μ of expected asset returns, this means that μ is located among a collection of possible outcomes, the *uncertainty set* $\mathcal{U}_\mu \subset \mathbb{R}^d$. In the same fashion, the covariance matrix of the d risky assets is assumed to be located in a subset of $\mathbb{R}^{d\times d}$, yielding the uncertainty set $\mathcal{U}_\Sigma$.

Then, the minimax approach applied to the mean-variance model (1) reads

$$\max_{w} \left\{ \min_{\mu,\Sigma} \left\{ \mu' w - \frac{\lambda}{2} w' \Sigma w \right\} \right\} \quad \text{s.t. } \mu \in \mathcal{U}_\mu,\ \Sigma \in \mathcal{U}_\Sigma,\ w \in \mathcal{C}. \tag{3}$$

As such, when solving the above problem, the investor chooses the portfolio w that stays optimal for even the most disadvantageous values of μ and Σ captured by the respective uncertainty sets. Halldórsson and Tütüncü (2003) study models of this type in more detail.

As suggested by Chopra and Ziemba (1993), the 'largest losses' in data-driven portfolio allocation can be associated with the estimation of the mean vector μ, whereas estimation of the covariance matrix Σ by its sample counterpart $\hat{\Sigma}$ or by some shrinkage estimator $\tilde{\Sigma}$ yields quite suitable results. Thus, whenever a reliable estimator for the covariance structure is at hand, the application of the minimax approach can be restricted to the uncertainty inherent in the mean vector of asset returns, while using some point estimate $\bar{\Sigma}$ for the covariance structure, viz.

$$\max_{w} \left\{ \min_{\mu} \left\{ \mu' w - \frac{\lambda}{2} w' \bar{\Sigma} w \right\} \right\} \quad \text{s.t. } \mu \in \mathcal{U}_{\mu},\ w \in \mathcal{C},$$

which is equivalent to

$$\max_{w} \left\{ \min_{\mu} \left\{ \mu' w \right\} - \frac{\lambda}{2} w' \bar{\Sigma} w \right\} \quad \text{s.t. } \mu \in \mathcal{U}_{\mu},\ w \in \mathcal{C}. \tag{4}$$

By carrying out the optimization over all values of μ in the uncertainty set $\mathcal{U}_{\mu}$, the usage of a wrong point estimate of the parameter μ is avoided. As such, drawbacks from estimation error in μ are mitigated, and the solution is less influenced by unexpected outcomes of μ.

When the uncertainty sets $\mathcal{U}_{\mu}$ and $\mathcal{U}_{\Sigma}$ are chosen to be convex and compact, and when the set of feasible portfolio weight vectors $\mathcal{C}$ is convex, the Min-Max Theorem guarantees that the order of minimization and maximization does not influence the solution of problem (3) or (4). Thus, due to the combination of maximization and minimization, problem formulations like (3) or (4) are called *minimax problems*, as described in the famous work of Savage (1972). The rationale behind the minimax approach from a decision theoretic point of view is to *avoid disappointment*, i.e., to obtain a solution w^{opt} that is optimal for even the most disadvantageous market evolution that is captured by the uncertainty set.

2.2 The Minimax Approach in Mean-Risk Analysis

As indicated in the introduction, the trade-off between portfolio return $w'\mu$ and its variance $w'\Sigma w$ can be replaced by analyzing the trade-off between portfolio return and some other measure of risk for the portfolio return characteristics. When such a risk measure ρ is selected, and a certain threshold return $\bar{\mu}$ is aimed at, the problem can be formulated as

$$\max_{w} \left\{ \big(\mathbb{E}(\tilde{\mu})\big)' w \right\} \quad \text{s.t. } \rho\left(\tilde{\mu} w - \bar{\mu}\right) \leq 0,\ w \in \mathcal{C}. \tag{5}$$

Note that the uncertain mean vector of asset returns is modeled as a random variable $\tilde{\mu}$. In a recent paper by Bertsimas and Brown (2009), it is shown that whenever ρ is a coherent measure of risk, the above can be reformulated as

$$\max_{w} \left\{ \min_{\mu} \left\{ \mu' w \right\} \right\} \quad \text{s.t. } \mu \in \mathcal{U}_{\rho}(\mathcal{R}),\ w \in \big(\mathcal{W}_{\bar{\mu},\rho}(\mathcal{R}) \cap \mathcal{C}\big), \tag{6}$$

with properly defined sets $\mathcal{U}_\rho(\mathcal{R})$ and $\mathcal{W}_{\bar{\mu},\rho}(\mathcal{R})$ that depend on the risk measure ρ, the threshold return $\bar{\mu}$, and on the data $\mathcal{R}$ available. For brevity, these sets will only be denoted by $\mathcal{U}_\rho$ and by $\mathcal{W}_{\bar{\mu}}$, but it should be kept in mind that they depend on the data of the problem, $\mathcal{R}$, as well.

The exact definitions of the sets $\mathcal{U}_\rho$ and $\mathcal{W}_{\bar{\mu}}$ will be presented in detail in Section 4. Again, for convex and compact $\mathcal{U}_\rho$ and convex $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, the Min-Max Theorem is applicable, justifying the classification of the above problem as a minimax problem. The construction of uncertainty sets that arise by the equivalence of (5) and (6) will also be investigated in Section 4.

3 Coherent Risk Measures

Let Ω be some set (of states of the world), and let $\mathcal{F}$ be a σ-algebra over Ω, so that $(\Omega, \mathcal{F})$ becomes a measurable space. When $(\Omega, \mathcal{F})$ is endowed with a probability measure $\mathbb{P}$, then the triple $(\Omega, \mathcal{F}, \mathbb{P})$ defines a probability space. Furthermore, let $\mathcal{X}$ denote the set of all bounded random variables on that space. A typical element X of $\mathcal{X}$ can be thought of as the difference of the known present value of some financial position and the random future value of that position, or just as the unknown future value of the position. Föllmer and Schied (2004) refer to X as the *payoff profile*. As such, negative values of X are interpreted as losses, whereas positive values of X indicate gains in the financial position.

A *risk measure* is now expected to assign a number to each financial position X, with riskier positions being assigned a higher value. The number assigned to X by the risk measure can then be interpreted as the amount of capital that must be reserved in order to limit the risk inherent in that position to some predefined low level.

Definition 3.1 *A functional $\rho : \mathcal{X} \to \mathbb{R}$ is called a* coherent risk measure *it fulfills the following axioms for any $X, Y \in \mathcal{X}$.*

- Monotonicity:
  Whenever $X(\omega) > Y(\omega)\ \forall \omega \in \Omega$, it holds that $\rho(Y) > \rho(X)$.
- Translation Equivariance:
  For any constant $c \in \mathbb{R}$, it holds that $\rho(X + c) = \rho(X) - c$.
- Convexity:
  For $\lambda \in [0, 1]$, it holds that $\rho\big(\lambda X + (1 - \lambda)Y\big) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$.
- Positive Homogeneity:
  For $\lambda \geq 0$, it holds that $\rho(\lambda X) = \lambda\,\rho(X)$.

The axioms have the following, immediate interpretations. *Monotonicity* implies that, if the future value of position one is larger than the future value of position two in any state, then position two is more risky. The addition of a sure amount of capital $c > 0$ to a risky position lowers the risk of that position. For negative c, the risk grows, as a sure loss is added. *Translation Equivariance*, or *Cash Invariance*, as referred to this axiom in Föllmer and Schied (2004), thus implies the previous interpretation of $\rho(X)$ as *risk capital*. While the axiom of *Positive Homogeneity* states that the risk of a multiple of some position is the multiple of its risk, adding *Convexity* states that diversification of a risky position will never increase the risk inherent in its components.

Whenever some risk measure ρ obeys the axioms of positive homogeneity and convexity, it also fulfills *Subadditivity*, that is, for X and $Y \in \mathcal{X}$, it follows that

$$\rho\,(X\,+\,Y)\;\leq\;\rho\,(X)\,+\,\rho\,(Y)\,.$$

Example 3.2 For a random variable $X \in \mathcal{X}$ and a given confidence level α, the *Conditional Value-at-Risk* of X at level $\alpha > 0$, *CVaR*$_\alpha$(X), is defined as

$$\text{CVaR}_\alpha(X)\;=\;\frac{1}{\alpha}\int_0^\alpha \text{VaR}_\xi(X)\,d\xi\;=\;-\frac{1}{\alpha}\int_0^\alpha q_X^+(\xi)\,d\xi. \tag{7}$$

Thus, the *CVaR*$_\alpha$ of a random variable is the mean of X when restricted to values that exceed *VaR*$_\alpha$, the *Value-at-Risk* at level α. Here, $q_X^+ : (0,1) \to (-\infty,\infty)$ denotes the *quantile function* of X, as defined in Föllmer and Schied (2004).

The main theorem for coherent risk measures shall be stated subsequently. While its mathematical content has already been proven in Huber (1981, Chapter 10), its prominence is due to the paper by Artzner et al. (1999). The formulation below coincides with the one given in Föllmer and Schied (2004).

Theorem 3.3 *A functional $\rho: \mathcal{X} \longrightarrow \mathbb{R}$ is a coherent risk measure if and only if there exists a family of probability measures $\mathcal{Q}$ on $(\Omega,\mathcal{F})$, such that*

$$\rho(X)\;=\;\sup_{Q\in\mathcal{Q}}\mathbb{E}_Q(-X),\quad X\in\mathcal{X}. \tag{8}$$

$\mathbb{E}_Q(-X)$ *denotes the expectation of the random variable $(-X)$ under the probability measure Q.*

In fact, it suffices to choose $\mathcal{Q}$ as the set of all finitely additive set functions on $(\Omega,\mathcal{F})$, which constitutes a subset of all probability measures on that measurable space, cp. Proposition 4.14 in Föllmer and Schied (2004). In the discrete case, the interpretation of $\mathcal{Q}$ is given as a set of *scenarios*. In each of these scenarios, probabilities are assigned to the states of nature. With this interpretation at hand, it makes sense to define any risk measure as the amount of capital to be reserved for the *worst-case scenario* contained in $\mathcal{Q}$, which is exactly the statement of (8).

3.1 Coherent Distortion Risk Measures

For this paper, a subclass of the set of coherent risk measures is of special interest, introduced in the next definition. Coherent distortion risk measures are studied in more detail by Acerbi (2002).

Definition 3.4 *Let $X, Y \in \mathcal{X}$. A coherent risk measure ρ is called a* coherent distortion risk measure *if it fulfills the following two axioms.*

- Law Invariance:

  Let F_X, F_Y denote the respective distribution functions under $\mathbb{P}$. It holds that

$$F_X(x) \;=\; F_Y(x) \;\; \forall\, x \in \mathbb{R} \quad \Longrightarrow \quad \rho(X) \;=\; \rho(Y).$$

- Comonotonicity:

  For all comonotone random variables X and Y, it holds that

$$\rho\,(X \,+\, Y) \;=\; \rho(X) \,+\, \rho(Y).$$

Note that for the axiom of *Comonotonicity*, the usual definition of comonotone random variables applies. Two random variables X and $Y \in \mathcal{X}$ are said to be *comonotone* if

$$\big(X(\omega_1) \,-\, X(\omega_2)\big)\,\big(Y(\omega_1) \,-\, Y(\omega_2)\big) \;\geq\; 0 \quad \forall\,(\omega_1, \omega_2) \in \Omega \times \Omega.$$

Informally, a risk measure that obeys *Law Invariance* assigns the same risk to random variables with the same distribution. A risk measure is thus said to be comonotone when the risk of the sum of two comonotone random variables is strictly additive rather than subadditive. Both properties are useful requirements for a risk measure, *Law Invariance* allows for consistent estimation of the risk measure from data; for an example where this property is violated, see Bertsimas and Brown (2009, p. 6). *Comonotonicity* penalizes non-diversified portfolios, and is a sensible postulation, as comonotone risks can never be used as a hedge against each other.

The special structure of coherent distortion risk measures allows for a more concrete representation. In fact, every coherent distortion risk measure can be written as an expectation of the random variable X with respect to *distorted* probabilities. Before turning to this representation, a proper definition of distortion functions will be given.

Definition 3.5 *A* distortion function $g \,:\, [0,1] \,\longrightarrow\, [0,1]$ *is a non-decreasing function such that* $g(0) = 0$ *and* $g(1) = 1$.

The distortion function, that is, the function that reweighs the original probabilities of the outcomes of X, must be concave in order for the risk measure to be coherent, cp. Föllmer and Schied (2004, Theorem 4.88). Actually, the set function $g \circ \mathbb{P}$ only must satisfy the weaker condition of *submodularity* for the risk measure to be coherent. Whenever g is concave, then $g \circ \mathbb{P}$ is submodular, cp. Föllmer and Schied (2004, Proposition 4.69). The set function $g \circ \mathbb{P}$ will be examined in Remark 3.7 in more detail. More general, any risk measure that can be calculated as the expectation with respect to distorted probabilities is called a distortion risk measure. With this classification, also the VaR at some level α can be called a distortion risk measure. In this paper, only the coherent distortion risk measures are of interest, so the distinction between distortion risk measures and coherent distortion risk measures is seminal.

Now, a special representation theorem for coherent distortion risk measures is given. Theorem 3.3 applies, but can be stated in a more precise form.

Theorem 3.6 *For $X \in \mathcal{X}$, and a concave distortion function g, the functional ρ_g given by*

$$\rho_g(X) \;:=\; \int_{-\infty}^{0} g\big(F_X(x)\big)dx \;+\; \int_{0}^{\infty}\Big(g\big(F_X(x)\big) \;-\; 1\Big)dx \tag{9}$$

is a coherent distortion risk measure in the sense of Definition 3.4.

Remark 3.7 Let a - not necessarily concave - distortion function g be given, and define a *set function* $c_g : \mathcal{F} \longrightarrow [0,1]$ via $c_g \equiv g \circ \mathbb{P}$, that is,

$$c_g(A) \;:=\; g\big(\mathbb{P}(A)\big) \quad \textit{for all} \;\; A \in \mathcal{F}. \tag{10}$$

Then c_g obeys the properties $c_g(\varnothing) \;=\; 0,\; c_g(\Omega) \;=\; 1$, and

$$A \,\subset\, B \quad \Longrightarrow \quad c_g(A) \;\leq\; c_g(B) \;\; \textit{for all} \;\; A, B \in \mathcal{F}.$$

Clearly, c_g is not a probability measure anymore, and is thus called a *capacity*, or a *non-additive measure*. It holds that for any concave distortion function g, the coherent distortion risk measure ρ_g equals the *Choquet integral* of $(-X)$ with respect to the capacity c_g, viz.

$$\rho_g(X) \;=\; \int (-X)\, dc_g \,.$$

For a proof, see Föllmer and Schied (2004, Chapter 4.6). Whenever the capacity c_g is a probability measure, usual integration and Choquet integration coincide, and as such, the above expression is well-defined.

The key issue about (coherent) distortion risk measures is the reweighing of the probabilities of the outcomes, which is accomplished by application of some (concave) distortion function g. Typically, unfavorable outcomes, i.e., losses, are assigned a higher weight than favorable outcomes. This type of reweighing is translated into concavity of g. For the special case of equally likely outcomes of a random variable X with finite support, which is the case for empirical observations of prices, calculation of the coherent distortion risk measure can as well be conducted by using the following theorem.

Theorem 3.8 (Bertsimas and Brown, 2009) *A risk measure ρ is a coherent distortion risk measure if and only if there exists $q \in \Delta^T := \{p \in \mathbb{R}_+^T : \underline{1}'p = 1\}$ with $q_1 \geq \ldots \geq q_T$ (the* generator*), such that*

$$\rho(X) \;:=\; -\sum_{j=1}^{T} q_j \, x_{(j)}, \tag{11}$$

where $x_{(j)}$ denotes the j-th order statistic of the random variable X, that is, $x_{(1)} \leq \ldots \leq x_{(T)}$.

3.2 Application of Risk Measures in Portfolio Selection

Whenever an investor only cares about the downside deviation of his return from a certain threshold return $\bar{\mu}$, measured by some risk measure ρ, his optimization problem reads

$$\max_w w'\mathbb{E}(\tilde{\mu}) \quad \text{s.t.} \;\; \rho(\, w'\tilde{\mu} \,-\, \bar{\mu}\,) \,\leq\, 0, \; w \in \mathcal{C}. \tag{12}$$

As already mentioned in the introduction, the above problem (12) is referred to as the *mean-risk portfolio selection problem*. A comparison of the mean-variance portfolio selection problem with the introduced methodology of portfolio selection via VaR or CVaR is given by Alexander and Baptista (2004).

Problem (12) with the measure ρ given by the CVaR at some level α has generated much interest over the last decade, compare, for example, the primary works by Rockafellar and Uryasev (2000) or Krokhmal et al. (2002). For the risk measure CVaR, the paper by Rockafellar and Uryasev (2000) delivers a solutions to problem (12) by linearizing the risk constraint. The approach taken in this work is different from their way of solving it, and is applicable to the whole set of coherent distortion risk measures.

This does not only open the investment decision to a much broader set of risk measures than only the CVaR. Coherent distortion risk measures are broadly applied in the investment and insurance industry, especially because of their simple structure and their economic justification.

4 Uncertainty Sets Generated by Risk Measures

When avoiding assumptions about the assets' return distribution, the only source of information for the investor is historical data or analysts' forecasts. Both can be coded in a series $(R_1, \ldots, R_T)$, $R_t \in \mathbb{R}^d$ for each t, which in this section shall be the only information that is available to the investor. To be consistent with the literature, let the matrix $R := [R_1 \cdots R_T]$ or the set $\mathcal{R} := \{R_1, \ldots, R_T\}$ be defined as the *data of the problem*. Furthermore, as the mean vector of asset returns μ is the uncertain quantity in the portfolio optimization problem, about which the only source of information is $\mathcal{R}$, let $\tilde{\mu}$ be a d-variate random variable on the finite probability space $(\Omega, 2^\Omega, \mathbb{P})$. Let $\mathcal{Q}^*$ describe the set of all probability measures on $(\Omega, 2^\Omega)$. In this setting, the cardinality of Ω

is T, written $\#\Omega = T$, and $\tilde{\mu}(\omega_i) = R_i$ for each $\omega_i \in \Omega$, so that the support of $\tilde{\mu}$ is $\mathcal{R}$. Moreover, let $\mathbb{P}(\omega_t) = \frac{1}{T}$ for $t = 1, \dots, T$. Consequently, the random variable $\tilde{\mu}$ attains the values $R_1, \dots, R_T$, each with probability $\frac{1}{T}$. As $\#\Omega = T$, it follows that each element of $\mathcal{Q}^*$ is fully described by a vector $q \in \Delta^T$. Let $\mathcal{Q} \subset \Delta^T$ denote the set of all such q. In this setting, all theorems from the section about coherent (distortion) risk measures are applicable.

4.1 Shape of the Uncertainty Set $\mathcal{U}_\rho$

A recently introduced idea for the construction of uncertainty sets for optimization problems of the type (5) can be found in Bertsimas and Brown (2009). In that paper, a main deficiency of previous studies is seen in the circumstance that uncertainty sets are often chosen "*ad hoc, with emphasis usually placed on sets that preserve computational tractability.*" According to the authors, the primary reason for using uncertainty sets in the context of optimization with uncertain input parameters should be to draw a precise picture of the underlying data, though.

Consequently, some coherent measure of risk ρ is taken as a starting point, which at this point is not necessarily a coherent distortion risk measure, and then it is shown that this risk measure induces a convex uncertainty set $\mathcal{U}_\rho$.

To be more precise, Bertsimas and Brown (2009) state the following central result.

Theorem 4.1 (Bertsimas and Brown, 2009) *Let the random return $\tilde{\mu}$ of d risky assets be modeled as above, and let ρ be a coherent risk measure. It holds that*

$$\left\{ w \in \mathbb{R}^d : \rho(\tilde{\mu}'w - \bar{\mu}) \le 0 \right\} = \left\{ w \in \mathbb{R}^d : \mu'w \ge \bar{\mu} \text{ for all } \mu \in \mathcal{U}_\rho \right\}, \tag{13}$$

where $\mathcal{U}_\rho = \text{conv}\left(\left\{ R\,q : q \in \mathcal{Q} \right\}\right)$. $\mathcal{Q} \subset \Delta^T$ is the family of generating measures for ρ, as introduced in Theorem 3.3.

Conversely, if $\mathcal{U} \subset \text{conv}(\mathcal{R})$, then (13) *holds with the coherent risk measure generated by the family of probability measures given by $\mathcal{Q} = \left\{ q \in \Delta^T : \exists \mu \in \mathcal{U} \text{ s.t. } R\,q = \mu \right\}$.*

To apply this approach in the context of portfolio optimization, let there be an investor who is uncertain about the estimated asset returns μ. Thus, he wants to conduct an optimization over a set $\mathcal{U}_\rho$ of possible values for μ, such that the solution w^{opt} stays optimal even for the worst-case scenario μ_*. This means that the risk of the worst-case return $(\mu_*)'w$ falling short of the threshold $\bar{\mu}$ is controlled for.

Let ρ be a coherent risk measure that corresponds to the investor's risk preferences. Then, Theorem 4.1 ensures that the solutions to the following problems coincide.

$$\arg\max_w \; \{ \mathbb{E}(\tilde{\mu})'w \} \quad \text{s.t. } \rho\big(\tilde{\mu}'w - \bar{\mu}\big) \le 0,\; w \in \mathcal{C}, \tag{14}$$

$$\arg\max_w \; \{ \mu'w \} \quad \text{s.t. } \mu'w \ge \bar{\mu} \text{ for all } \mu \in \mathcal{U}_\rho,\; w \in \mathcal{C}, \tag{15}$$

where $\mathcal{U}_\rho$ is the uncertainty set defined in Theorem 4.1.

For a *chance constraint*, i.e., a constraint on the probability of the threshold return being failed, it holds that

$$\begin{aligned}\{w \in \mathbb{R}^d : \mathbb{P}(\tilde{\mu}'w \leq \bar{\mu}) \leq \alpha\} &= \{w \in \mathbb{R}^d : \text{VaR}_\alpha(\tilde{\mu}'w) \leq -\bar{\mu}\} \\ &\supset \{w \in \mathbb{R}^d : \text{CVaR}_\alpha(\tilde{\mu}'w) \leq -\bar{\mu}\},\end{aligned}$$

and as such, a solution w^* to problem (14) with the coherent distortion risk measure $\rho \equiv \text{CVaR}_\alpha$ at threshold return $\bar{\mu}$ will also satisfy the chance constraint $\mathbb{P}(\tilde{\mu}'w \leq \bar{\mu})$. Clearly, though, w^* will not necessarily be a solution to the chance constrained maximization problem

$$\arg\max_w \;\{\, \mu'w \,\} \quad \text{s.t. } \mathbb{P}(\tilde{\mu}'w \leq \bar{\mu}) \leq \alpha.$$

Obviously, when $\mathcal{U}_\rho$ is compact, any solution w^{opt} to problem (15) is also a solution to the problem

$$\arg\max_w \left\{ \{\, \mu'w \,\} \quad \text{s.t. } \min_{\mu \in \mathcal{U}_\rho} \{\, \mu'w \,\} \geq \bar{\mu} \right\}, \quad w \in \mathcal{C}, \tag{16}$$

which again shows the minimax character of these problems.

For the special class of coherent distortion risk measures, it holds that each one is generated by a family of probability distributions

$$\mathcal{Q} = \left\{ \tilde{q} \in \Delta^T : \; \tilde{q}_t \;=\; q_{\sigma(t)}, \; \sigma \in S(T) \right\} \tag{17}$$

for some generator $q \in \Delta^T$ such that $q_1 \geq \ldots \geq q_T$, cp. Theorem 3.8.

Example 4.2 The coherent distortion risk measure *CVaR*$_\alpha$ is generated by $q \in \Delta^T$ with the following entries:

$$q_j \;=\; \begin{cases} \frac{1}{T\alpha} & \text{for } j < T - \lceil T(1-\alpha) \rceil + 1 \\ \frac{\lceil T(1-\alpha) \rceil - T(1-\alpha)}{T\alpha} & \text{for } j = T - \lceil T(1-\alpha) \rceil + 1 \\ 0 & \text{for } j > T - \lceil T(1-\alpha) \rceil + 1, \end{cases} \tag{18}$$

where $\lceil x \rceil$ denotes the smallest integer larger than or equal to x. It is noteworthy that these weights come out as the increments of the distortion function g_{CVaR_α}, defined by

$$g_{\text{CVaR}_\alpha}(u) \;:=\; \begin{cases} \frac{u}{\alpha} & \text{for } u \leq \alpha \\ 1 & \text{for } u > \alpha\,, \end{cases} \tag{19}$$

when evaluated at the points $0, \frac{1}{T}, \frac{2}{T}, \ldots, \frac{T-1}{T}, 1$.

First, the theorem connecting coherent distortion risk measures and their uncertainty sets is given, and afterwards the link towards the concept of *weighted-mean trimmed regions* is explained.

Theorem 4.3 (Bertsimas and Brown, 2009) *If a risk measure ρ is a coherent distortion risk measure with generator q, the following equality holds:*

$$\left\{w \in \mathbb{R}^d : \rho\left(\tilde{\mu}'w - \bar{\mu}\right) \leq 0\right\} = \left\{w \in \mathbb{R}^d : \mu'w \geq \bar{\mu} \text{ for all } \mu \in \mathcal{U}_\rho\right\}, \tag{20}$$

where the uncertainty set $\mathcal{U}_\rho$ takes the form

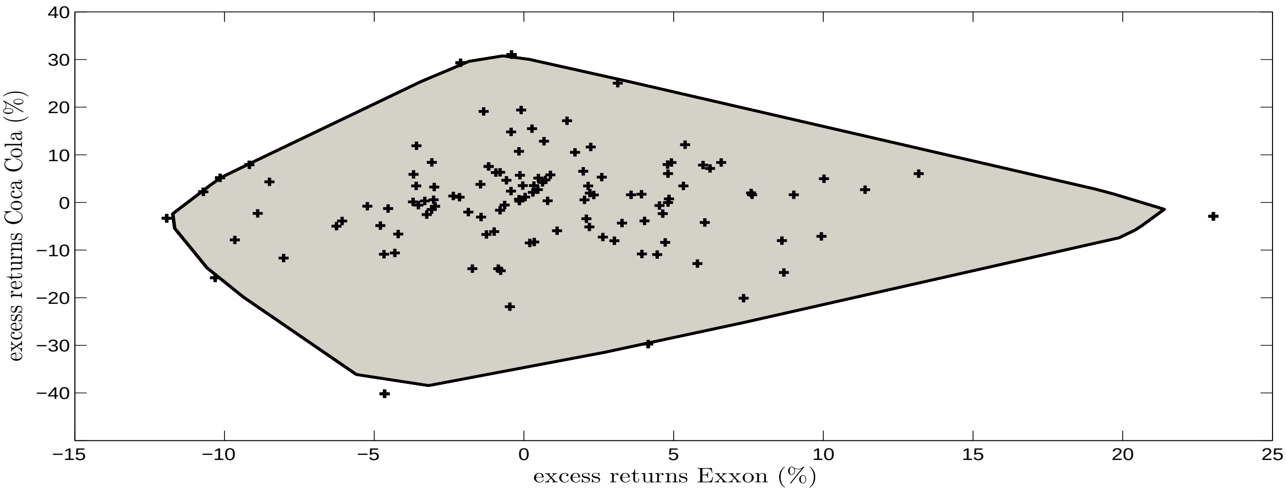
$$\mathcal{U}_\rho := \mathit{conv}\left(\left\{\sum_{t=1}^{T} q_{\sigma(t)} R_t : \sigma \in S(T)\right\}\right). \tag{21}$$

Definition 4.7 in Bertsimas and Brown (2009) calls $\mathcal{U}_\rho$ as above the *q-permutohull* generated by the data $\mathcal{R}$ and the coherent distortion risk measure ρ. The interpretation of the linkage between the generator q of ρ and the size of $\mathcal{U}_\rho$ is as follows.

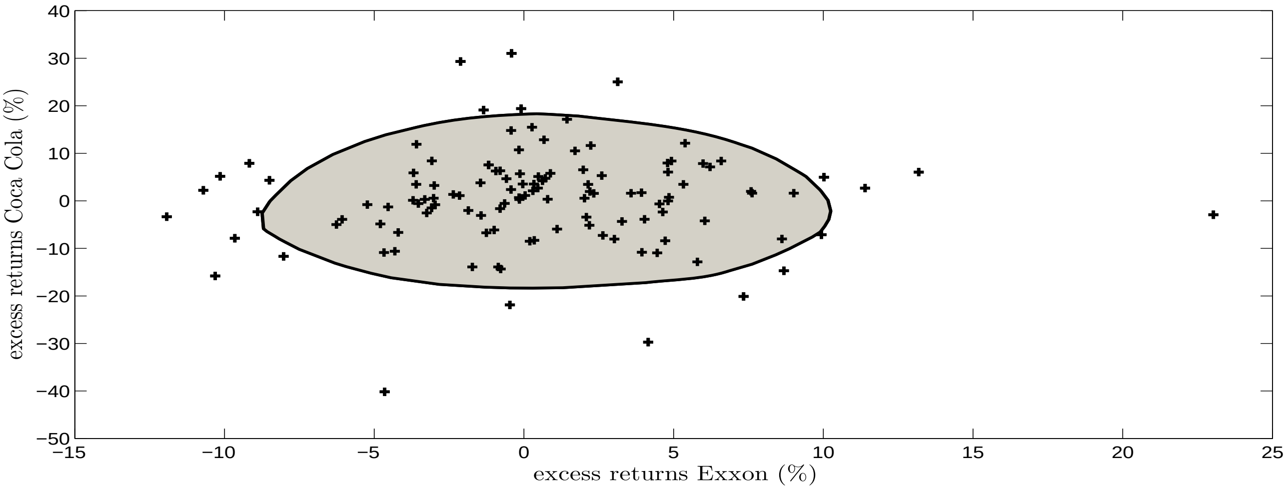
An investor who is uncertain about the mean of the parameter $\tilde{\mu}$ wishes to control for the underperformance of his portfolio return $\bar{\mu}$ by the application of a risk measure ρ. In the equivalent formulation with the risk constraint, i.e., $\rho(\tilde{\mu}'w \leq \bar{\mu}) \leq 0$, the risk measure can be chosen such that a large set of possible outcomes of $\tilde{\mu}$ is covered. In the case of ρ being a coherent distortion risk measure, this is indicated by choosing the generator q in such a way that it puts large probability masses on few data records only.

Conversly, let the investor be rather confident in the sample estimate $\hat{\mu}$ of $\tilde{\mu}$. This implies the choice of an investor-specific coherent distortion risk measure ρ such that the associated uncertainty set is smaller. In this case, the generator of ρ should not put too much emphasis on single observations from the past, but rather smooth out the historical observations via similar probability masses for all entries of q. In the case of $\rho \equiv \mathrm{CVaR}_\alpha$, the first case is expressed by choosing a small α, while in the latter case, the investor chooses larger values for α. A visualization of the connection between a risk measure for an *uncertain* investor and a risk measure for a more *confident* investor with the respective uncertainty sets is presented in Figures 1(a) and 1(b), respectively. While Figure 1(a) includes nearly all historical return observations, Figure 1(b) excludes some of them, yielding a smaller set.

The uncertainty sets $\mathcal{U}_\rho$, generated via (21) from a coherent distortion risk measure with generator q and data $\mathcal{R}$, take a special form of the recently introduced *weighted-mean trimmed regions*, denoted $D_\beta(\mathcal{R})$, as defined by Dyckerhoff and Mosler (2010b). For D_β, the trimming parameter $\beta \in]0,1]$ defines the *depth* of the region, and is connected to the generator q in the case of ρ being a coherent distortion risk measure. The definition of weighted-mean trimmed regions allows for all types of weights that a generator of a coherent distortion risk measure might consist of, cp. Dyckerhoff and Mosler (2010b, Definition 2). A more practical expression for the calculation of its extreme points is given in their Proposition 2, while an algorithm for the calculation of these points is already made available by Bazovkin and Mosler (2010).



(a) Uncertainty set $\mathcal{U}_{\text{CVaR}}$ corresponding to $\text{CVaR}_{1\%}$



(b) Uncertainty set $\mathcal{U}_{\text{CVaR}}$ corresponding to $\text{CVaR}_{10\%}$

Figure 1: Uncertainty sets $\mathcal{U}_\rho$ for $\rho \equiv \text{CVaR}_\alpha$ with different levels of α. The data for both uncertainty sets consists of $T = 120$ data points, representing the monthly excess returns of the Exxon Mobil Corp. and the Coca Cola Enterprises. With growing parameter α, the uncertainty set $\mathcal{U}_{\text{CVaR}}$ decreases.

Remark 4.4 It is noteworthy that *nestedness* of the respective regions is a property common to all uncertainty sets generated by coherent distortion risk measures. In fact, Proposition 3 in Dyckerhoff and Mosler (2010b) implies that whenever the generators q^1 and q^2 of two coherent distortion risk measures ρ^1 and ρ^2 are such that

$$\sum_{j=1}^{k} q_j^1 \;\leq\; \sum_{j=1}^{k} q_j^2 \quad \text{for all } \; k = 1, \ldots, T,$$

then the respective uncertainty sets of the form (21) are *nested*, that is,

$$\mathcal{U}_{\rho^1} \;\subset\; \mathcal{U}_{\rho^2}\,.$$

Several other important analytical features are presented in Dyckerhoff and Mosler (2010b) as well, along with examples for different kinds of such regions. The connection of weighted-mean trimmed regions and probability distributions is explored in Dyckerhoff and Mosler (2010a). It is noteworthy that the q-permutohulls from Theorem 4.3 are equivalent to the special case where the weighted-mean trimmed regions take the form of *zonoid regions*, introduced earlier by Koshevoy and Mosler (1997).

Formalizing (14) for some coherent distortion risk measure ρ, and leaving the set of constraints $\mathcal{C}$ aside for the moment, one obtains the expression

$$\max\ \mu' w \quad \text{s.t. } w \in \mathcal{W}_{\bar{\mu}},\ \mu \in \mathcal{U}_\rho, \tag{22}$$

where

$$\mathcal{W}_{\bar{\mu}} := \left\{ w \in \mathbb{R}^d :\ w'\mu \geq \bar{\mu} \text{ for all } \mu \in \mathcal{U}_\rho \right\}, \quad \text{and} \tag{23}$$

$$\mathcal{U}_\rho = \text{conv}\left(\left\{ \sum_{t=1}^{T} q_{\sigma(t)}\, R_t :\ \sigma \in S(T) \right\} \right). \tag{24}$$

Here, q denotes the generator for ρ, so $\mathcal{U}_\rho$ is generated by q and R_t, with $t = 1, \ldots, T$, wich are the data available to the investor.

The uncertainty set $\mathcal{U}_\rho$ is convex by definition. In the case of a coherent distortion risk measure ρ, (24) indicates that $\mathcal{U}_\rho$ is given as the convex hull of a finite set of points. As such, the set of extreme points of $\mathcal{U}_\rho$, denoted by $\mathcal{U}_\rho^e$, must be finite and contained in the set of points generated by q and $\mathcal{R}$, viz.

$$\mathcal{U}_\rho^e \subset \left\{ \sum_{i=1}^{T} q_{\sigma(t)}\, R_t :\ \sigma \in S(T) \right\}. \tag{25}$$

Problem (22) is a nonlinear problem, as μ and w are entangled with each other. In the sequel, a way of solving this problem is presented, making use of the special forms of the uncertainty set $\mathcal{U}_\rho$ and the set $\mathcal{W}_{\bar{\mu}}$.

4.2 Shape of the Set $\mathcal{W}_{\bar{\mu}}$

Before turning to the set $\mathcal{W}_{\bar{\mu}}$, the following concepts from the geometry of convex sets shall be reviewed, as they turn out to be closely connected to the sets $\mathcal{U}_\rho$ and $\mathcal{W}_{\bar{\mu}}$. A standard textbook about convex analysis is Rockafellar (1970).

Let $C \in \mathbb{R}^d$ be a closed, convex set. If C is bounded (and thus compact), C is given as the convex hull of its extreme points C^e, that is, $C = \text{conv}\,(C^e)$.

Definition 4.5 *For any closed set $K \subset \mathbb{R}^d$, the* polar set *K^* is defined as*

$$K^* := \{y \in \mathbb{R}^d :\ y'x \leq 1 \text{ for all } x \in K\}. \tag{26}$$

The polar set K^* of some set K is thus the intersection of all halfspaces generated by the elements of K. Note that for a particular element $k \in \mathbb{R}^d$, the boundary of the halfspace

$$\{k\}^* = \{y \in \mathbb{R}^d : \ y'k \leq 1\}$$

is orthogonal to k, and $\underline{0}$ is contained in $\{k\}^*$.

A characterization of the polar set of some closed, convex set is given now.

Proposition 4.6 *Let $K \subset \mathbb{R}^d$ be closed and convex.*

1. *K^* is bounded if and only if $\underline{0} \in \mathrm{int}(K)$.*

2. *K is bounded if and only if $\underline{0} \in \mathrm{int}(K^*)$.*

Moreover, as any compact and convex set K is given by the convex hull of its extreme points, $K = \mathrm{conv}(K^e)$, it holds that K^* is given as the polar set of the set K^e of extreme points of K, viz.

$$K^* \ = \ \{y \in \mathbb{R}^d : \ y'x^e \leq 1 \text{ for all } x^e \in K^e\} \ = \ \bigcap_{x^e \in K^e} \{x^e\}^*. \tag{27}$$

Clearly, if the set of extreme points is finite, i.e., $\#K^e = n < \infty$, then the intersection of the halfspaces given by the polars of the n extreme points entirely describes K^*. Burdet (1973) gives a generalization of polar sets along with some results.

Definition 4.7 (Burdet 1973) *Let P denote a closed set in $\mathbb{R}^d$. For a given paramter $\kappa \in \mathbb{R}$, define the* polaroid *$P^*(\kappa)$ by*

$$P^*(\kappa) \ = \ \{ \ y \in \mathbb{R}^d : \ x'y \ \leq \ \kappa \ \text{ for all } x \in P \ \}.$$

For any closed set $K \subset \mathbb{R}^d$, it holds that its polar set coincides with the polaroid at $\kappa = 1$, i.e., $K^* = K^*(1)$.

Theorem 4.8 (Burdet 1973) *For any closed sets P and Q in $\mathbb{R}^d$, one has the following implications:*

1. *If $Q \subset P$, then $Q^*(\kappa) \supset P^*(\kappa)$.*

2. *If $\kappa_1 \geq \kappa_2$, then $P^*(\kappa_1) \supset P^*(\kappa_2)$.*

For $\kappa \leq 1$, Theorem 4.8 ensures that $P^*(\kappa) \subset P^*(1) = P^*$, and using the assertion from Proposition 4.6 it is immediate to obtain

Lemma 4.9 *Let $P \in \mathbb{R}^d$ be closed and convex, and let $\kappa \in \mathbb{R}$ such that $\kappa \leq 1$.*

$$P^*(\kappa) \text{ is bounded if and only if } \underline{0} \in \mathrm{int}(P).$$

These general observations and concepts are now directly applicable to the portfolio optimization problem as given by (22). The connection between $\mathcal{U}_\rho$ and $\mathcal{W}_{\bar\mu}$ is formalized by

Lemma 4.10 *It holds that* $\mathcal{W}_{\bar\mu} \;=\; -\big(\mathcal{U}^*_\rho(-\bar\mu)\big)$,

which is easily proved by noting that

$$\begin{aligned}
\mathcal{W}_{\bar\mu} &= \{\, w \in \mathbb{R}^d :\; w'\mu \;\geq\; \bar\mu \;\text{ for all } \mu \in \mathcal{U}_\rho \,\} \\
&= \{\, -w \in \mathbb{R}^d :\; w'\mu \;\leq\; -\bar\mu \;\text{ for all } \mu \in \mathcal{U}_\rho \,\} \\
&= -\{\, w \in \mathbb{R}^d :\; w'\mu \;\leq\; -\bar\mu \;\text{ for all } \mu \in \mathcal{U}_\rho \,\} \;=\; -\big(\mathcal{U}^*_\rho(-\bar\mu)\big)\;.
\end{aligned}$$

With the connection of $\mathcal{U}_\rho$ and $\mathcal{W}_{\bar\mu}$ at hand it is immediate to obtain the following result in the light of portfolio optimization with risk constraints given by a coherent distortion risk measure ρ from the previous statements.

Proposition 4.11 *Given a compact and convex uncertainty set $\mathcal{U}_\rho$, the set of feasible portfolio weight vectors $\mathcal{W}_{\bar\mu}$ has the following properties:*

1. *$\mathcal{W}_{\bar\mu}$ is convex and closed.*
2. *For $\bar\mu \geq -1$, $\mathcal{W}_{\bar\mu}$ is bounded, and as such, compact, if and only if $\underline{0} \in \operatorname{int}\mathcal{U}_\rho$.*
3. *For $\bar\mu_1 \geq \bar\mu_2$, it holds that $\mathcal{W}_{\bar\mu_2} \supset \mathcal{W}_{\bar\mu_1}$.*
4. *For the uncertainty sets $\mathcal{U}^1_\rho$ and $\mathcal{U}^2_\rho$, let $\mathcal{W}^1_{\bar\mu}$ and $\mathcal{W}^2_{\bar\mu}$ denote the respective sets of feasible portfolio weight vectors for threshold return $\bar\mu$. It holds that*

$$\mathcal{U}^1_\rho \subset \mathcal{U}^2_\rho \;\Longrightarrow\; \mathcal{W}^1_{\bar\mu} \supset \mathcal{W}^2_{\bar\mu}.$$

From a practical point of view, Statement 3 in the above proposition can be made more precise. When the threshold return $\bar\mu$ is multiplied by some constant number $c > 0$, then the new set of feasible portfolio weight vectors $\mathcal{W}_{c\bar\mu}$ is a multiple of the original set, $\mathcal{W}_{\bar\mu}$. Leaving other constraints on the set of feasible portfolio weight vectors aside for the moment, it is an immediate consequence that the the new optimal portfolio will thus be a multiple of the old portfolio. This observation is formalized in

Lemma 4.12 *Let $\mathcal{W}_{\bar\mu}$ be the set of feasible portfolio weight vectors associated with the uncertainty set $\mathcal{U}_\rho$, and let $c > 0$. Then*

$$\mathcal{W}_{c\bar\mu} \;=\; c\,\mathcal{W}_{\bar\mu}.$$

Clearly, for practical portfolio applications, this observation will be of use only in most exceptional cases. Whenever additional constraints are in place, optimal portfolios cannot easily be in- or deflated by some scalar.

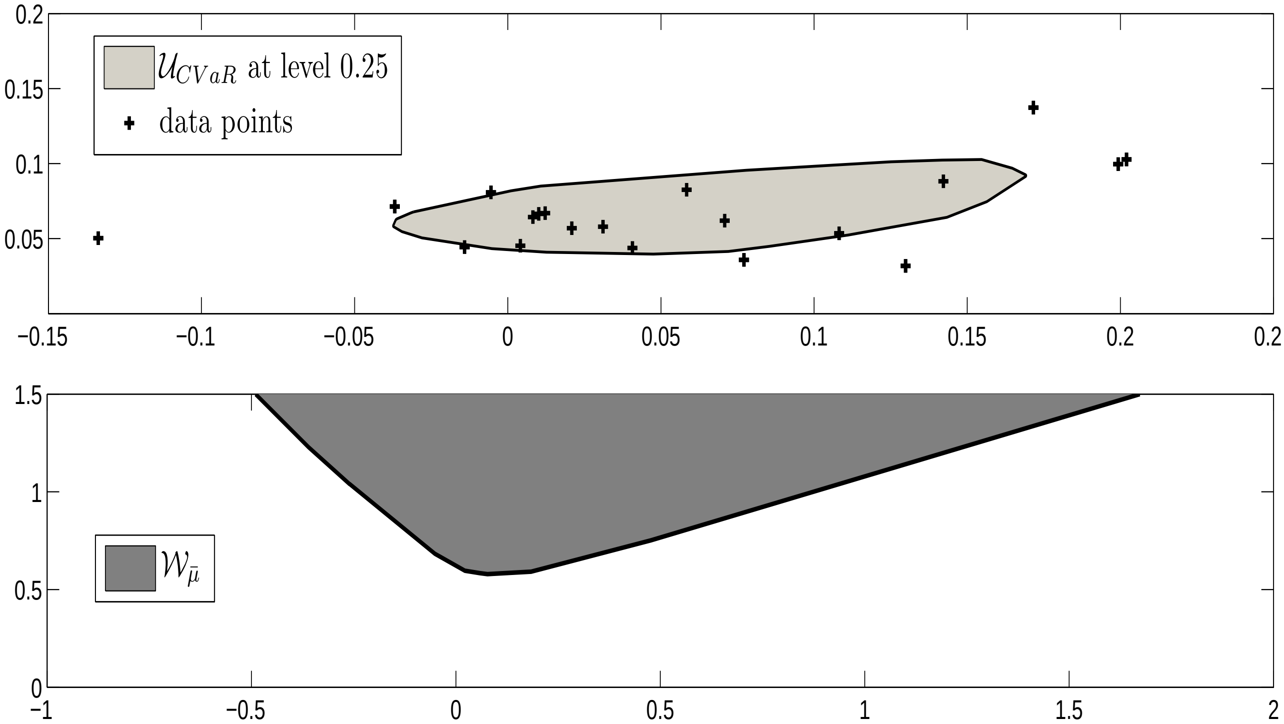


Figure 2: For the shape of the uncertainty set $\mathcal{U}_\rho$ in the upper figure, the convex set $\mathcal{W}_{\bar{\mu}}$ of feasible portfolio weight vectors becomes unbounded.

Note that Lemma 4.10 already yields the a way of constructing $\mathcal{W}_{\bar{\mu}}$. It holds that

$$\begin{aligned} -\left(\mathcal{U}_\rho^*(-\bar{\mu})\right) &= -\{\, w \in \mathbb{R}^d :\ w'\mu \ \leq\ -\bar{\mu} \ \text{ for all } \mu \in \mathcal{U}_\rho \,\} \\ &= \{\, w \in \mathbb{R}^d :\ w'(-\mu^e) \ \leq\ -\bar{\mu} \ \text{ for all } \mu^e \in \mathcal{U}_\rho^e \,\} \\ &= \bigcap_{j=1}^{n} \left(\{-\mu^{e,j}\}\right)^* (-\bar{\mu}), \end{aligned}$$

indicating that $\mathcal{W}_{\bar{\mu}}$ is the intersection of halfspaces generated by the n extreme points of $\mathcal{U}_\rho$ and the threshold return $\bar{\mu}$. By slight abuse of notation, let $-\mathcal{U}_\rho^e$ also denote the $(n \times d)$-matrix that contains the negative of these n extreme points as rows, so it holds that

$$\mathcal{W}_{\bar{\mu}} \ =\ \{\, w \in \mathbb{R}^d :\ -\mathcal{U}_\rho^e\, w \ \leq\ -\bar{\mu}\underline{1} \,\}. \tag{28}$$

The set $\mathcal{W}_{\bar{\mu}}$ is convex, but not necessarily bounded, as indicated by Proposition 4.11. Note that the structure of $\mathcal{W}_{\bar{\mu}}$ depends on the given uncertainty set $\mathcal{U}_\rho$ and on $\bar{\mu}$, the threshold return. Of course, for an unbounded set $\mathcal{W}_{\bar{\mu}}$, there does not necessarily exist a solution to the portfolio optimization problem (22).

An example for an unbounded set of feasible portfolio weight vectors, $\mathcal{W}_{\bar{\mu}}$, along with the respective uncertainty set $\mathcal{U}_\rho$ with the risk measure ρ given as the CVaR at level 0.25, is given in Figure 2.

In practical situations, though, the set of feasible portfolio weight vectors is always bounded by some constraints $\mathcal{C}$. Most prominently, it is bounded below by the prohibition of short sales, which means $w \geq 0$. An upper bound is frequently given by $w_i \leq 1$, or more general, by $w_i \leq \zeta_i$, $\zeta_i \in \mathbb{R}$, for each $i = 1, \ldots, d$. The budget constraints, i.e., the requirement that all of the capital is invested among the assets, reads $w'\underline{1} \leq 1$.

For the following discussion, $\mathcal{C}$ is assumed to be compact and convex, and that the number of restrictions in $\mathcal{C}$ is finite, which is a realistic assumption.

As such, the intersection $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ constitutes the set of all weight vectors that fulfill all constraints for the portfolio optimization problem (22). To distinguish the sets $\mathcal{W}_{\bar{\mu}}$ and $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ in the sequel, the latter is called the set of *all practically feasible portfolio weight vectors*. As an intersection of closed and convex sets, $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ is closed and convex, and its boundedness follows from the boundedness of $\mathcal{C}$.

Figure 3 demonstrates the effect of adding a a compact constraint set to $\mathcal{W}_{\bar{\mu}}$, yielding a compact and convex set of all practically feasible portfolio weight vectors.

The set $\mathcal{W}_{\bar{\mu}}$ of feasible portfolio weight vectors is given as the intersection of a finite number n of halfspaces, as described by (28). Note that by the above assumptions, the set $\mathcal{C}$ of constraints can as well be written as the intersection of a finite number of halfspaces, so

$$\mathcal{C} \;=\; \{\, w \in \mathbb{R}^d : \; Cw \;\leq\; \gamma \,\} \tag{29}$$

for some appropriate $(M \times d)$-matrix C, and a vector $\gamma \in \mathbb{R}^M$. Then, the set $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ of all practically feasible portfolio weight vectors is given by

$$(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C}) \;=\; \{\, w \in \mathbb{R}^d : \; Aw \;\leq\; b \,\}, \tag{30}$$

where

$$A := \begin{bmatrix} -\mathcal{U}_\rho^e \\ C \end{bmatrix} \in \mathbb{R}^{(n+M)\times d} \quad \text{and} \quad b := \begin{bmatrix} -\bar{\mu}\underline{1} \\ \gamma \end{bmatrix} \in \mathbb{R}^{(n+M)}. \tag{31}$$

4.3 Methodology for the Portfolio Optimization Problem

With the descriptions of the uncertainty set $\mathcal{U}_\rho$ and the set of all practically feasible portfolio weight vectors, $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, it is now straightforward to find the solution to the portfolio optimization problem (14).

To this end, let $\mathcal{V}$ denote the set of all scalar products of elements of $\mathcal{U}_\rho$ and elements of $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, i.e.

$$\mathcal{V} \;:=\; \{\, \mu' w : \; \mu \in \mathcal{U}_\rho, \; w \in (\mathcal{W}_{\bar{\mu}} \cap \mathcal{C}) \,\} \;\;\subset\; \mathbb{R}. \tag{32}$$

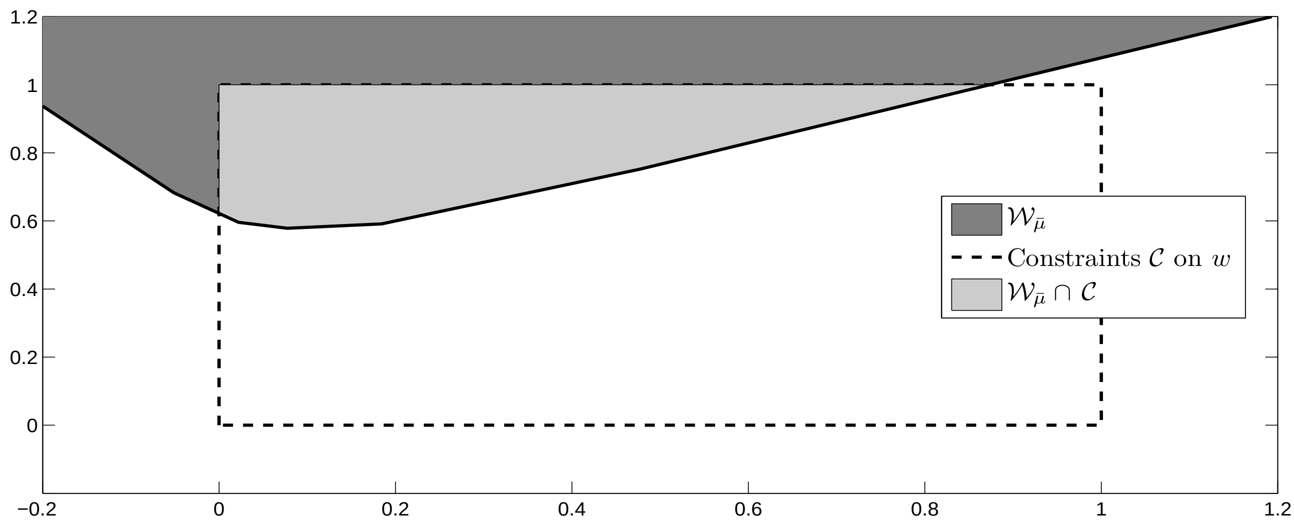


Figure 3: The set of all feasible portfolio weight vectors is always bounded by some constraints $\mathcal{C}$, which, in this case, are given by $\{ w \in \mathbb{R}^2 : \; w \geq \underline{0}, \; w \leq \underline{1} \}$. The convex set generated by intersection of $\mathcal{C}$ and $\mathcal{W}_{\bar{\mu}}$ is the set of all practically feasible portfolio weight vectors, $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$.

As $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ and $\mathcal{U}_\rho$ are convex sets, they are connected. The scalar product is a continuous function of elements of both sets, and as such, the set $\mathcal{V}$ is connected, and as a connected subset of $\mathbb{R}$, $\mathcal{V}$ is convex. Compactness of $\mathcal{V}$ follows from compactness of both $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$ and $\mathcal{U}_\rho$.

As a convex and compact set in $\mathbb{R}$, $\mathcal{V}$ possesses a maximum, denoted by m, and the following holds true by the property $\mathcal{U}_\rho = \text{conv}\left(\mathcal{U}_\rho^e\right)$ and by convexity of $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$:

$$
\begin{aligned}
m &= \max \left\{ \mu' w : \; \mu \in \mathcal{U}_\rho, \; w \in (\mathcal{W}_{\bar{\mu}} \cap \mathcal{C}) \right\} \\
&= \max \left\{ (\mu^e)' w : \; \mu^e \in \mathcal{U}_\rho^e, \; w \in (\mathcal{W}_{\bar{\mu}} \cap \mathcal{C}) \right\} \\
&= \max \left\{ (\mu^e)' w : \; \mu^e \in \mathcal{U}_\rho^e, \; Aw \leq b \right\} \\
&= \max \left\{ \underbrace{\max_w \left\{ (\mu^{e,j})' w \text{ s.t. } Aw \leq b \right\}}_{\text{linear optimization problem for each } j} \text{ for } \mu^{e,j} \in \mathcal{U}_\rho^e, \; j = 1, \ldots, \#\mathcal{U}_\rho^e \right\}.
\end{aligned}
$$

Denoting by μ^{opt} and w^{opt} the elements from $\mathcal{U}_\rho^e$ and $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, respectively, such that $(\mu^{opt})' w^{opt} = m$, it follows that w^{opt} is the solution to (14).

Thus, the portfolio optimization problem with a risk constraint given by a coherent distortion risk measure can always be written as a set of linear optimization problems. Of course, the number of constraints described by $Aw \leq b$ grows rapidly with the dimensionality of the problem and with the number of data collected. For given dimensionality d and data size T, it depends heavily on the risk measure chosen for the risk constraint, which in turn induces the uncertainty set and the number of its extreme points.

However, with the advent of an exact and efficient algorithm for the construction of weighted-mean trimmed regions, see Bazovkin and Mosler (2010), the problem of optimizing some linear objective function with respect to a risk constraint is solved by this procedure for the entire class of coherent distortion risk measures.

Moreover, for the real-life situation where short sales are prohibited, the number of extreme points to be examined can be drastically reduced. Note that when $w \geq \underline{0}$ is enforced, all extreme points μ_l^e of $\mathcal{U}_\rho$ that fulfill the componentwise comparison $\mu_l^e \geq \mu^e$ for any other $\mu^e \in \mathcal{U}_\rho^e$ need not be accounted for. Rough introspection of Figures 1(a) and (b) give rise to the general presumption that when short sales are prohibited, only $2^{-d}(\#\mathcal{U}_\rho^e)$ of the extreme points of $\mathcal{U}_\rho \subset \mathbb{R}^d$ need to be visited. Moreover, in this case the number of restrictions defining $\mathcal{W}_{\bar{\mu}}$ is reduced by the same factor. Applications of this methodology to the problem of portfolio optimization with respect to observed financial data as well as simulated data are given in the empirical section.

5 Empirical Study

In this section, the above developed methodology is applied to both financial and simulated data. As the CVaR is the most commonly used coherent risk measure for decision makers in the financial industry, it serves as the starting point for the construction of the uncertainty sets with empirical data. In the section with simulated data, the compatibility of the approach with any coherent distortion risk measure is demonstrated by using the risk measures arising via the Proportional Hazard Transform and the Wang Transform, see Section 3.1.

5.1 Mean-Risk Portfolio Optimization on Financial Data

In the last decade, investments in the so-called BRIC countries have gained interest. The BRIC countries are Brazil, Russia, India and China, and the growth of their respective economies is deemed to outperform the developed countries' growth.

Let a Euroland investor consider to hold a portfolio of the BRIC currencies, that are, the Brazilian Real, the Russian Ruble, the Indian Rupee and the Chinese Yuan. The object of interest is the returns of the respective Euro exchange rates of these currencies. They are retrieved from the European Central Bank (2010), and as such, $T = 60$ bilateral monthly exchange rates are obtained and their respective monthly returns are calculated, yielding a series $(R_1, \ldots, R_{60})$ of 4-dimensional data points. Summary statistics for the BRIC data set are given in Table 1.

As the mean return of the exchange rates is negative for two of the currencies, the Brazilian Real and the Indian Rupee, the investor cannot necessarily expect a solution for his portfolio optimization problem whenever the benchmark return is set to a positive number. He can rather impose a risk constraint in order to establish a lower bound which the currency portfolio should not underperform, as assured by the CVaR at a level of his choice.

Let the investor impose the constraint that his loss should be at most 2.8%, as measured by CVaR$_\alpha$, where α takes the values 1%, 2%, ..., 10%. The respective uncertainty sets $\mathcal{U}_{\text{CVaR}}$ are calculated using Mosler et al. (2009), and Lemma 4.10 gives the construction of the respective set of feasible portfolio weight vectors $\mathcal{W}_{\bar{\mu}}$.

	Brazilian Real	Russian Ruble	Indian Rupee	Chinese Yuan
mean (%)	-0.1880	0.3761	-0.2585	0.0450
minimum (%)	-6.2997	-3.1845	-4.7214	-7.3088
maximum (%)	13.2234	11.6842	6.6472	6.0437
std. dev. (%)	3.5554	2.6577	2.5184	2.6369

Table 1: Summary statistics of monthly returns of the exchange rates for the respective BRIC currencies to the Euro. The dataset consists of the exchange rate returns ranging from November 2005 through October 2010, yielding $T = 60$ data points of dimension $d = 4$.

In addition, the portfolio weight restrictions are given by the short sale constraint and the budget constaint, yielding $4+2$ additional inequality constraints on the set of all practically feasible portfolio weight vectors. For each α, the portfolio optimization problem to be solved thus reads

$$\max_{w} \; \left\{ \big(\mathbb{E}(\tilde{\mu})\big)' w \right\} \quad \text{s.t. } \mathrm{CVaR}_{\alpha}\big(\,\tilde{\mu} w\,\big) \;\leq\; 0.028,\; w \geq 0,\; w'\underline{1} \;=\; 1.$$

For each level of α between 1% and 10%, the extreme points of the uncertainty sets $\mathcal{U}_{\text{CVaR}}$ are obtained. In this setting, with $T = 60$ and $d = 4$, the number of extreme points for each value of α is shown in Table 2. It is noteworthy that whenever αT is an integer, the number of extreme points is substantially smaller than for non-integer values of αT. In Table 2, this can be seen for α =5% and α =10%. This observation is in line with the theory of weighted-mean trimmed regions, cp. Bazovkin and Mosler (2010).

In the equivalent formulation as a set of linear optimization problems, not all extreme points of the respective uncertainty sets need to be checked for optimality. As short sales are not allowed, all extreme points of $\mathcal{U}_{\text{CVaR}}$ that are greater (by componentwise comparison) than some other extreme point can be discarded. As a result, the numbers of extreme points to be checked vary between 9 (for α =1%) and 324 (for α =9%).

The set of all practically feasible portfolio weight vectors, $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, given by the matrix A and the vector b as described by (31), completes the setup for the optimization. Thus, the number of constraints is given by the number of rows of A, which in this case equals at most $324 + 4 + 2 = 330$, and the number of optimizations subject to these constraints is at most 324.

The optimal investment policies for the investments in the currencies of Brazil, Russia, India and China, respectively, are presented in Figure 4. For each level of α, the optimal portfolio subject to the CVaR_{α} risk constraint is dominated by the investment into the Russian Ruble. It is interesting to observe, though, that for the smallest values of α, a good portion of the portfolio weights goes into the Chinese Yuan and the Brazilian Real. An explanation for this phenomenon can be traced back to the fact that the set of all practically feasible portfolio weight vectors is bounded also by the constraint set $\mathcal{C}$.

α	1%	2%	3%	4%	5%	6%	7%	8%	9%	10%
$\#\mathcal{U}^e_{\text{CVaR}}$	24	244	244	853	357	2 099	6 954	6 932	6 897	3 153

Table 2: For the BRIC dataset with $T = 60$ data points of dimension $d = 4$, and each of the given levels of α, the number of extreme points of the respective uncertainty set $\mathcal{U}_{\text{CVaR}}$ is shown.

For growing α, the uncertainty set $\mathcal{U}_{\text{CVaR}_\alpha}$ shrinks towards the singleton given by the sample mean, $\{\hat{\mu}\}$. As the set of feasible portfolio weight vectors, $\mathcal{W}_{\bar{\mu}}$, is the negative of the polaroid of $\mathcal{U}_{\text{CVaR}_\alpha}$ at threshold return $(-\bar{\mu})$, cp. Lemma 4.10, it grows for larger values of α, as ensured by Proposition 4.11.

Clearly, a solution w^{opt} for the mean-risk portfolio optimization problem is always found on the boundary of $(\mathcal{W}_{\bar{\mu}} \cap \mathcal{C})$, and once the set $\mathcal{W}_{\bar{\mu}}$ grows larger than the set $\mathcal{C}$ in some direction, a corner of $\mathcal{C}$ is attained as solution. This transition towards a corner value for the investment decision can be seen to happen for the investment into the Brazilian Real and the Russian Ruble in the above example. For values of α greater than 4%, the portfolio weight for the Brazilian Real declines from 15% to only 2%, and it is clear that for even larger values of α, the portfolio will entirely consist of the Russian Ruble.

This observation leads to the conclusion that a higher aversion against risk leads - at least to some extent - to a diversified portfolio, which is a well-documented fact in the finance literature.

5.2 Mean-Risk Portfolio Optimization on Simulated Data

The optimization methodology for the portfolio optimization problem with a risk constraint given by a coherent distortion risk measure is now illustrated on simulated data. More exactly, $T = 12$ data points from a 3-variate t-distribution with mean vector $\mu = 0.03$, a prespecified covariance structure Σ, and 3.6 degrees of freedom are simulated.

In a first step, the coherent distortion risk measures ρ_{PHT} arising by applying the Proportional Hazard Transform, defined by

$$g_\lambda(u) \;:=\; u^{\frac{1}{\lambda}}. \tag{33}$$

with parameter $\lambda = 2$ is used. In contrast to the CVaR risk measure used in the preceding paragraph, all $T = 12$ entries of the generator q are non-zero. In this setting, even for these small numbers of T and especially d, as many as $\#\mathcal{U}^e_{\rho_{\text{PHT}}} = 3\,850$ extreme points constitute the respective uncertainty set. With the set $\mathcal{C}$ of additional constraints given again as $\{w \in \mathbb{R}^3 : w \geq \underline{0},\ w'\underline{1} = 1\}$, only 125 of the original extreme points need to be considered in the optimization. For a benchmark return of $\bar{\mu} \leq 0.0257$, the optimal portfolio is found as the corner solution $w^{opt} = (1 \quad 0 \quad 0)'$, whereas for values of $\bar{\mu} > 0.0257$, there is no solution to the portfolio optimization problem.

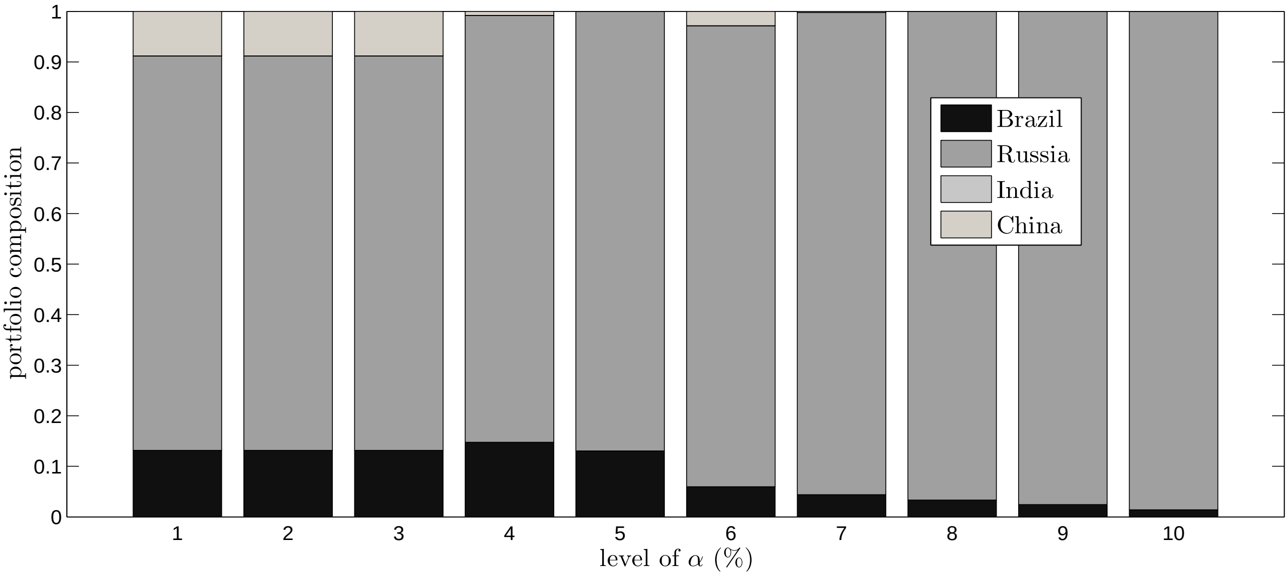


Figure 4: Portfolio weights for the investments in the respective BRIC countries' currencies. The respective risk constraint is given by the CVaR for different levels of α, as given on the horizontal axis. Note that the Indian Rupee is not invested into at any level of α, whereas the Russian Ruble dominates each of the portfolios.

When using the Proportional Hazard Transform with parameter $\lambda = 4$, the resulting uncertainty set is given by the convex hull of 4 615 extreme points, of which only 137 need to be included in the optimization. For $\bar{\mu} \leq 0.021$, the resulting optimal portfolio is given by the corner solution $w^{opt} = (1 \quad 0 \quad 0)'$, whereas for $\bar{\mu} > 0.021$, no feasible point is found.

This effect is in line with the theory for coherent distortion risk measures from Section 3. A larger λ induces a generator q that puts larger probability mass on the worst observed scenario, cp. Theorem 3.8. To provide some intuition, for $\lambda = 4$, the first 3 entries of that generator are $(q_1 \quad q_2 \quad q_3)' = (0.54 \quad 0.1 \quad 0.07)'$, whereas for $\lambda = 2$, these entries amount to $(0.29 \quad 0.12 \quad 0.09)'$. As the condition given in Remark 4.4 is met when comparing the generators for $\lambda = 2$ and $\lambda = 4$, the uncertainty set grows in the parameter λ. By Proposition 4.11, the set of all practically feasible portfolio weight vectors therefore shrinks.

The second coherent distortion risk measure considered on the above data is ρ_{WT}, for which the Wang Transform (34) with $\gamma = 1$ is used for distorting the original probabilities, viz.

$$g_\gamma(u) \quad := \quad \Phi\big(\Phi^{-1}(u) + \gamma\big), \tag{34}$$

where Φ denotes the cumulative distribution function of the standard normal distribution. Again, the generator q has 12 non-zero entries, and as such, the calculated uncertainty set is given as the convex hull of $\#\mathcal{U}^e_{\rho_{\mathrm{WT}}} = 3\,852$ extreme points. As above, by the additional constraint, only 151 of these points need to be considered, and again, the corner solution $w^{opt} = (1 \quad 0 \quad 0)'$ is obtained for every $\bar{\mu} \leq 0.0225$. Larger threshold values imply infeasibility of the linear program.

6 Conclusion

Downside risk measures are of central interest in the financial industry, and thus have become a standard tool for risk management. The reason is twofold, as not only the investor wishes to control the risk of downside deviations in his portfolio's value, but also because regulating authorities require the financial institutions to provide guard against risk. Among these downside risk measures, the class of coherent distortion risk measures plays a special role, as of its economic relevance and simple structure. For the portfolio optimization problem, risk constraints given by such measures are suitable and vividly applied. For the CVaR risk measure, Rockafellar and Uryasev (2000) shows the equivalence of CVaR_α constrained portfolio optimization to a linear program.

This paper extends their results. It is shown that for any portfolio optimization problem under risk constraints given by a coherent distortion risk measure ρ, there exists an equivalent formulation as a linear program. The key reference is Bertsimas and Brown (2009), connecting optimization under a specific risk constraint to linear optimization with parameters given in an uncertainty set $\mathcal{U}_\rho$. As it turns out, these uncertainty sets have the same structure as the recently introduced weighted-mean trimmed regions, see Dyckerhoff and Mosler (2010b). Combining both results, the portfolio optimization problem with a risk constraint given by a coherent distortion risk measure can always be solved by means of linear optimization. Furthermore, the uncertainty set $\mathcal{U}_\rho$ generated by observed data and by the coherent distortion risk measure ρ receives a geometric interpretation as a collection of scenarios. As this uncertainty set is convex, only its extreme points, which can be interpreted as unique scenarios, must be examined for the optimization problem. All other points in the uncertainty set can be seen as mixtures of these unique scenarios, and as such they do not generate extreme portfolio returns. Once the uncertainty set is constructed, the set of all feasible portfolio weight vectors is established, and due to the work Bertsimas and Brown (2009) it can be represented as a scaled version of the polaroid of the uncertainty set. As the set of all (practically) feasible portfolio weight vectors is also convex and finitely generated, a linear program gives the optimal portfolio. The fact that the uncertainty set is constructed without any distributional assumption on the data-generating process should underline the usefulness of the introduced method.

Computational issues prohibit large-scale application of the given methodology at this point in time, especially because the computation of the uncertainty sets is time-consuming. Algorithms that compute the weighted-mean trimmed regions within a prespecified tolerance of precision, but therefore faster and with fewer extreme points, might be of enormous interest for the kind of application presented in this work. However, with the exact algorithm introduced in Bazovkin and Mosler (2010), small portfolios can be optimized for decent sizes of historical return data.

References

Acerbi, Carlo, "A coherent representation of subjective risk aversion," *Journal of Banking and Finance*, 2002, *26* (7), 1505–1518.

Alexander, Gordon J. and Alexandre M. Baptista, "A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model," *Management Science*, 2004, *50* (9), 1261–1273.

Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath, "Coherent measures of risk," *Mathematical Finance*, 1999, *9* (3), 203–228.

Bazovkin, Pavel and Karl Mosler, "An exact algorithm for weighted-mean trimmed regions in any dimension," 2010. Discussion Paper, University of Cologne, Department of Economic and Social Statistics.

Bertsimas, Dimitris and David B. Brown, "Constructing uncertainty sets for robust linear optimization," *Operations Research*, 2009, *57* (6), 1483–1495.

Burdet, Claude-Alain, "Polaroids: a new tool in non-convex and in integer programming," *Naval Research Logistics Quarterly*, 1973, *20* (1), 13–24.

Chopra, Vijay K. and William T. Ziemba, "The effect of errors in means, variances, and covariances on optimal portfolio choice," *The Journal of Portfolio Management*, 1993, *19* (2), 6–11.

Dyckerhoff, Rainer and Karl Mosler, "Weighted-mean regions of a probability distribution," 2010. Mimeo, University of Cologne, Department of Economic and Social Statistics.

_ and _ , "Weighted-mean trimming of multivariate data," *Journal of Multivariate Analysis*, 2010, *102* (3), 405–421.

European Central Bank, "Statistical Data Warehouse," `http://sdw.ecb.europa.eu/`, retrieved November 10 2010.

Ferguson, Kathleen W. and Brian M. Rom, "Post-modern portfolio theory comes of age," *The Journal of Investing*, 1994, *3* (3), 11–17.

Föllmer, Hans and Alexander Schied, *Stochastic finance*, de Gruyter, 2004.

Garlappi, Lorenzo, Raman Uppal, and Tan Wang, "Portfolio selection with parameter and model uncertainty: a multi-prior approach," *The Review of Financial Studies*, 2007, *20* (1), 41–81.

Halldórsson, Bjarni V. and Reha H. Tütüncü, "An interior-point method for a class of saddle-point problems," *Journal of Optimization Theory and Applications*, 2003, *116* (3), 559–590.

Huber, Peter J., *Robust Statistics*, John Wiley & Sons, New York, 1981.

Koshevoy, Gleb and Karl Mosler, "Zonoid trimming for multivariate distributions," *The Annals of Statistics*, 1997, *25* (5), 1998–2017.

Krokhmal, Pavlo, Jonas Palmquist, and Stanislav Uryasev, "Portfolio optimization with Conditional Value-at-Risk objective and constraints," *Journal of Risk*, 2002, *4* (2), 11–27.

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

Mosler, Karl, Tatjana Lange, and Pavel Bazovkin, "Computing zonoid trimmed regions of dimension $d > 2$," *Computational Statistics and Data Analysis*, 2009, *53* (7), 2500–2510.

Nawrocki, David, "A brief history of downside risk measures," *Journal of Investing*, 1999, *8* (3), 9–25.

Ogryczak, Wlodzimierz and Andrzej Ruszczynski, "From stochastic dominance to mean-risk models: semideviations as risk measures," *European Journal of Operational Research*, 1999, *116* (1), 33–50.

Pflug, Georg, "A value-of-information approach to measuring risk in multi-period economic activity," *Journal of Banking and Finance*, 2006, *30* (2), 695–715.

_ **and David Wozabal**, "Ambiguity in portfolio selection," *Quantitative Finance*, 2007, *7* (4), 435–442.

Rockafellar, R. Tyrell, *Convex Analysis*, John Wiley & Sons, New York, 1970.

_ **and Stanislav Uryasev**, "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2000, *2* (3), 21–41.

Savage, Leonard J., *The Foundations of Statistics*, Dover Publications, Inc., 1972.