CONDITIONAL VALUE AT RISK

CONDITIONAL VALUE AT RISK AS A CRITERION

FOR

OPTIMAL PORTFOLIO SELECTION

By

MARIA BETCHEVA

A Project

Submitted to the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree

Master of Science

McMaster University

MASTER OF SCIENCE (2005) McMaster University

(Mathematics) Hamilton, Ontario

TITLE: Conditional Value at Risk as a Criterion for Optimal Portfolio Selection

AUTHOR: Maria Betcheva

SUPERVISOR: Professor T. R. Hurd

NUMBER OF PAGES: vi, 55

ABSTRACT

The focus of my master's project research involves Conditional Value at Risk (or Expected Shortfall) as a risk measure for optimal portfolio selection. The project is organized as follows. In the first chapter, we introduce and discuss the quantile based risk measures, Value at Risk (VaR) and Conditional Value at Risk (CVaR), with respect to axiomatic characterization of coherent risk measures. As an alternative to VaR, CVaR has been attracting attention since it is a coherent measure. The properties and advantages of CVaR are analyzed. The second chapter deals with mean-risk models of portfolio optimization. The common idea in all asset allocation models is the minimization of some measure of risk while simultaneously maximizing portfolio expected return. Portfolio optimization in a mean-CVaR framework has been actively discussed recently. CVaR is a numerically tractable measure, allowing optimal portfolios to be computed by means of convex programming. Most importantly for applications, however, a mean-CVaR model can be used with scenario simulation of loss distributions. We investigate the convergence of the Monte-Carlo based CVaR optimal portfolio algorithm when an analytical solution of the optimization problem can be obtained. The last chapter considers the benchmark (or relative) portfolio selection problem in terms of a multiobjective problem. Tracking error optimization in a mean-multirisk framework allows implementation of an interactive decision making and taking into account of the investor's preferences.

ACKNOWLEDGEMENTS:

I cordially thank my supervisor Professor Dr. T. Hurd, who has shown a great and sincere interest in my project. Our numerous scientific discussions and his many valuable comments have considerably improved this work.

Special thanks are due to the faculty of the Department of Mathematics and Statistics at McMaster University for a truly wonderful learning experience. My deep gratitude goes to Dr. Shui Feng, Dr. T. Terlaky, Dr. M. Grasselli, Dr. M. Min-Oo, Dr. L. Bronsard, Dr. E. Sawyer, Dr. F. Hoppe, and Dr. B. Protas.

All my fellow students at PhiMAC are thanked for their friendship.

Contents

CHAPTER 1

Risk Measures

1.1. Risk Measures and Coherence

Risk in finance has no single definition. It is usually understood as the future chance or probability of losing or not gaining in value. Therefore, it is better to speak of measures that describe risk, or measures that give the manager or decision maker a "quantitative" tool to compare different alternatives. In risk management, risk measures are considered as mathematical objects, whereas risk is a concept from real life subject to different interpretations.

1.1.1. Risk Measures - Definition, Classification.

In order to introduce some examples and basic concepts from risk measurement, we first give a formal definition of risk measures. We assume that the financial consequences of economic activities can be assessed on the basis of a random variable X. This random variable may represent the absolute or relative return of an investment, the profit or the return on capital of a company, or the accumulated return for a portfolio of risky assets. In general, we think of X as a profit-and-loss (payoff or return) random variable, which may have positive, as well as negative values. We may also specify the loss of a financial position as the random variable $Y = -X$. Let now $(\Omega, \mathcal{A}, P)$ be some probability space, where the elements ω of Ω represent future states, or individual scenarios; $\mathcal{A}$ is the field of measurable subsets of Ω and P is a probability measure on $\mathcal{A}$. Denote by $\triangle t$ a fixed time interval and consider a set $\mathcal{V}$ of all real-valued $\mathcal{A}-$measurable random variables X (on the $(\Omega, \mathcal{A}, P)$), which we interpret as the possible profit-and-losses of some financial position over the time horizon $\triangle t$.

DEFINITION 1.1.1. *A* risk measure *is any mapping* $\rho: \mathcal{V} \rightarrow \mathbb{R}$.

There are great number of risk measures considered in the financial literature. We present here a classification, given in [Albrecht, 2004]. Risk measures are plased into two main categories: risk as the magnitude of deviations from a target; and risk as necessary capital respectively necessary premium. We now explain these two categories:

1. *Risk as the magnitude of the deviation from a target*

The target is an arbitrary deterministic value, for example the expected value. When the target is the mean of the random variable, risk measures are called *deviation measures*. The first category includes:

(a) Two-sided risk measures - penalize negative as well as positive deviations. Some examples are the classical measures, *variance* and *standard deviation*, as well as the *mean absolute deviation*, given by:

$$MAD(X) = \mathbb{E}\left[|x - \mathbb{E}(X)|\right]$$

or the more general risk measures

$$\rho(X) = \mathbb{E}\left[|x - \mathbb{E}(X)|^k\right],$$

$$\rho(X) = \mathbb{E}\left[|x - \mathbb{E}(X)|^k\right]^{1/k},$$

$$\rho(X) = \mathbb{E}\left[f\left(x - \mathbb{E}(X)\right)^k\right]^{1/k}.$$

The last type, with f a monotonic function, allows a different weighting of positive and negative deviations from the expected value, and is considered by Rockafellar et al. [2002].

(b) One-sided (downside) measures - penalize only the negative deviation from a target.

A general class of risk measures of this type is the class of lower partial moments of degree k ($k = 0, 1, 2, ...$) :

$$LPM_k(c,\, X) = \mathbb{E}\left[\max(c - X,\, 0)^k\right],$$

where c denotes the reference level from which the deviation is measured.

The expected regret, considered by Dembo [1991], is the case $k = 1$:

$$ER = \mathbb{E}[\max(c - X,\, 0)].$$

The lower-semi-absolute deviation is the case $c = \mathbb{E}(X)$, $k = 1$:

$$LSAD = \mathbb{E}[\max(\mathbb{E}(X) - X,\, 0)],$$

and also considered by Ogryczak and Ruszczynski [1999, 2001] and Gotoh and Konno [2000] with relation to the theory of stochastic dominance and mean-risk analysis.

The semivariance, proposed by Markowitz [1987], is the case $c = \mathbb{E}(X)$, $k = 2$:

$$\rho(X) = \mathbb{E}\left[\max\left(\mathbb{E}(X) - X,\, 0\right)^2\right].$$

2. *Risk as necessary capital (or necessary premium)*

Here, the risk is regarded as either:

(a) necessary capital (in terms of capital to be added to a financial position to make it riskless) in case the value of the risk measure is positive; or

(b) necessary premium (to be withdrawn from a position without endangering safety) in case the risk measure is negative

Note that these risk measures can have both positive and negative values, whereas the measures in the first category have only positive ones. Examples are Value at Risk (VaR) and Conditional Value at Risk. For an overview of the risk measures and their properties, the reader can see also [Cheng et al., 2004].

In the financial industry, VaR is a widely used measure for quantifying the future losses. For example, the Basel Committee on Banking Supervision requires banks to use VaR to determine the minimum capital to support their portfolios. Moreover, Hull [2003] notes that VaR is also used by dealers, fund managers, and financial institutions. With the emerging of VaR as one of the most popular concepts in Risk Management, researchers have been extensively criticized the use of VaR as a "good" measure of risk. In fact, before the appearance of the seminal works of Artzner et al. [1999], Delbaen [2002], the notion of a risk measure was not clearly defined, and there was no list of the properties a good (from a financial point of view) measure should have. Artzner et al. [1999] stated the foundations of the axiomatic approach to the characterization of risk measures. They proposed a list of structural properties, called *axioms of coherence*, a reasonable measure should satisfy.

1.1.2. Axiomatic Characterization of Risk Measures.

We now introduce the axioms of coherence, and give their financial interpretation.

DEFINITION 1.1.2. *A mapping $\rho: \mathcal{V} \to \mathbb{R}$ is called a* coherent risk measure *if it satisfies the following axioms*

(a) Translation invariance: if $m \in \mathbb{R}$, then $\rho(X + m) = \rho(X) - m$, $\forall X \in \mathcal{V}$;

(b) Monotonicity: if $X \leq Y$, then $\rho(X) \geq \rho(Y)$, $\forall X, Y \in \mathcal{V}$;

(c) Positive homogeneity: if $m \geq 0$, then $\rho(mX) = m\,\rho(X)$, $\forall X \in \mathcal{V}$;

(d) Subadditivity: $\rho(X + Y) \leq \rho(X) + \rho(Y)$, $\forall X, Y \in \mathcal{V}$;

Coherent risk measures were extended to *convex risk measures* [Follmer and Schied, 2002], also called *weakly coherent measures*, by relaxing the constraint of subadditivity and positive homogeneity, and instead requiring *convexity*:

(e) Convexity: $\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y)$, $X, Y \in \mathcal{V}$, $\lambda \in [0, 1]$;

DEFINITION 1.1.3. *A mapping* $\rho : \mathcal{V} \rightarrow \mathbb{R}$ *is called a* convex risk measure *if it satisfies the axioms of translation invariance (a), monotonicity (b), and convexity (e).*

Axiom (a) states that adding a deterministic quantity m to the initial position decreases risk by that amount. In particular, we have $\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$. Thus, once we have added $\rho(X)$ to a position, the resulting position is *acceptable* (i.e. has nonpositive risk) without any further changes.

Monotonicity (b) means that if $X(\omega) \leq Y(\omega)$ for every state of nature then X is riskier than Y because of the higher loss potential.

Subadditivity (d) requires the risk of a combined position to be less than the sum of the risks of the separate positions. In fact, this axiom reflects the idea that risk can be reduced by diversification, one of the principles in finance. At the same time, subadditivity is the most debated axiom because it rules out certain risk measures popular in practice, such as semivariance and Value at Risk.

Axiom (c) implies that the risk of a certain multiple of a financial position is identical with corresponding multiple of the risk of the position. This axiom is natural in conjuction with the previous axiom. Subadditivity implies that for $m \in \mathbb{N}$,

$$\rho(mX) = \rho(X + ... + X) \leq m\rho(X).$$

Since there is no diversification in this portfolio, it is natural to require that equality holds above, which leads to positive homogeneity.

The axiom of convexity (e) takes into account the situations where the risk of a position increases in a nonlinear way with the size of the position.

The axioms of coherence have been very influential. We already mentioned that VaR is not a coherent measure since it fails to satisfy the subadditivity property, that is, the VaR of a portfolio of two securities may be larger than the sum of the VaR of each of the securities in the portfolio. Starting from coherent risk measures, several quantile-based alternatives to VaR were proposed in the literature: firstly, "tail conditional expectation" (TCE) (see [Artzner et al., 1999]; and later, "expected shortfall" (ES) (see Aserbi and Tashe [2002b]) and "conditional value at risk" (CVaR) (see [Rockafellar and Uryasev, 2002]). Research in this direction has received much attention in the last few years. In the next section, we introduce different mathematical definitions of VaR and its quantile-based alternatives, given in the literature, and discuss the relations between the measures. The main questions that we address in the next section are the following: 1) What mathematical definitions of VaR and its quantile-based alternatives exist in the literature? 2) Is there a relation between these measures? and 3) What properties of CVaR make it a risk measure, dominating and therefore preferable to VaR as a risk management tool?

1.2. Quantile Type Risk Measures

1.2.1. Quantiles.

DEFINITION. Let X be a real random variable defined on a probability space $(\Omega, \mathcal{A}, P)$ and $\alpha \in [0, 1]$. Then the α quantiles of X are the elements of the set:

$$Q_\alpha(X) = \{x \in \mathbb{R},\ P(X < x) \le \alpha \le P(X \le x)\}.$$

DEFINITION 1.2.1. **Upper and Lower Quantiles**
Let $(\Omega, \mathcal{A}, P)$ *be a probability space and* $\alpha \in [0, 1]$. *For* X, *a random variable defined on* $(\Omega, \mathcal{A}, P)$, *the* upper quantile of order α *of* X is

$$q_\alpha^+(X) = \sup\{x \in \mathbb{R},\ P(X < x) \le \alpha\};$$

the lower quantile of order α *of* X is

$$q_\alpha^-(X) = \inf\{x \in \mathbb{R},\ P(X \le x) \ge \alpha\}.$$

Furthermore $q_\alpha^-(X)$ and $q_\alpha^+(X)$ are the left and resp. the right end of the $\alpha-$quantile set:

$$Q_\alpha(X) = [q_\alpha^-(X),\, q_\alpha^+(X)].$$

The next proposition provides an alternative characterization of quantiles.

PROPOSITION 1.2.2. (in [Laurent, 2003])

Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\alpha \in [0,1]$ and X a random variable. Then,

$$q_\alpha^+(X) = \inf\{x \in \mathbb{R},\, P(X \leq x) > \alpha\}$$

and

$$q_\alpha^-(X) = \sup\{x \in \mathbb{R},\, P(X < x) < \alpha\}.$$

The next proposition [Laurent, 2003] gives the relation between the quantiles of random variables X and $-X$.

PROPOSITION 1.2.3. *Let $(\Omega, \mathcal{A}, P)$ be a probability space, $\alpha \in [0,1]$ and X a random variable. Then, the α quantiles of $-X$ are the opposites of the $(1-\alpha)$ quantiles of X. Thus,*

$$Q_\alpha(-X) = -Q_{1-\alpha}(X)$$

and therefore

$$q_\alpha^-(-X) = -q_{1-\alpha}^+(X), \quad q_\alpha^+(-X) = -q_{1-\alpha}^-(X). \tag{1.2.1}$$

For discrete distributions, we have:

PROPOSITION 1.2.4. *Let X be a discrete random variable taking values among $x_i \in \mathbb{R},\ i = 1, \ldots, n$. Then*

$$x_i = q_\alpha^+(X) \iff P(X < x_i) \leq \alpha < P(X \leq x_i)$$

and

$$x_i = q_\alpha^-(X) \iff P(X < x_i) < \alpha \leq P(X \leq x_i)$$

Next, we introduce Value at Risk, Conditional Value at Risk, and Expected Shortfall, and give the conditions under which Conditional Value at Risk and Expected Shortfall represent the same risk measure.

1.2.2. Value at Risk.

Let us first introduce the cumulative distribution function (cdf) $F(x) = P(X \leq x)$ of the random variable X. According to def.1.2.1 and prop.1.2.2, in terms of $F(x)$ we get

$$q_\alpha^+(X) = \inf\{x \in \mathbb{R},\ F(x) > \alpha\} \tag{1.2.2}$$

and

$$q_\alpha^-(X) = \inf\{x \in \mathbb{R},\ F(x) \geq \alpha\}. \tag{1.2.3}$$

Let the above random variable X represents some portfolio's *profit and loss* $(P\&L)$ random variable. Now define the first quantile-based risk measure: Value at Risk (VaR). We join here Aserbi and Tashe [2002a], Delbaen [2002], Artzner et al. [1999] by taking VaR_α as the upper α quantile of the $(P\&L)$ distribution. Thus, we have

DEFINITION 1.2.5. **Value at Risk.**

Let X be a profit and loss $(P\&L)$ random variable, defined on a *probability space* $(\Omega, \mathcal{A}, P)$. *For a fixed level $\alpha \in (0,1)$, the* Value at Risk of X for the level α, $VaR_\alpha(X)$ is:

$$VaR_\alpha(X) = -\inf\{x \in \mathbb{R},\ F(x) > \alpha\} = -q_\alpha^+(X) \tag{1.2.4}$$

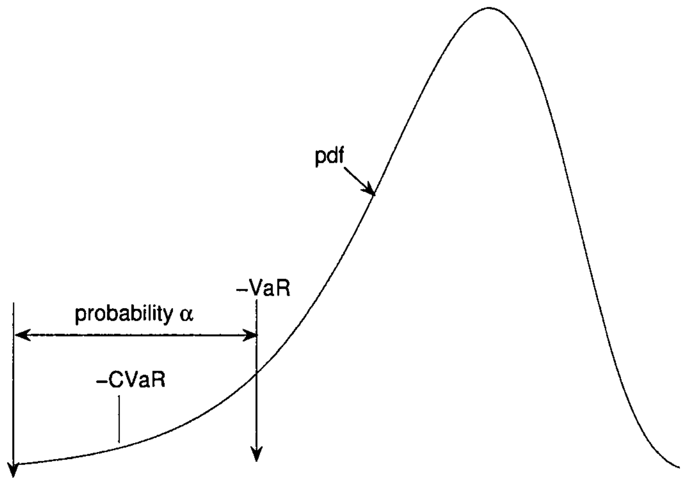

Figure 1: *Value at Risk of X at level* α

Note that for the random variable $-X$, representing portfolio *loss*, the following relation holds:

$$VaR_{1-\alpha}(-X) = -VaR_{\alpha}(X)$$

REMARK 1.2.6. Value at Risk, defined as a limit from the left is given in [Aserbi and Tashe, 2002b]:

$$VaR_{\alpha}(X) = -\sup\{x \in \mathbb{R},\, P(X < x) \leq \alpha\}$$

This formulation of VaR, however, is not used very often.

To explain the VaR concept, we consider figure 1. Let q_α is the $\alpha-$quantile of a continuous $(P\&L)$ distribution, i.e.

$$P(X \leq q_\alpha) = \alpha, \quad P(X > q_\alpha) = 1 - \alpha$$

Given the threshold z_α, the $VaR_\alpha(X)$ is by definition the amount of money, which is needed to cover *losses* up to $-q_\alpha$,i.e $VaR_\alpha(X) = -q_\alpha(X)$. Therefore, $VaR_\alpha(X)$ is implicitly defined by

$$P(X \leq -VaR_\alpha(X)) = \alpha \; if \; P(X < 0) > \alpha \quad and$$

$$VaR_\alpha(X) = 0 \; if \; P(X < 0) \leq \alpha$$

Usually, when working with $(P\&L)$ distribution, we take α between 0.01 and 0.05 (1%-5%).

1.2.3. CVaR and Expected Shortfall.

The term Conditional Value at Risk was introduced in [Rockafellar and Uryasev, 2000, Uryasev, 2000] for the case of continuous loss distributions. The concept of CVaR for general probability distributions, including discrete distributions, was later developed by Rockafellar and Uryasev [2002]. They define Conditional Value at Risk of X at the level α, denoted by $CVaR_\alpha(X)$, as a weighted average of $VaR_\alpha(X)$ and $CVaR_\alpha^+(X)$, called "upper CVaR".

DEFINITION 1.2.7. **Conditional Value at Risk**

Let X be a $P\&L$ random variable, defined on a probability space $(\Omega, \mathcal{A}, P)$ with finite expectation and $\alpha \in (0,1)$. Then CVaR of X at the level α *is given by:*

$$CVaR_\alpha(X) = \lambda \, VaR_\alpha(X) + (1-\lambda) \, CVaR_\alpha^+(X), \tag{1.2.5}$$

where

$$CVaR_\alpha^+(X) = -\mathbb{E}[X \mid X \leq q_\alpha^+(X)] = -\mathbb{E}[X \mid X \leq -VaR_\alpha(X)]$$

and

$$\lambda = \frac{\alpha - P(X \leq q_\alpha^+(X))}{\alpha} = \frac{\alpha - P(X \leq -VaR_\alpha(X))}{\alpha} \leq 1$$

Independently, Aserbi and Tashe [2002b,a] introduced the term Expected Shortfall (ES).

DEFINITION 1.2.8. **Expected Shortfall**

Let X be a $P\&L$ random variable, defined on a *probability space* $(\Omega, \mathcal{A}, P)$ *with finite expectation and* $\alpha \in (0,1)$. *Then the* Expected shortfall of X at the level α *is defined by:*

$$ES_\alpha(X) = -\alpha^{-1}\{\mathbb{E}[X\mathbf{1}_{\{X\leq q_\alpha^+(X)\}}] - q_\alpha^+(X)(P[X \leq q_\alpha^+(X)] - \alpha)\} \tag{1.2.6}$$

where $\mathbf{1}_E$ *denotes the indicator function of a set* E.

In the same paper, the authors adopt the term tail conditional expectation, $(TCE_\alpha(X))$, for $CVaR_\alpha^+(X)$:

$$TCE_\alpha(X) = -\mathbb{E}[X/X \leq q_\alpha^+(X)]$$

PROPOSITION 1.2.9. *Let X be a real random variable defined on a probability space* $(\Omega, \mathcal{A}, P)$, $s \in \mathbb{R}$ *and* $\alpha \in (0,1)$ *be fixed. Then*

$$ES_\alpha(X) = CVaR_\alpha(X)$$

PROOF. The expressions given by right hand sides of (1.2.5) and (1.2.6) are equivalent. For example, (1.2.5) can be derived from (1.2.6) multiplying and dividing by $P(X \leq q_\alpha^+)$ and expressing $\mathbb{E}[X\mathbf{1}_{\{X\leq q_\alpha^+\}}]$ as a

$$\mathbb{E}[X\mathbf{1}_{\{X\leq q_\alpha^+\}}] = \mathbb{E}[X/X \leq q_\alpha^+(X)]\, P(X \leq q_\alpha^+).$$

□

NOTE. The term $q_\alpha^+(P(X \leq q_\alpha^+) - \alpha)$ in (1.2.6) is interpreted as the exceeding part to be subtracted from the expected value $\mathbb{E}[X\mathbf{1}_{\{X\leq q_\alpha^+\}}]$ when $P(X \leq q_\alpha^+) \geq \alpha$. When the *distribution is continuous*, $P(X \leq q_\alpha^+) = \alpha$ and therefore $ES_\alpha(X) = TCE_\alpha(X)$. By analogy, when $P(X \leq q_\alpha^+) = \alpha$, the coefficient λ equals zero and $CVaR_\alpha(X) =$

$CVaR^+_\alpha(X)$ (see 1.2.5). In general, $CVaR_\alpha(X) \geq CVaR^+_\alpha(X)$ (equivalently $ES_\alpha(X) \geq TCE_\alpha(X)$). This can be easily seen if we rewrite (1.2.5) as

$$CVaR_\alpha(X) = CVaR^+_\alpha(X) + (\mu - 1)\left(CVaR^+_\alpha(X) - VaR_\alpha(X)\right)$$

where

$$\mu = \frac{P(X \leq q^+_\alpha)}{\alpha} \geq 1.$$

Note also that $CVaR_\alpha(X)$ is conditional tail estimation, while $VaR_\alpha(X)$ is a quantile function. Figure 1 displays $CVaR_\alpha(X)$ as a average value of the losses, given the losses are larger than $VaR_\alpha(X)$.

Keeping in mind that the terms "CVaR" and "ES" are interchangeable, further in our discussions we will use only the notion Conditional Value at Risk.

There are two alternative representations of the CVaR, which are given by the following propositions:

PROPOSITION 1.2.10. ***CVaR by optimization***

Let X be a P&L random variable defined on a probability space $(\Omega, \mathcal{A}, P)$ with expectation $E[X^-] < \infty$, where $X^- = \max(0, -X)$. Let $\alpha \in (0,1)$. $CVaR_\alpha(X)$ is the solution of the minimization problem:

$$CVaR_\alpha(X) = \min_{s \in \mathbb{R}} \frac{1}{\alpha}\,\mathbb{E}[(X - s)^-] - s. \tag{1.2.7}$$

The above minimization formula was first developed for the case of continuous loss distribution function in [Rockafellar and Uryasev, 2000]. Later Pflug [2000] proved the validity of the formula for discontinuous c.d.f. as well.

PROPOSITION 1.2.11. ***Spectral representation of CVaR***

Let X be a real random variable defined on a probability space $(\Omega, \mathcal{A}, P)$, *and* $\alpha \in (0, 1)$ *be fixed. Then*

$$CVaR_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha q_u^+(X)\,du. \tag{1.2.8}$$

Formula 1.2.8 shows that $CVaR_\alpha(X)$ *can be expressed as a* VaR_α *average. Because of that, Follmer and Schied [2002] proposed another name: "average value at risk". The last formulation of* $CVaR_\alpha$*(1.2.8) is related to the so-called spectral risk measures, considered in [Aserbi, 2004].*

The next proposition shows the relation between $VaR_\alpha(X)$ and $CVaR_\alpha(X)$, expressed by eq. (1.2.7).

PROPOSITION 1.2.12. (see [Aserbi and Tashe, 2002a, Rockafellar and Uryasev, 2002, Chabaane et al.])

Let X be a real random variable defined on a probability space $(\Omega, \mathcal{A}, P)$ *with finite expectation. Let* $\alpha \in [0, 1]$ *and* $s \in \mathbb{R}$. *Then the function*

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

is minimal on the quantile set $Q_\alpha(X)$, *i.e.*

$$\arg\min_s F_\alpha(s) = [q_\alpha^-(X),\, q_\alpha^+(X)]$$

In particular, one always has $|VaR_\alpha(X)| \in \arg\min_s F_\alpha(s)$.

PROOF. Let $Z_\alpha(s) = \alpha\,(X-s)^+ + (1-\alpha)\,(X-s)^-$, where $(X-s)^+ = \max(X-s,\, 0)$ and $(X-s)^- = \max(s-X,\, 0)$. The set of *minimizers* of

$$H_\alpha = \mathbb{E}(Z_\alpha(s)) = \alpha\,\mathbb{E}[(X-s)^+] + (1-\alpha)\,\mathbb{E}[(X-s)^-]$$

is the $\alpha-$quantile set (see the proof in Chabaane et al.), i.e. $\arg\min_s H_\alpha(s) = Q_\alpha(X)$. Then using the equality $x^+ = x + x^-$, we can write

$$Z_\alpha(s) = \alpha\,(X-s) + \alpha\,(X-s)^- + (1-\alpha)\,(X-s)^- = \alpha\,[X + \frac{1}{\alpha}\,(X-s)^- - s].$$

Thus note the following equivalent representation for $H_\alpha(s)$:

$$H_\alpha(s) = \alpha\,\mathbb{E}[X] + \alpha\,\left(\frac{1}{\alpha}\,\mathbb{E}[(X-s)^-] - s\right).$$

Then the function $F_\alpha(s)$ can be expressed as:

$$F_\alpha(s) = \frac{1}{\alpha}\,\mathbb{E}[(X-s)^-] - s = \frac{1}{\alpha}\,H_\alpha(s) - \mathbb{E}[X].$$

As a consequence, its minimum is attained for $s \in Q_\alpha(X) = [q_\alpha^-(X),\, q_\alpha^+(X)]$. □

1.2.4. Properties of CVaR.

Before we state some properties of $CVaR_\alpha(X)$, we define properties of risk measures in terms of preference sructures induced by dominance relations.

DEFINITION 1.2.13. [Pflug, 2000]

(1) A relation between two random variables X_1,X_2, denoting the profit-and-loss of two portfolios, is of stochastic dominance of order 1, $X_1 \prec_{SD(1)} X_2$, *iff*

$$\mathbb{E}[f(X_1)] \leq \mathbb{E}[f(X_2)]$$

for all monotonic nondecreasing functions f. We say that X_2 dominates X_1(or X_2 is prefered to X_1) iff the above inequality holds.

(2) A relation between two random variables X_1,X_2 is of stochastic dominance of order 2, $X_1 \prec_{SD(2)} X_2$, *iff*

$$\mathbb{E}[f(X_1)] \leq \mathbb{E}[f(X_2)]$$

for all concave, monotonic nondecreasing functions f.

PROPOSITION 1.2.14. *(1) $CVaR_\alpha$ is translation invariant, i.e.*

$$CVaR_\alpha(X+m) = CVaR_\alpha(X) - m.$$

for m any real number.

(2) $CVaR_\alpha$ is positively homogeneous, i.e.

$$CVaR_\alpha(mX) = m\, CVaR_\alpha(X)$$

for $m > 0$.

(3) $CVaR_\alpha$ is convex, i.e. for arbitrary random variables X_1, X_2 and $\lambda \in [0,\, 1]$,

$$CVaR_\alpha(\lambda X_1 + (1-\lambda)X_2) \leq \lambda CVaR_\alpha(X_1) + (1-\lambda)CVaR_\alpha(X_2).$$

(4) $CVaR_\alpha$ is monotonic w.r.t. stochastic dominance of order 2 (and therefore w.r.t. stochastic dominance of order 1), i.e. if

$$X_1 \prec_{SD(2)} X_2 \quad \textit{then} \quad CVaR_\alpha(X_1) \geq CVaR_\alpha(X_2).$$

PROOF. Properties (1) and (2) are immediate from the characterization of $CVaR_\alpha$. To prove the convexity, we fix numbers s_i such that

$$CVaR_\alpha(X_i) = \frac{1}{\alpha}\mathbb{E}(X_i - s_i)^- - s_i.$$

Since the function $x \longrightarrow (x-s)^-$ is convex, we have

$$\begin{aligned} CVaR_\alpha(\lambda X_1 + (1-\lambda)X_2) &= \frac{1}{\alpha}\,\mathbb{E}[\lambda X_1 + (1-\lambda)X_2 - \lambda s_1 - (1-\lambda)s_2]^- - \lambda s_1 - (1-\lambda)s_2 \\ &\leq \frac{1}{\alpha}\,\mathbb{E}(X_1 - s_1)^- + \frac{1-\lambda}{\alpha}\,\mathbb{E}(X_2 - s_2)^- - \lambda s_1 - (1-\lambda)s_2 \\ &= \lambda CVaR_\alpha(X_1) + (1-\lambda)CVaR_\alpha(X_2) \end{aligned}$$

Note that monotonicity w.r.t. stochastic dominance of order 1 is the axiom (b) of coherence. The proof of the ordering properties follows from the fact that the function $x \longrightarrow (x-s)^-$ is convex and monotone. $\square$

In contrast to $CVaR_\alpha$, VaR_α is only translation equivariant, positively homogeneous, and monotonic w.r.t. stochastic dominance of order 1, but is not convex. Moreover, stochastic dominance of order 1 does not account for the decision maker's risk aversion. Let us consider the function f in Definition 1.2.13 as a utility function. The condition that f is nondecreasing and concave means that f represents a nonsatiated ($f' > 0$) and risk-averse ($f'' < 0$) preference. Then, if X_2 dominates X_1 in the sense of second order stochastic dominance every risk-averse investor chooses X_2 over X_1. The connection between utility theory and stochastic dominance is given in [Levy, 1998].

CHAPTER 2

Optimization of CVaR

2.1. Mean-Risk Portfolio Selection

2.1.1. Preliminaries.

Consider a given portfolio ψ from a universe of n risky instruments, for example, a book of n derivatives, or collection of n stocks or bonds. We denote the *value* of this portfolio at a present time $t_0 = 0$ by $V^{\psi}(0)$. For a given time horizon $\Delta t = [0,\, t]$, such as one or ten days, the portfolio *profit and loss* X over the period is given by the change in a portfolio market value:

$$X^{\psi} = \Delta V^{\psi} = V^{\psi}(t) - V^{\psi}(0).$$

While X is observable at time t, it is random from the viewpoint of time 0. The distribution of the random variable X is the so-called profit & loss (P&L) distribution.

Furthermore the change in portfolio value is expressed as:

$$\Delta V^{\psi} = \sum_{i=1}^{n} \psi_i(0)\,[P_i(t) - P_i(0)]$$

where $\psi_i(0)$ is number of units of asset $i$held at time 0, and $P_i(.)$ is the price of asset i at a fixed time.

A vector $\psi = (\psi_1(0),\, \psi_2(0), ..., \psi_n(0))$ is called a trading strategy, since it characterizes the investor's decision.

Calculating the distribution of future prices (prices at time t) is a modeling issue. Following risk management practice, the asset's price usually is modelled as a function of stochastic market factors, which are observable at time 0. Examples of the most frequently used risk factors are the logarithmic prices of stocks, yields, exchange and interest rates.

Instead of working with prices, we often use *return*. The rate of return $r_i(t)$ of a security i at time t is defined by $r_i = P_i(t)/P_i(0) - 1$. Now, instead of the trading strategies defined for absolute security prices, we use the *normalized* strategies $\phi = (\phi_1(0), \phi_2(0), ..., \phi_n(0))^T$, where

$$\phi_i(0) = \frac{\psi_i(0)\, P_i(0)}{V(0)}$$

is the fraction of $V(0)$ in asset i at time 0 and also $\sum_{i=1}^{n} \phi_i = 1$ holds.

Thus the portfolio *return* R^ϕ is considered as a function of the vector of assets' weights ϕ and the vector of assets' relative returns $r = (r_1, r_2, ..., r_n)^T$, the latter one expressing the uncertanty in the price change:

$$R^\phi = \frac{V^\phi(t) - V^\phi(0)}{V^\phi(0)} = \sum_{i=1}^{n} \phi_i(0)\, r_i(t)$$

Unless stated otherwise, we assume that the set of all admissible portfolios is $\Phi = \{\phi \in \mathbb{R}^n \mid \sum_{i=1}^{n} \phi_i = 1,\ \phi_i \geq 0 \,\text{for all}\, i\}$, (i.e. no short-sales are allowed). Additional constraints $\mathcal{B}$ can be introduced, as for example diversification constraints $\mathcal{B} = \{\phi \in \mathbb{R}^n \mid \phi_i \in (\delta_i^-, \delta_i^+) \text{ for } i = 1, ..., n\}$, which give the limits for the weights.

Note that we will consider mean-risk models of portfolio optimization based on using a normalized strategy. Now we introduce some basic formulas.

Let $\mu_i = \mathbb{E}(r_i)$ denotes the expected return of asset i for $i = 1, ..., n$, and $\mu = (\mu_1, \mu_2, ..., \mu_n)^T$. For two risky assets r_j and r_k ($j \neq k$), $\sigma_{jk} = Cov(r_j, r_k)$ is the covariance between them. Then $V = (\sigma_{jk})_{1 \leq j, k \leq n}$ is the variance-covariance matrix, since $\sigma_{jk} = Var(r_j) = \sigma_j^2$ when $j = k$. For a portfolio ϕ, we have the portfolio expected rate of return and the variance, calculated by the formulas:

(2.1.1) $$\mathbb{E}(R^\phi) = \mu_p^\phi = \phi^T \mu$$

and

$$Var(R^{\phi}) = (\sigma_p^{\phi})^2 = \phi^T V \phi \tag{2.1.2}$$

2.1.2. Mean-Risk Models.

The common idea in all asset allocation mean-risk models is the minimization of some measure of risk while simultaneously maximizing portfolio expected return. For example, the variance of portfolio returns is used as a risk measure in the classical Markowitz approach. Variance has been very attractive as a risk measure in general and in the portfolio optimization problem in particular because it is a basic probability concept, well understood and easy to calculate (eq.2.1.2). Further, it allows for development of closed form solutions for finding optimal allocations, or for optimization using standard quadratic optimization techniques. However, variance as a risk measure has some obvious drawbacks. For example, variance is a measure of the first type in our classification, i.e. it penalizes negative as well as positive returns. Also, one of the fundamental assumptions of the traditional mean-variance problem is multi-normality of the return distributions - which does not take into account fat tails and skewness. With the progress in risk measurement, a number of alternative risk measures have been suggested for solving the asset allocation problem. Portfolio optimization in a mean-CVaR framework has been actively studied during the last years. This problem, and all mean-risk models in general, denoted by $(\mu(R), \rho(R))$, can be formulated as problems of the bicriterion mathematical programming problem (see Appendix):

$$\left| \begin{array}{ll} \max\limits_{\phi} \mathbb{E}(R) & \qquad (MR*) \\ \min\limits_{\phi} \rho(R) & \\ \text{s.t.}\ \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

We state the following definition:

DEFINITION 2.1.1. *A portfolio* ϕ_1 *is* mean-risk, (μ_p, ρ), efficient, *and therefore a solution of the above problem, if and only if no portfolio* ϕ_2 *exists such that* $\mathbb{E}(R^{\phi_1}) \leq \mathbb{E}(R^{\phi_2})$ *and* $\rho(R^{\phi_1}) \geq \rho(R^{\phi_2})$, *where at least one of the inequalities is strict. The* (μ_p, ρ)-efficient frontier *is the subset of* $\mathbb{R}^2$, *defined by all pairs* $(\mathbb{E}(R^{\phi}), \rho(R^{\phi}))$, *where* ϕ *is an efficient portfolio.*

PROPOSITION 2.1.2. *Let us consider the following three optimization problems:*

$$\min_{\phi} \rho(R) - \lambda \mathbb{E}(R), \qquad \phi \in \Phi, \ \lambda \geq 0 \qquad (P1)$$

$$\min_{\phi} \rho(R), \qquad \mathbb{E}(R) \geq \mu_p^*, \qquad \phi \in \Phi, \qquad (P2)$$

$$\min_{\phi} -\mathbb{E}(R), \qquad \rho(R) \leq \rho_p^*, \qquad \phi \in \Phi, \qquad (P3)$$

Varying the parameters λ, μ_p^*, *and* ρ_p^* *traces the efficient frontiers for the problems (P1), (P2) and (P3) accordingly. If* $\rho(R)$ *is convex, and the set* Φ *is convex (note that* $\mathbb{E}(R)$ *is linear), then the three problems (P1)-(P3) generate the same efficient frontier.*

PROOF. Proof is given in [Krokhmal et al., 2002] □

2.1.3. CVaR Minimization.

The minimization problem of $CVaR$ has been investigated by Rockafellar and Uryasev [2000, 2002], Uryasev and Rockafellar [2001], Pflug [2000]. The fundamental result of these papers is that the minimization of $CVaR_\alpha(R(\phi))$ is shown to be equivalent to the problem of the minimization of the following function containing an auxiliary variable s:

$$F_\alpha(R(\phi), s) = \frac{1}{\alpha} \mathbb{E}[R(\phi) - s]^- - s \qquad (2.1.3)$$

Note, that we already discussed the equivalence of the representations of $CVaR_\alpha$, given by formulas (1.2.5), (1.2.6), (1.2.7) and (1.2.8). In fact, $CVaR_\alpha$ was given as the minimum of the function $F_\alpha(R,\, s)$ in s (see eq.(1.2.7) and Prop.1.2.12). Moreover, Prop. 1.2.12 stated that the minimum of $F_\alpha(R,\, s)$ is attained when the auxiliary variable s takes any value in the quantile set, i.e.

$$\min_s F_\alpha(R,\, s) = F_\alpha(R,\, s^*) = CVaR_\alpha(R) \quad \forall s^* \in [q_\alpha^-(R),\, q_\alpha^+(R)]$$

The following proposition summarizes these results:

PROPOSITION 2.1.3. *1. $F_\alpha(R(\phi),\, s)$ is convex in s, as well as in ϕ, if $R(\phi)$ is convex in ϕ.*

2. if $R(\phi)$ is linear in ϕ, i.e. if the parameters ϕ_i are the assets weights (which is the case here) then the function $F_\alpha(R(\phi),\, s)$ is convex in the extended set of parameters $\{\phi,\, s\}$.

3.The minimum of $CVaR_\alpha(R(\phi))$ in ϕ coincides with the minimum of $F_\alpha(R(\phi),\, s)$ in the extended set of parameters $\{\phi,\, s\}$:

$$\min_\phi CVaR_\alpha(R(\phi)) = \min_{\phi,\, s} F_\alpha(R(\phi),\, s).$$

Furthermore, if the constraints are such that Φ is a convex set, the joint minimization is an instance of convex programming.

When the multivariate distribution of the return variables of the portfolio assets is given in empirical form by a sample of N scenarios with equal probabilities, (2.1.3) can be replaced by:

$$F_\alpha^{(N)}(R(\phi),\, s) = \frac{1}{N\alpha}\sum_{l=1}^{N}[R_l(\phi) - s]^- - s \tag{2.1.4}$$

where $R_l(\phi) = \phi_1 r_{1,l} + \phi_2 r_{2,l} + \ldots + \phi_n r_{n,l}, \quad l = 1, 2, \ldots, N.$

Since $R_l(\phi)$ is linear in ϕ, the function $F_\alpha^{(N)}$ is convex and piecewise linear.

Let's now discuss the optimization problem $(P2)$ with the Conditional value at risk as a risk measure $\rho(R)$. The *sample objective function* $F_\alpha^{(N)}(R(\phi),\, s)$ can be minimized by the means of standard nonlinear optimization approaches. The nonlinear optimization problem is therefore the following:

$$\left|\begin{array}{ll} \min\limits_{\phi,s} \frac{1}{N\alpha} \sum_{l=1}^{N} [R_l(\phi) - s]^- - s & (NP) \\ \text{s.t. } \mathbb{E}(R(\phi)) \geq \mu_p^* & \\ \phi \in \Phi \subset \mathbb{R}^n & \\ s \in \mathbb{R} & \end{array}\right.$$

Further by using auxiliary variables $z_i = 1, ..., N$, the nondifferentiable sample function $F_\alpha^{(N)}(R(\phi),\, s)$ can be replaced by a linear function and a set of linear constraints:

$$\left|\begin{array}{ll} \min\limits_{\phi,s,z} \frac{1}{N\alpha} \sum_{l=1}^{N} z_l - s & (LP) \\ \text{s.t. } z_l \geq s - R_l(\phi),\ l = 1, ..., N & \\ z_l \geq 0\ l = 1, ..., N & \\ \mathbb{E}(R(\phi)) \geq \mu_p^* & \\ \phi \in \Phi \subset \mathbb{R}^n & \\ s \in \mathbb{R} & \end{array}\right.$$

The two optimization problems (NP) and (LP) are equivalent. The latter one is linear programming problem since $R_l(\phi)$ is linear in ϕ. The number of variables of the linear problem is $n + N + 1$ while in the nonlinear case is $n + 1$.

An alternative of the (LP) problem is proposed in [Alexander et al., 2003]. The approach is based on approximation of the piecewise linear function $z^+ = \max(z,\, 0)$ by a continuously differentiable piecewise quadratic function $\eta_\epsilon(z)$; given a resolution parameter $\epsilon > 0$

$$\eta_\epsilon(z) = \begin{cases} z & if\ z \geq \epsilon \\ \frac{z^2}{4\epsilon} + \frac{1}{2}z + \frac{1}{4}\epsilon & if\ -\epsilon \leq z \leq \epsilon \\ 0 & otherwise \end{cases}$$

Thus, the sample objective function $F_\alpha^{(N)}(R(\phi),\ s)$ is replaced by the *approximate sample objective function*

(2.1.5)
$$F_\alpha^{a(N)}(R(\phi),\ s) = \frac{1}{N\alpha}\sum_{l=1}^{N}\eta_\epsilon(s - R_l(\phi)) - s.$$

We have the following convex nonlinear programming problem:

$$\left|\begin{array}{ll} \min\limits_{\phi,s} \frac{1}{N\alpha}\sum_{l=1}^{N}\eta_\epsilon(s - R_l(\phi)) - s & (NP^a) \\ \text{s.t. } \mathbb{E}(R(\phi)) \geq \mu_p^* & \\ \phi \in \Phi \subset \mathbb{R}^n & \\ s \in \mathbb{R} & \end{array}\right.$$

The $CVaR$ minimization technique, described in this section, can be used with different schemes for generating scenarios. For example, one can assume a joint distibution for the return process for all assets and generate scenarios in a Monte Carlo simulation. As well, the approach allows using of historical data without assuming a particular distribution. Additionally, an approximate value of $VaR_\alpha(R(\phi))$ when $CVaR_\alpha(R(\phi))$ reaches its minimum is obtained. It is given as the negative value of s^* in the minimum.

REMARK 2.1.4. 1. Alexander et al. [2003] observe that (NP^a) is more efficient than the linear programming method with up to 1100% efficiency speedup. The computational efficiency of (NP^a) is important when we are dealing with large number of scenarios and assets. On the other hand, it is shown [Alexander et al., 2003] that the proposed smoothing method (NP^a) yields accurate solutions.

2. Equivalent formulations of the problems $(P1)$ and $(P3)$ with $CVaR_\alpha(R(\phi))$ can be stated in a similar way.

2.2. CVaR and Monte-Carlo simulation

One of the advantages of the mean-CVaR model is that it allows using of scenarios. To simulate scenarios from a known distribution, Monte Carlo method is used. Our goal is to test the convergence of the Monte-Carlo based CVaR optimal portfolios when an analytical solution of the optimization problem is known. We use the approach, suggested in [Abad and Hurd, 2004]. The discussion is based on [De Giorgi, 2002] and [Kamdem, 2004], where formulas for $VaR_\alpha(R)$ and $CVaR_\alpha(R)$ as explicit functions of the portfolios weights are derived. Let's assume that the vector r of assets' relative returns is multivariate Gaussian distributed, and consider the following optimization problem:

$$\left| \begin{array}{ll} \min\limits_{\phi} CVaR_\alpha(R(\phi)) & \qquad (CVaR*) \\ \text{s.t. } \mu_p(R(\phi)) = \mu_p^* > 0 & \\ \phi^T e = 1 & \end{array} \right.$$

PROPOSITION 2.2.1. (VaR_α and $CVaR_\alpha$ for normal distributions)*: Let $\alpha \in (0,\, 0.5)$ and the vector r is multivariate Gaussian distributed with mean μ and variance-covariance matrix V. Then $R^\phi \sim N(\mu_p,\, \sigma_p^2)$, where μ_p and σ_p^2 are given by formulas 2.1.1 and 2.1.2, and*

$$VaR_\alpha^\phi(R) = z_{\alpha,1}\, \sigma_p^\phi - \mu_p^\phi, \tag{2.2.1}$$

with $z_{\alpha,1} = \Phi^{-1}(1-\alpha)$, $\Phi(.)$ – cdf of the standard normal distribution

$$CVaR_\alpha^\phi(R) = z_{\alpha,2}\, \sigma_p^\phi - \mu_p^\phi \tag{2.2.2}$$

with $z_{\alpha,2} = \rho(z_{\alpha,1})/\alpha$, $\rho(.)$ - density of the standard normal distribution

Since the coefficients z_α are positive, it is evident that the mean-CVaR optimization problem leads to the classical mean-variance Markowitz problem. An analytical solution of the problem

$$\left|\begin{array}{l} \min_{\phi} \sigma_p^2(R(\phi)) \qquad\qquad (MV*) \\ \text{s.t. } \mu_p(R(\phi)) = \mu_p^* > 0 \\ \phi^T e = 1 \end{array}\right.$$

where $e = (1, ..., 1)^T \in \mathbb{R}^n$, is given in [Merton, 1972]. We will formulate it according to De Giorgi [2002]. Under the assumption that the variance-covariance matrix V is strictly positive definite, a portfolio ϕ solves the problem $(MV*)$ iff:

(2.2.3)
$$\phi(\mu_p^*) = \mu_p^* \phi_0 - \phi_1$$

where

$$\phi_0 = \frac{1}{d}\,(b\,V^{-1}\,\mu - c\,V^{-1}\,e)$$

$$\phi_1 = \frac{1}{d}\,(c\,V^{-1}\,\mu - a\,V^{-1}\,e)$$

with

(2.2.4)
$$a = \mu^T\,V^{-1}\,\mu$$

$$b = e^T\,V^{-1}\,e$$

$$c = e^T\,V^{-1}\,\mu$$

$$d = a\,b - c^2$$

Formula 2.2.3 also sets the relation between the target expectation μ_p^* and the optimal standard deviation $\sigma(\mu_p^*)$, which is

$$\frac{\sigma^2}{1/b} - \frac{(\mu_p^* - c/b)^2}{d/b^2} = 1 \tag{2.2.5}$$

From the last equation it follows that the global minimum variance on the mean-variance boundary can be obtained with $\mu_p^* = c/b$. Hence we have the global minimum variance portfolio given by

$$\phi_{minVariance} = \frac{1}{c} V^{-1} \mu \tag{2.2.6}$$

PROPOSITION 2.2.2. *The global minimum CVaR portfolio exists iff* $z_{\alpha,2} > \sqrt{d/b}$. *It is given by*

$$\phi_{minCVaR} = \mu_{pminCVaR}\, \phi_0 - \phi_1 \tag{2.2.7}$$

where

$$\mu_{pminCVaR} = \frac{c}{b} + \sqrt{\frac{d}{b}\left(\frac{z_{\alpha,2}^2}{b\, z_{\alpha,2}^2} - \frac{1}{b}\right)} \tag{2.2.8}$$

REMARK 2.2.3. When $z_{\alpha,2} < \sqrt{d/b}$ and decreases, the $(\mu_p, CVaR_\alpha)$ boundary approches to a straight line with a slope -1. Therefore, an efficient portfolio doesn't exist. Moreover, a $(\mu_p, CVaR_\alpha)$ portfolio is an efficient portfolio only if we choose $\mu_p^* \geq \mu_{pminCVaR}$.

When the multivariate Gaussian distributed vector r is simulated, the numerical solution of $(CVaR*)$, denoted by ϕ^a, is an approximate solution depending on the number of Monte

Carlo simulations. We want to investigate the convergence of Monte Carlo based solutions with the increase of the simulation number. For this purpose, we observe the behaviour of the distance between the portfolios ϕ^a and ϕ^e. Research has been done for the case of a portfolio of stocks.

Case Study: Portfolio of stocks.

A commonly used model for stock price behavior is the geometric Brownian motion, which is also consistent with the Black-Scholes option-pricing model. We assume that the stock price process

$$S_t = \left(S_t^1,\, S_t^2, \dots, S_t^n\right)$$

is a multivariate geometric Brownian motion:

$$dS_t^i = \mu^i S_t^i dt + \sum_{j=1}^{n} \sigma^{ij} S_t^i \, dW_t^j, \tag{2.2.9}$$

where $\mu = (\mu^1,\, \mu^2, \dots, \mu^n)^T$ is a vector of instantaneous rates of return, and $(\sigma^{ij})_{1 \le i, j \le n}$ is a matrix of stock volatilities. The solution of the stochastic differential equation (2.2.9) is:

$$S_t^i = S_0^i \exp\left[\left(\mu^i - \frac{1}{2}(\sigma\sigma^T)^{ii}\right) t + \sum_{j=1}^{n} \sigma^{ij} W_t^j\right]$$

where S_0^i is the initial price. The vector of relative stock returns over a short time interval $\Delta t = [0,\, t]$, $r = (r_1,\, r_2, \dots, r_n)^T$, where $(r_i = \frac{S_t^i - S_0^i}{S_0^i})$, is approximately multivariate normally distributed, i.e $r \sim N(\mu\Delta t,\, (\sigma\sigma^T)\Delta t)$.

Next, for our real life application, we constructed a portfolio of $n = 10$ stocks from S&P100 index. The criterion for choosing an equity was last year's return > 5%. As well, the portfolio was diversified by including small (Black & Decker Corp. (BDK), Toys R Us Inc. (TOY)), mid (The AES Corp. (AES), Dell Inc. (DELL), Dow Chemical Co. (DOW), United Technologies Corp. (UTX)), large (Allstate Corp. (ALL)) and mega (Exxon Mobil Corp. (XOM), General Electric Co.(GE), Johnson & Johnson Inc. (JNJ)) caps. We collected 5 years of data set, consisting of 1252 adjusted closing daily prices for each equity, that is 1251 relative returns for the period May 8/00 to May 4/05. In other words, we utilized an historical time-series of $T = 1251$ returns for each stock. Our holding period is one day, and we suppose standing on the 4th of May, 2005. The estimated daily expected returns and covariances are given in table 1.

By applying the analytical method, described above, we solve the $(CVaR*)$ problem. For a chosen confidence level $\alpha = 0.01$, we obtain $z_{0.01} = 2.6652$. Since $z_{0.01} > \sqrt{\frac{d}{b}} = 0.0519$, a global minimum $CVaR_{0.01}$ exists, and $\mu_{pminCVaR} = 6.8965 \times 10^{-4}$. For an appropriate target expectation $\mu_p^* = 0.0008$, we have the efficient portfolio ϕ^e, displayed below, as a unique solution of our problem.

$equity$	AES	ALL	BDK	$DELL$	DOW	XOM	GE	JNJ	TOY	UTX
$weights, \phi^e$	-0.0023	0.3000	0.1257	0.0192	0.0137	0.2042	-0.1541	0.3585	0.0557	0.0792

The value of the objective function is $CVaR_{0.01}^{\phi^e} = 0.0282$. Also, the value of $VaR_{0.01}(R(\phi))$ when $CVaR_{0.01}(R(\phi))$ reaches its minimum is $VaR_{0.01}^{\phi^e} = 0.0245$.

The second part of our experiment is in generating scenarios for the random vector of the future daily returns for each stocks in our portfolio, $r = (r_1, r_2, ..., r_n)$. We use Monte Carlo simulation from the multivariate Gaussian distribution, thus satisfying the hypothesis of our model. As well, to solve $(CVaR*)$ we apply a linear programming approach from Section 2.1.3. We have

$$\left| \begin{array}{ll} \min\limits_{\phi,s,z} \frac{1}{N\alpha} \sum_{l=1}^{N} z_l - s & (LP*) \\ \text{s.t. } z_l \geq s - R_l(\phi), \; l = 1, ..., N & \\ z_l \geq 0 \; l = 1, ..., N & \\ \mathbb{E}(R(\phi)) = \mu_p^* & \\ \phi^T e = 1 & \\ s \in \mathbb{R} & \end{array} \right.$$

TABLE 1. *Variance-covariance matrix and expected returns*

equity	AES	ALL	BDK	$DELL$	DOW	XOM	GE	JNJ	TOY	UTX
$C(AES,.)$	0.00277	0.00015	0.00015	0.00029	0.00020	0.00014	0.00023	0.00006	0.00026	0.00023
$C(ALL,.)$	0.00015	0.00031	0.00009	0.00010	0.00010	0.00007	0.00012	0.00003	0.00008	0.00009
$C(BDK,.)$	0.00015	0.00009	0.00044	0.00019	0.00019	0.00008	0.00017	0.00005	0.00018	0.00019
$C(DELL,.)$	0.00029	0.00010	0.00019	0.00090	0.00015	0.00008	0.00026	0.00004	0.00022	0.00020
$C(DOW,.)$	0.00020	0.00010	0.00019	0.00015	0.00049	0.00011	0.00020	0.00007	0.00019	0.00021
$C(XOM,.)$	0.00014	0.00007	0.00008	0.00008	0.00011	0.00023	0.00011	0.00008	0.00008	0.00011
$C(GE,.)$	0.00023	0.00012	0.00017	0.00026	0.00020	0.00011	0.00042	0.00009	0.00017	0.00023
$C(JNJ,.)$	0.00006	0.00003	0.00005	0.00004	0.00007	0.00008	0.00009	0.00022	0.00005	0.00007
$C(TOY,.)$	0.00026	0.00008	0.00018	0.00022	0.00019	0.00008	0.00017	0.00005	0.00073	0.00017
$C(UTX,.)$	0.00023	0.00009	0.00019	0.00020	0.00021	0.00011	0.00023	0.00007	0.00017	0.00042
μ	0.00069	0.00096	0.00087	0.00020	0.00055	0.00052	0	0.00054	0.00073	0.00066

TABLE 2. *Minimum error versus* $\log_2$ *simulation size, LP problem*

	$\phi^e_{(.)}$	$\phi^{a(8)}_{(.)}$	$\phi^{a(9)}_{(.)}$	$\phi^{a(10)}_{(.)}$	$\phi^{a(11)}_{(.)}$	$\phi^{a(12)}_{(.)}$
AES	-0.0023	0.0131	0.0366	-0.0181	-0.0267	0.0106
ALL	0.3000	0.2518	0.2578	0.2974	0.2948	0.3259
BDK	0.1257	-0.0032	0.1123	0.1461	0.1291	0.1074
$DELL$	0.0192	0.0267	0.1421	-0.0647	0.0606	0.0299
DOW	0.0137	-0.0238	-0.0137	0.0479	0.0015	0.0409
XOM	0.2042	0.2655	0.2004	0.2843	0.2463	0.2362
GE	-0.1541	-0.0751	-0.2186	-0.1365	-0.0997	-0.1492
JNJ	0.3585	0.3877	0.3696	0.3425	0.3486	0.2959
TOY	0.0557	-0.0152	0.0225	0.0009	0.0070	0.0427
UTX	0.0792	0.1726	0.0911	0.1003	0.0384	0.0596
$\varepsilon_{min} = \left\Vert \phi^{a(.)} - \phi^e \right\Vert$	0	0.5712	0.3694	0.3466	0.2825	0.2271
$VaR_{0.01}(\phi^{(\cdot)})$	0.0245	0.0224	0.0248	0.0240	0.0241	0.0242
$CVaR_{0.01}(\phi^{(\cdot)})$	0.0282	0.0224	0.0285	0.0255	0.0273	0.0274

where the portfolio return function for each scenario is given by:

$$R_l(\phi,\, y_l) = \phi_1 r_{1,l} + \ldots + \phi_{10} r_{10,l}, \quad l = 1, \ldots, N$$

The aproximate solution ϕ^a of our optimization problem, given by the solution of $(LP*)$, depends on the number of generated scenarios N. Let us assume $N = 2^k$. Our interest is in observing whether the Monte Carlo approximate $\phi^{a(k)}$ converges to the true solution ϕ^e as k increases. In particular, we observe the L_1-distance between the vectors $\phi^{a(k)}$ and ϕ^e, $\varepsilon = \left\Vert \phi^{a(k)} - \phi^e \right\Vert = \sum_{i=1}^{n} \left| \phi_i^{a(k)} - \phi_i^e \right|$ for each value of $k = \log_2(N)$ in some fixed range. We use the linear programming problem $(LP*)$ for our data set, simulating scenarios for $k \in [8,\, 12]$ and executing the program 20 times for each k. These calculations yield 20 estimates $\phi^{a(k)}$ for each k in the range. The error values, denoted by $\varepsilon^{(k)}$, are displayed on figure 2. The average error is shown as a function of the $\log_2$ simulation size. We use the Matlab optimization toolbox to solve the minimization problem.

Table 2 displays the exact portfolio ϕ^e, and the portfolios $\phi^{a(k)}$ for which a minimum error is obtained in the set of experiments.

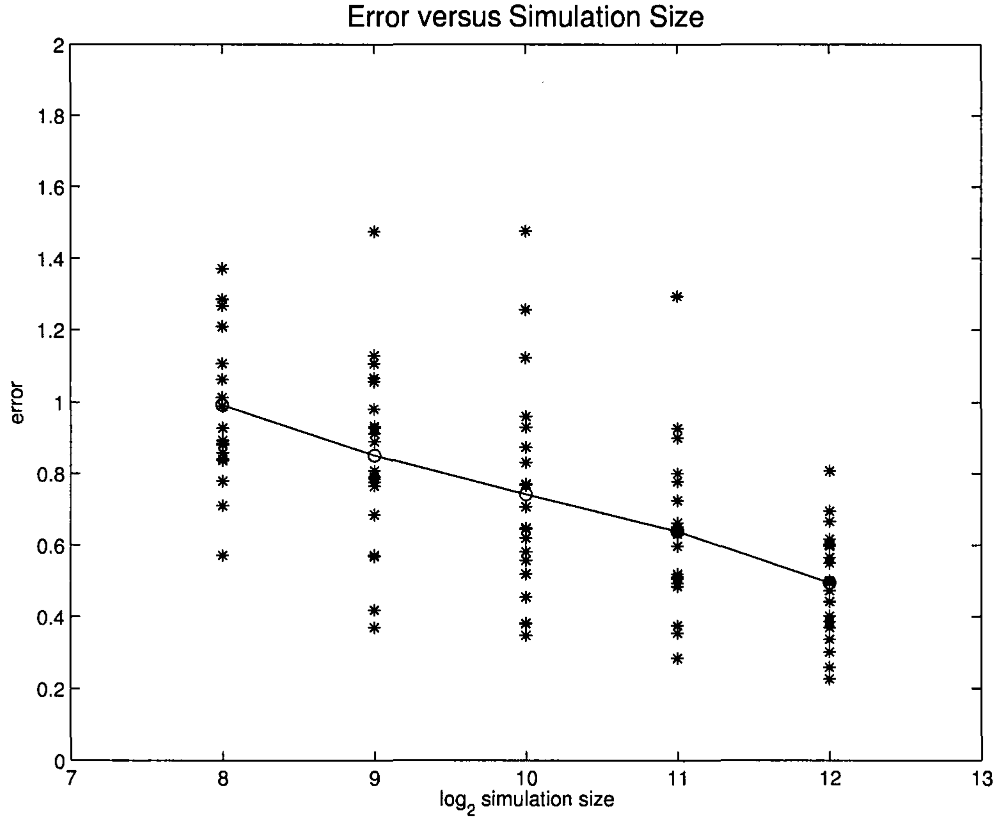

Figure 2: *Observed average error as a function of* $\log_2$ *simulation size, LP problem,* $\alpha = 0.01$

TABLE 3. *Average estimates versus simulation size, LP problem*

N	error (av)	$VaR^{(N)}$ (av)	$VaR^{(N)}$ (diff.%)	$CVaR^{(N)}$(av)	$CVaR^{(N)}$(diff.%)
2^8	0.9922	0.0228	-6.76	0.0231	-18.05
2^9	0.8510	0.0242	-1.27	0.0255	-9.70
2^{10}	0.7415	0.0249	1.53	0.0270	-4.09
2^{11}	0.6389	0.0250	1.98	0.0279	-1.01
2^{12}	0.4962	0.0249	1.68	0.0281	-0.51

As well, we compare the average values of the estimates for VaR_α and $CVaR_\alpha$ with the corresponding exact values. The obtained results are displayed in table 3. We see that while $VaR_\alpha^{(N)}(av.)$ values differ from the exact value by only few percentages, the convergence of the $CVaR_\alpha^{(N)}(av.)$ estimates is slower. Besides, $CVaR_\alpha$ appears to be underestimated in the most cases.

Since the smoothing method, where the sample objective function $F_\alpha^{(N)}(R(\phi),\, s)$ is replaced by the approximate sample objective function $F_\alpha^{a(N)}(R(\phi),\, s)$ (eq.2.1.5), is more computationally efficient compared to the linear programming method, we conduct the same experiments for $k = [14,\ 17]$ by using the following convex nonlinear program.

$$
\left|
\begin{array}{ll}
\min\limits_{\phi,s} \frac{1}{N\alpha} \sum_{l=1}^{N} \eta_\epsilon(s - R_l(\phi)) - s & \qquad (NP^*) \\
\text{s.t. } \mathbb{E}(R(\phi)) = \mu_p^* & \\
\phi^T e = 1 & \\
s \in \mathbb{R} &
\end{array}
\right.
$$

Moreover, when the number of samples N increases the difference between the function $F_\alpha^{(N)}(\phi,\, s)$ and $F_\alpha^{a(N)}(\phi,\, s)$ becomes smaller. This difference depends also on the resolution parameter ϵ, typically set to have a value ≤ 0.05 [Alexander et al., 2003]. Smaller values of ϵ lead to a better approximation when simulation size N is large. On the other hand, smaller values of ϵalso lead to increased computational time. The resolution parameter used here is 0.0001. The implementation of the both methods is made on a machine with Pentium IV processor and 1GB RAM. The average CPU time for a linear method with number of simulations $N \leq 2^{10}$ is less than 60 sec, whereas it is 10 min for $N = 2^{11}$ and respectively 110 min for $N = 2^{12}$. Due to less memory requirement, the nonlinear method (NP^a) with $N = 2^{17}$ can be solved in less than 25 CPU min. For comparison, the (NP^a) requires an average CPU time of 1 min for number of simulations $N = 2^{14}$, 2 min for $N = 2^{15}$, and 5 min when $N = 2^{16}$.

Note that the implementation of the smoothing method is based on a standard algorithm for nonlinear minimization in Matlab. The observed results are displayed on figure 3 and in tables 4 and 5. We conclude that $CVaR$ asset allocation problem based on Monte Carlo simulation yields a good approximate solution of $CVaR$ when the sample size is relatively large.

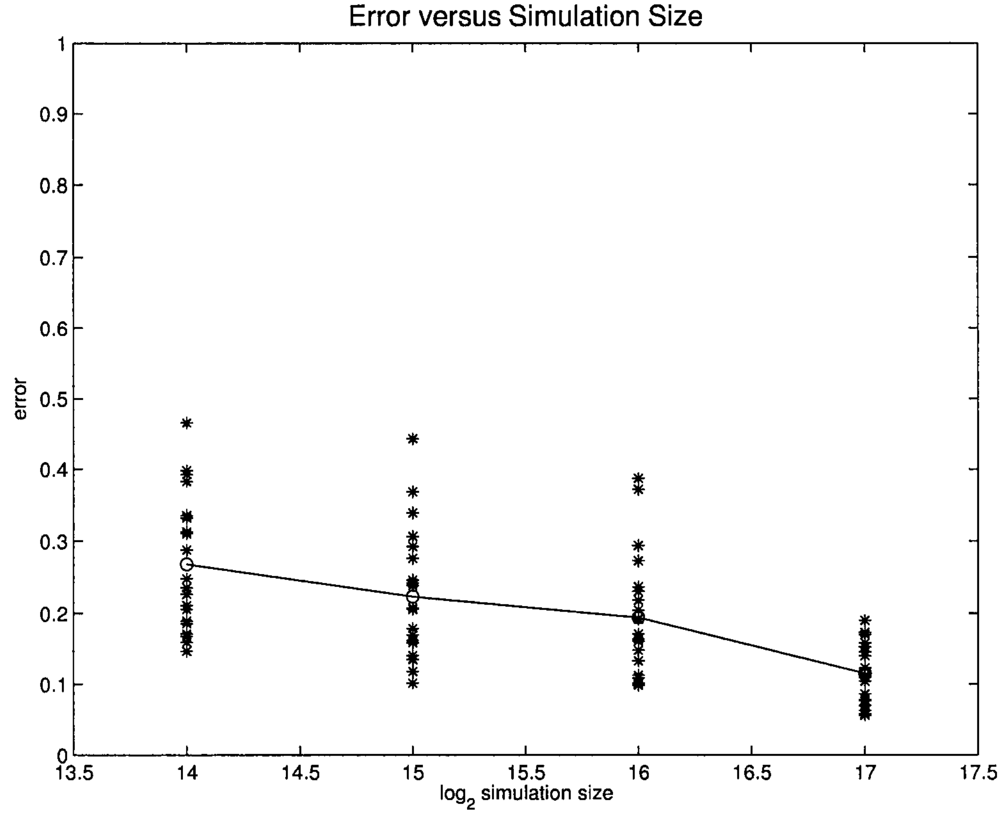

Figure 3: *Observed average error as a function of* $\log_2$ *simulation size, CVaR problem solved as a convex nonlinear programming problem,* $\alpha = 0.01$,$\epsilon = 0.0001$

TABLE 4. *Minimum error versus* $\log_2$ *simulation size, NP problem*

	$\phi^e_{(.)}$	$\phi^{a(14)}_{(.)}$	$\phi^{a(15)}_{(.)}$	$\phi^{a(16)}_{(.)}$	$\phi^{a(17)}_{(.)}$
AES	-0.0023	-0.0098	-0.0027	-0.0024	-0.0043
ALL	0.3000	0.2666	0.3004	0.2969	0.2945
BDK	0.1257	0.1155	0.1184	0.1169	0.1444
$DELL$	0.0192	0.0128	0.0234	0.0109	0.0108
DOW	0.0137	0.0205	0.0441	0.0206	0.0102
XOM	0.2042	0.2346	0.1844	0.2424	0.1963
GE	-0.1541	-0.1205	-0.1382	-0.1545	-0.1502
JNJ	0.3585	0.3561	0.3435	0.3399	0.3506
TOY	0.0557	0.0578	0.0540	0.0599	0.0561
UTX	0.0792	0.0663	0.0728	0.0694	0.0845
$\varepsilon_{min} = \left\|\phi^{a(\cdot)} - \phi^e\right\|$	0	0.1466	0.1014	0.0984	0.0562
$VaR_{0.01}(\phi^{(\cdot)})$	0.0245	0.0244	0.0243	0.0243	0.0245
$CVaR_{0.01}(\phi^{(\cdot)})$	0.0282	0.0278	0.0282	0.0278	0.0281

TABLE 5. *Average estimates versus simulation size, NP problem*

N	error (av)	$VaR^{(N)}$ (av)	$VaR^{(N)}$ (diff.%)	$CVaR^{(N)}$(av)	$CVaR^{(N)}$(diff.%)
2^{14}	0.2681	0.02453	0.10	0.02831	0.37
2^{15}	0.2231	0.02460	0.39	0.02822	0.05
2^{16}	0.1948	0.02454	0.14	0.2828	0.26
2^{17}	0.1154	0.02447	-0.14	0.0281	-0.37

CHAPTER 3

CVaR Relative Portfolio Optimization

3.1. Relative Portfolio Optimization in Practice

Relative portfolio optimization has emerged from practice with the implementation of various portfolio optimization algorithms. Since the asset managers' performances are often evaluated with reference to a benchmark, managers are concerned not only with absolute return, but also with return relative to the benchmark. In fact, instead of implementing the classical mean-variance approach, they optimize portfolios focusing on the difference in a managed portfolios' return and the return of a chosen benchmark portfolio. The random variable describing this difference is known as "excess", "relative", "active" or "differential" portfolio return, as well as a "tracking error". Although the tracking error is mostly understood as the standard deviation of the relative return, we use the term "tracking error" to refer to the random variable itself. We define the variable tracking error R_E by the expression [Roll, 1992]:

$$R_E = R - R_B,$$

where R denotes the managed portfolio return, and R_B is the benchmark return.

When evaluating relative performance against the benchmark in reward and risk terms, the reward parameter is therefore represented by the mean tracking error, and the risk measure is represented by the tracking error variance during the evalution period. The main assumption is that the portfolio manager desires to minimize the tracking error variance $\sigma^2(R_E(\phi)) = (\phi - b)^T V(\phi - b)$, while maximizing the expected excess return $\mathbb{E}(R_E(\phi)) = (\phi - b)^T \mu$. Similarly to the mean-risk model $(MR*)$, we formulate the tracking error problem as a problem of the biobjective programming:

$$\left| \begin{array}{ll} \max_{\phi} \mathbb{E}(R_E(\phi)) & \qquad (TEV) \\ \min_{\phi} \sigma^2(R_E(\phi)) & \\ \text{s.t. } \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

where $b = (b_1, ..., b_n)$ is the vector of fixed benchmark weights, and $b_i > 0$ for $i = 1, ..., n$.

In the literature, models of relative portfolio optimization have gained interest since the publication of Roll's work [Roll, 1992], considering the optimization problem:

$$\left| \begin{array}{ll} \min_{\phi} \sigma^2(R_E(\phi)) & \qquad (TEV*) \\ \text{s.t. } \mathbb{E}(R_E(\phi)) = \mu_E^* > 0 & \\ (\phi - b)^T e = 0 & \end{array} \right.$$

An analytical solution of the problem is given in the above paper. Note also, that if the constant tracking error expectation μ_E equals zero, $\phi = b$ is a trivial solution. An overview of the recent models can be found in [Wagner, 2003].

3.2. Relative CVaR

All the models in [Wagner, 2003] have in common that the relative portfolio risk is represented by the tracking error variance. The *linear models* of tracking error minimization are introduced by Markus et al. [1999]. The authors argue that the linear deviations between the benchmark and portfolio returns give a more accurate description of the investors' risk attitude than the squared deviations. Their models are based on minimizing the absolute deviations between the returns. *Expected regret* is another risk measure which is utilized in a model for benchmark portfolio optimization [Dembo and Rosen, 2000]. Dembo and Rosen [2000] solve the problem of mean-expected regret portfolio optimization by using the scenario approach. In contrast to the mean-tracking error variance optimization, the scenario approach allows for general non-normal distributions (see the discussion in the previous chapter). Another application of the risk measures to the excess portfolio return is presented in [Watson and Mina, 2000]. A measure, labeled *"Relative VaR_α"*, is defined as the level of tracking error that will not be exceeded over the chosen time period with an assigned confidence level α. As well, Jorion [2003] notes that constraining tracking error variance is equivalent to constraining *tracking error* VaR_α (or *relative* VaR_α) when the expected return is omitted in computing VaR_α. In this case, however, normal distributions of the returns are assumed. Taking into account the advantages of the Conditional Value at Risk as a risk measure, we propose CVaR techniques to be applied to the distribution of the excess return, or in a more general setting, to the profit and loss distribution of the so called "differential" portfolio. Moreover, we don't make any assumption about the form of the distribution. We call this measure *"Relative $CVaR_\alpha$"* and explain the similarity with the formulas for $CVaR_\alpha$, measuring total portfolio risk. For example, for continuous distributions *relative* $CVaR$ at confidence level α is given by the formula:

$$CVaR_\alpha(R_E) = -\mathbb{E}[R_E/R_E \leq -VaR_\alpha(R_E)],$$

where $VaR_\alpha(R_E)$ denotes *relative* VaR_α

$$VaR_\alpha(R_E) = -\inf\{x \in \mathbb{R},\ F_{R_E}(x) > \alpha\}$$

We formulate the mean-*relative* $CVaR$ optimization problem as:

$$\left| \begin{array}{ll} \max\limits_{\phi} \mathbb{E}(R_E(\phi)) & \qquad (RCVaR) \\ \min\limits_{\phi} CVaR_\alpha(R_E(\phi)) & \\ \text{s.t. } \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

Note that the discussions from Section 2.1.3, in terms of using scenarios and sample objective function, hold here.

Case study: Relative CVaR.

A benchmark is simply a reference portfolio. Let us suppose that our benchmark is a portfolio, which includes the 10 equities from Chapter 2, weighted according their market capitalization. Note, that this is the principle of construction of a market index. Standing on the 4th of May, 2005, we have:

$equity$	AES	ALL	BDK	$DELL$	DOW	XOM	GE	JNJ	TOY	UTX
$weights, b$	0.0087	0.0316	0.0056	0.0715	0.0374	0.3070	0.3200	0.1703	0.0046	0.0433

The objective is to track the given target portfolio as closely as possible. We want to illustrate the form of the mean-*relative* $CVaR_\alpha$ efficient frontier for different values of the confidence level α. In order to construct it, we minimize the relative risk (*relative* $CVaR_\alpha$) subject to attaining an expected excess return larger or equal to a given level. We design this case study as a demonstration of the methodology of using a sample objective function, representing *relative* $CVaR_\alpha$, rather than as a demonstration of using a particular multivariate distribution of the vector of assets' returns. Therefore we can simply simulate from a multivariate Gaussian distribution, and apply one of the numerical approaches, decribed in Section 2.1.3. We want to show the results, obtained by using the smoothing method with relation to the *relative return* R_E. Hence, the *relative* $CVaR_\alpha$ minimization implies minimization of the approximate function $F_\alpha^{a(N)}(R_E(\phi),\, s)$. We have the following convex nonlinear problem

$$\left|\begin{array}{l} \min\limits_{\phi,s} \frac{1}{N\alpha} \sum_{l=1}^{N} \eta_\epsilon(s - R_{E,l}(\phi)) - s \qquad (RNP^a) \\ \text{s.t. } (\phi - b)^T \mu \geq \mu_E^* \\ \phi \in \Phi \subset \mathbb{R}^n \\ s \in \mathbb{R} \end{array}\right.$$

where

$$R_{E,l}(\phi,\, r_l) = (\phi_1 - b_1)\, r_{1,l} + \ldots + (\phi_{10} - b_{10})\, r_{10,l}, \quad l = 1, \ldots, N$$

TABLE 1. Mean-Relative CVaR

mean, $\mu(R_E)$	0.0001	0.0002	0.0003	0.0004	0.0005
$CVaR^{(N)}_{0.01}(R_E)$	0.0051	0.0102	0.0153	0.0203	0.0262
$CVaR_{0.01}(R_E)$	0.0050	0.0101	0.0152	0.0207	0.0287
$\varepsilon = \left\|\phi^{(N)}_{0.01} - \phi\right\|$	0.0996	0.1941	0.2893	0.2661	0.3968
$CVaR^{(N)}_{0.02}(R_E)$	0.0047	0.0093	0.0140	0.0186	0.0240
$CVaR_{0.02}(R_E)$	0.0046	0.0091	0.0137	0.0188	0.0260
$\varepsilon = \left\|\phi^{(N)}_{0.02} - \phi\right\|$	0.1005	0.2004	0.2994	0.2807	0.4252
$CVaR^{(N)}_{0.03}(R_E)$	0.0044	0.0087	0.0131	0.0175	0.0225
$CVaR_{0.03}(R_E)$	0.0043	0.0085	0.0128	0.0176	0.0244
$\varepsilon = \left\|\phi^{(N)}_{0.03} - \phi\right\|$	0.1000	0.2006	0.3001	0.3082	0.4304

The program is executed for number of simulations $N = 25000$, parameter $\epsilon = 0.0005$ and confidence levels $\alpha = 0.01$, $\alpha = 0.02$ and $\alpha = 0.03$.

Figure 4 clearly shows that the behaviour of *relative* $CVaR_\alpha$ is similar to that of $CVaR_\alpha$, measuring the total portfolio risk, where with the increase of the confidence level α the value of the risk measure also increases for the same value of expected return. It can be noted here, that small percentages of outperformance are related to small differences in the *relative* $CVaR$ values at different confidence levels. The obtained results are also displayed in table 1. Since we simulate from multivariate normal distribution, we can use quadratic programming solutions to the method of minimizing *relative* $CVaR_\alpha$in order to compare the results from the (RNP^a) program. The error between the approximate and quadratic programming solution is calculated at each value of the expected return, and the results are displayed in table1 (Mean - Relative CVaR).

3.3. CVaR and Mean-Multirisk Relative Portfolio Problem

We note first that in the literature the problem of mean-multirisk portfolio optimization has been investigated under the assumption of normal distribution of the portfolio return. All the models introduced in this section have this assumption. As noted by Roll [1992] and Jorion [2003] institutional investors often manage money against a benchmark. When

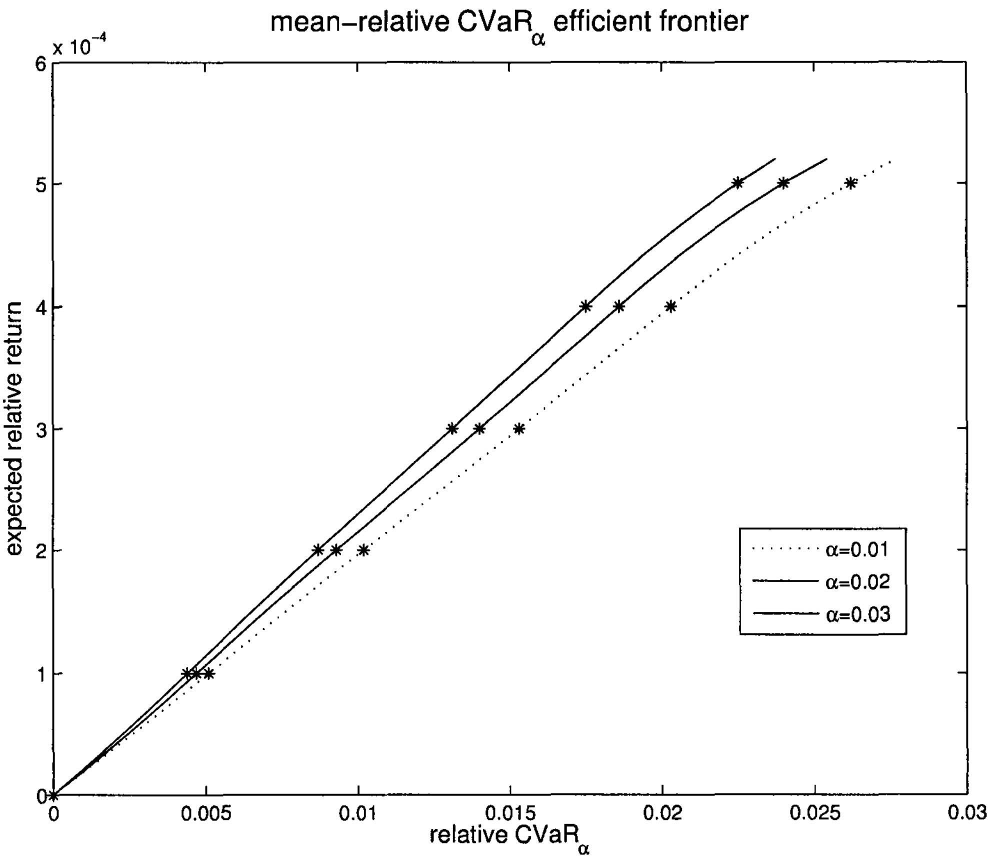

Figure 4: *Mean-Relative* $CVaR_\alpha$, confidence level $\alpha = 0.01$; 0.02; 0.03, simulation size $N = 25000$, parameter $\epsilon = 0.0005$

the investor's goal is to beat the return of a benchmark by a given percentage, the issue is whether the added value is in line with the risk undertaken. This leads to the formulation of the mean-risk problem in excess return space, which we discussed in the previous section. However, the problem with this setup is that the overall portfolio risk is totally ignored. Roll [1992] points out that for a given mean-variance inefficient benchmark portfolio, the excess return optimization $(TEV*)$ yields mean-variance inefficient solutions. In fact, one always assumes that the benchmark is not on the mean-variance efficient frontier; otherwise there is no need of outperformance. Moreover, the solution of $(TEV*)$ is independent of the benchmark. Jorion [2003] investigates whether the $(TEV*)$ problem can be corrected with additional restrictions on the relative portfolio. Since the tracking error variance constraint is widely used in practice, he considers the problem

$$\left| \begin{array}{ll} \max\limits_{\phi} \mathbb{E}(R_E(\phi)) & (TEV**) \\ \text{s.t. } \sigma^2(R_E(\phi)) = T > 0 & \\ (\phi - b)^T e = 0 & \end{array} \right.$$

and by imposing an additional total return volatility constraint above, in particular variance $Var(R(\phi)) = Var^*$, he shows that the selected portfolio dominates the benchmark in the mean-variance space. Another method for overcoming the problem of investing in inefficient portfolios is proposed in [Alexander and Baptista, 2005]. The authors show that adding a Value at Risk, $VaR_\alpha(R(\phi))$, constraint to the $(TEV*)$ model, examined by Roll [1992], leads to a selection of a portfolio that is also closer to the mean-variance efficient frontier. These two papers ([Alexander and Baptista, 2005, Jorion, 2003]) comprehensively interpret relative portfolio allocation solutions with an additional portfolio total risk constraint. Independently of them, several other models in a multirisk framework have been proposed in the literature. The model of Chow [1995] and Zhang [1998] can be regarded as a mean-multiple-variance model. It is given by the following minimization problem:

$$\left| \begin{array}{ll} \min\limits_{\phi} \sigma^2(R_E(\phi)) + k\,\sigma^2(R(\phi)) & (EMV) \\ \text{s.t. } \mathbb{E}(R(\phi)) = \mu_p^* & \\ \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

where $k \geq 0$.

Note that Chow, 1995 and Zhang, 1998 simply add the tracking error variance to the objective function of the classical Markowitz problem.

A slightly different approach is followed by Wagner [2002]. He derives the objective function from a multi-attribute utility theory. By introducing a regret aversion coefficient λ, he formulates the optimization problem:

$$\left| \begin{array}{ll} \min_{\phi} \sigma^2(R(\phi)) - \lambda Cov(R(\phi), R_B) & (EVC) \\ \text{s.t. } \mathbb{E}(R(\phi)) = \mu_p^* & \\ \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

It is easy to check that if $\lambda = \frac{2k}{1+k}$, the mean-variance-covariance model (EVC) is equivalent to (EMV). Note also that $\mathbb{E}(R) = \mathbb{E}(R_E) + \mu_B$, where $\mu_B = \mu^T b$ is a constant benchmark portfolio return. Therefore having $\mathbb{E}(R)$ in (EMV) and (EVC) doesn't affect our discussion. The common idea in all of the above models is that when optimizing with respect to a benchmark, the total risk should be taken under consideration. Moreover, Jorion [2003] shows that adding a volatility constraint to the problem of excess return maximization subject to a given level of tracking error variance $(TEV**)$ leads to the selection of a portfolio with smaller volatility but also with a smaller expected excess return (and with the same tracking error variance). Similar conclusions can be made for the others optimization problems too. In fact we are dealing with conflicting objectives in terms of multiple risks and expected return, and this is our motivation to investigate the discussed models as multiobjective optimization models. We summarize and develop a multiple criteria formulation of the relative portfolio allocation problem, represented by the above models:

$$\left| \begin{array}{ll} \max_{\phi} \mathbb{E}(R_E(\phi)) & (MTEV) \\ \min_{\phi} \sigma^2(R_E(\phi)) & \\ \min_{\phi} \rho\,(R(\phi)) \quad \text{s.t. } \phi \in \Phi \subset \mathbb{R}^n & \end{array} \right.$$

where $\rho(R)$ is given by the variance $\sigma^2(R)$, or Value at Risk $VaR_\alpha(R)$. Further, we propose Conditional Value at Risk, $CVaR_\alpha(R)$, to be used as a portfolio total risk measure in the relative portfolio optimization problem $(MTEV)$.

Case study: Mean - Multirisk Problem.

Our experiment is in generating a Pareto optimal (efficient) solution of the tricriterion $(MTEV)$ problem with $CVaR_\alpha$ as a portfolio risk measure and investigate the ideas, implemented in [Jorion, 2003, Alexander and Baptista, 2005]. For this purpose we consider the benchmark portfolio, described in the previous section. To generate an efficient solution we use an a priori method of multiobjectve optimization, in particular an approach of the goal programming. The method involves setting a set of goals, expressing the desicion maker's preferences for the values of the objective functions. The initial goals can be under- or overachieved with relative degree of under or overachievement, depending on a vector of positive weighting coefficients. Note that different solutions can be obtained by altering the weights. The scalarizing problem to be solved is:

$$\left| \begin{array}{l} \min\limits_{\phi,\gamma} \sum_{j=1}^{3} \gamma_j \qquad\qquad\qquad (GP) \\ \text{s.t. } f_1(\phi) - w_1\gamma_1 \leq goal_1 \\ f_2(\phi) - w_2\gamma_2 \leq goal_2 \\ f_3(\phi) - w_3\gamma_3 \leq goal_3 \\ \phi \in \Phi \subset \mathbb{R}^n \\ \gamma \in \mathbb{R}^3 \end{array} \right.$$

where the vector function is $F(\phi) = (-\mathbb{E}(R_E(\phi)), \sigma^2(R_E(\phi)), CVaR_{0.01}(R(\phi)))$, and γ is the vector of deviations from the goals.

Let us define the following goal point in the criterion space

$$goal = (-0.0005,\ 0.0002,\ 0.05)$$

and set the weighting vector $w = \text{abs}(goal)$. Note that calculating the benchmark statistics $\mu_B = 3.596\text{x}10^{-4}$, $\sigma_B^2 = 1.656x10^{-4}$, as well as $CVaR_{0.01}(R_B) = 0.0339$ assists us in setting the goals.

The obtained Pareto optimal solution for the criterion vector (representing a point from the efficient surface of $(MTEV)$) is:

$$\mu(R_E^{\phi_1}) = 3.7084\text{x}10^{-4},\ \sigma^2(R_E^{\phi_1}) = 5.3413\text{x}10^{-5},\ \text{and}\ CVaR_{0.01}(R^{\phi_1}) = 0.029.$$

The corresponding efficient solution is

$$\phi_1 = (0.0105,\ 0.2902,\ 0.1463,\ 0,\ 0,\ 0.1875,\ 0,\ 0.2047,\ 0.0454,\ 0.1153).$$

For this portfolio we have expected return

$$\mu(R^{\phi_1}) = 7.3046\text{x}10^{-4}$$

and variance

$$\sigma^2(R^{\phi_1}) = 1.2465\text{x}10^{-4}.$$

It is easy to check that this is a mean-variance, $(\mu,\ \sigma^2)$, inefficient, as well as a mean-$CVaR_{0.01}$, $(\mu,\ CVaR_{0.01})$, inefficient portfolio, since there is a portfolio ϕ^* such that

$$(\mu(R^{\phi^*}) = 7.3046\text{x}10^{-4},\ \sigma^2(R^{\phi^*}) = 1.1866x10^{-4})$$

is a point on the $(\mu,\ \sigma^2)$ efficient fronier and resp.

$$(\mu(R^{\phi^*}) = 7.3046\text{x}10^{-4},\ CVaR_{0.01}(R^{\phi^*}) = 0.0283)$$

is a point on the $(\mu,\ CVaR_{0.01})$ efficient frontier. However, we want to show that solving tracking error minimization problem (TE) with $\mu^*(R_E) \geq 3.7084\text{x}10^{-4}$ leads to a selection of portfolio, which is also $(\mu,\ \sigma^2)$ and $(\mu,\ CVaR_{0.01})$ inefficient, but dominated by the solution of the tricriterion problem. We have

$$\phi_2 = (0.0179,\ 0.2743,\ 0.1505,\ 0.0001,\ 0.0022,\ 0.2031,\ 0,\ 0.1521,\ 0.0428,\ 0.1572)$$

as a solution of

$$\left|\begin{array}{ll} \min\limits_{\phi} \sigma^2(R_E(\phi)) & (TE) \\ s.t.\ \mu(R_E(\phi)) \geq 3.7084\text{x}10^{-4} & \\ \phi \in \Phi & \end{array}\right.$$

and

$$\sigma^2(R_E^{\phi_2}) = 5.2031\text{x}10^{-5},\ \mu(R^{\phi_2}) = 7.3046\text{x}10^{-4},$$

$$\sigma^2(R^{\phi_2}) = 1.3186\text{x}10^{-4},\ CVaR_{0.01}(R^{\phi_2}) = 0.0299.$$

Obviously, for the objective vectors we have the following relations of preference:

$$(\mu(R^{\phi_1}),\ \sigma^2(R^{\phi_1})) \succ (\mu(R^{\phi_2}),\ \sigma^2(R^{\phi_2}))$$

and

$$(\mu(R^{\phi_1}),\ CVaR_{0.01}(R^{\phi_1})) \succ (\mu(R^{\phi_2}),\ CVaR_{0.01}(R^{\phi_2})).$$

Hence the benchmark portfolio selection problem considered with an additional objective for the total portfolio risk produces preferred mean-risk vectors. We note here that the assumption of normal distribution allowed us to compare the solutions in both mean-variance and mean-$CVaR$ space. As well, the choice of a particular risk measure depends only on the investor and therefore when working with Conditional Value at Risk, *relative* $CVaR_\alpha$ can be used for measuring the active risk.

When we address optimization problems with three or more criteria, the theory of multi-objective optmization is mostly concerned with interactive algorithms for repetitively sampling the Pareto optimal criterion (nondominated) set until a final solution is obtained. The problem of tracking error optimization in a mean-multirisk framework allows implementation of an interactive decision making and thus taking into account of the investor's preferences during the interactive process.

Concluding Remarks

Optimal portfolio selection is a longstanding issue in both practice and academic research on portfolio theory. From theoretical perspective, portfolio optimization can be managed in the framework of risk management. Risk management methodologies assume some measure of risk that impacts the assets' allocation in the portfolio. Because of the practical importance of "Value at Risk" measure, several alternatives to VaR, overcoming some of its shortcomings, have been proposed in the literature. This projects deals with the portfolio optimization model with the "Conditional Value at Risk" measure. CVaR develops the ideas incorporating in the Value at Risk concept, as for example measuring the downside risk, but also shows advantages over VaR. The purpose of the first chapter was to present and compare the different CVaR definitions, existing in the literature, and to discuss CVaR in terms of a coherent and therefore "relevant" risk measure. In the second chapter we introduced the formulation of the mean-risk models of portfolio selection, and presented the CVaR minimization as a convex programming problem. Our goal was in investigating the convergence of Monte Carlo based CVaR optimal solutions when two different approaches to representation of the sample objective function are used. This allowed solving of linear and convex nonlinear programming problems. The obtained results showed that a large number of simulation is required in order to obtain a satisfactory solution in terms of the error between the exact and approximate (Monte Carlo) solution. In the last chapter we extended our investigation on CVaR and proposed the use of "relative CVaR" measure in the problems of active portfolio management. As well, we proposed and showed that by adding a CVaR criterion to the classical relative portfolio optimization problem, preferred (in mean-total risk space) portfolios are generated. This supported the conclusion from the analytical analysis of benchmark optimization problem under normality assumptions. The suggested technique can be applied to non-normal distributions and future research on this problem can be done.

Appendix:

Multiobjective optimization

Problem formulation.

The mathematical formulation of the classical multiple objective optimization problem is:

$$\left| \begin{array}{l} \text{"opt"}\ f(x) \\ \text{s.t. } x \in S \subset \mathbb{R}^n \end{array} \right.$$

where $f : \mathbb{R}^n \longrightarrow \mathbb{R}^k$, $f(x) = (f_1(x),\ f_2(x), ..., f_k(x))$ is a vector of the *objective functions (criteria)* $f_i(x)$ for $i = 1, ..., k$ and $k \geq 2$. The *decision vectors* $x = (x_1,\ x_2, ..., x_n)$ belong to the set of feasible solutions (admissible decisions) S, and "opt" stands for optimization. For simplicity we assume that all the objective functions are to be minimized. If an objective function f_i is to be maximized, it is equivalent to minimize the function $-f_i$. Thus we study the problem

$$\left| \begin{array}{ll} \min\limits_{x} \{f_1(x),\ f_2(x), ..., f_k(x)\} & \qquad (MOP) \\ \text{s.t. } x \in S \subset \mathbb{R}^n. & \end{array} \right.$$

If there is no conflict between the objective functions, then a solution can be found where every objective function attains its optimum. To avoid such trivial cases, the formulation of each multiobjective programming problem suggests that there does not exist a solution, which minimizes all objective functions simultaneously. We are in the situation of trying to minimize each objective to the "greatest extent possible", i.e. we search for these solutions x^* where the increase of some objective functions is improved only by sacrificing some other functions. This means that the objective functions are at least partly conflicting.

DEFINITION. *The decision vector $x^* \in S$ is an Pareto optimal (efficient) solution if there does not exist another decision vector $x \in S$, such that $f_i(x) \leq f_i(x^*)$ for all $i = 1, ..., k$ and $f_j(x) < f_j(x^*)$ for at least one index j.*

The vector $f(x^*)$ corresponding to the Pareto optimal (efficient) solution x^* is called a Pareto optimal or nondominated vector function.

Decision space vs. Criterion space.

It can be seen that whereas in the single-objective programming there is only the feasible region S in decision space $\mathbb{R}^n$, in multiobjective programming there is the feasible region $Z = \{z/z = f(x),\ x \in S\}$ in criterion space $\mathbb{R}^k$. Each $x \in S$ in decision space has an image $z \in Z$ in criterion space, and each $z \in Z$ in criterion space has at least one inverse image $x \in S$ in decision space. Note that regardless of the dimension of S, the dimension of Z is k, usually much smaller number than n. Then a criterion vector $z^* \in Z$ is nondominated iff there does not exists another vector $z \in Z$ s.t. $z_i \leq z_i^*$ for all $i = 1, ..., k$ and $z_j < z_j^*$ for at least one index j. Furthermore $x^* \in S$ is an efficient solution iff its image criterion vector $z^* = f(x^*) = (f_1(x^*), f_2(x^*), ..., f_k(x^*))$ is nondominated. Multiple objective programming is mostly studied in criterion space. The set of all efficient points is called the efficient set. The nondominated set (image of the efficient set) is in the boundary of the feasible region Z in criterion space. The nondominated points in the bi-objective optimization problem form the so-called nondominated frontier, while the nondominated points in the case of higher dimension form the nondominated surface. Note: Despite the fact, that nondominance is a criterion space concept and efficiency is only a decision space concept in terms of multiple criteria optimization, "efficient frontier" and "efficient surface" will be used when considering portfolio optimization problems as problems of multiple criteria optimization.

Solving multiobjective optimization problem.

Multiobjective optimization (also called multicriteria optimization, multiple objective mathematical programming) is defined as a problem of finding a vector of decision variables which satisfies constraints and optimizes a vector function whose elements represent the objective functions. Since the objectives are with conflict with each other, the term "optimize" means finding such a solution, which would give the values of all the objective functions acceptable to the desicion maker. In other words, we want to determine from among the set of feasible solutions S a particular Pareto vector $x = (x_1^*,\ x_2^*, ..., x_n^*)$ that is most acceptable to a user or decision-maker. Assuming such a solution exists, it is called a final solution. Multiobjective optimization problems are usually solved by scalarization. Scalarization means that the problem is converted into a single or a family of single objective optimization problems with a real-valued objective function, called *scalarizing function*, depending possibly on some parameters. The scalarizing finction is required to cover Pareto solutions. Hence the Pareto optimal solutions of multiobjective optimization problems can be characterized as solutions of certain single objective optimization problems. In general, there are many Pareto solutions. The final decision is made among them taking the total balance over all criteria into account. This is a problem of value judgment of the decision-maker (DM). Balancing over all criteria is called *trade-off*.

Methods of multiobjective optimization [Miettinen, 1998].

Methods of solving multiobjective optimization problems can be classified in many ways according to different criteria. Here we give a cllassification according to the participation of the decision maker in the solution process:

1. Methods where no articulation of preference information is used (no-preference methods).

2. Methods where a posteriori articulation of preference information is used (a posteriori methods) - examples: weighting method, $\varepsilon-$constraint method. After the Pareto

optimal set (or part of it) has been generated, it is presented to the DM, who selects the most preferred among the alternatives.

3. Methods where a priori articulation of preference information is used (a priori methods) - examples: goal programming. The DM specifies his/her preferences before the solution process.

4. Methods where progressive articulation of preference information is used (interactive methods) - most developed class of methods. Procedures of this type are characterized by phases of decision alternating with phases of computation. At each computation phase, a solution, or a subset of solutions, is generated for examination in the decision phase. As a result of the judgment, the DM inputs some preferential information, which intends to "improve" the proposed solution(s) generated in the next computation phase. The DM is "learning" about the problem and therefore building his/her preference model during the interactive process.

Bibliography

J. Abad and T.R. Hurd. An Error Formula for Monte Carlo Based Portfolio Optimization. *Working paper, McMasterUniversity, Department of Mathematics and Statistics*, February 2004.

P. Albrecht. Risk Measures. In *Encyclopedia of Actuarial Science*. John Wiley and Sons, 2004.

G. Alexander and A. Baptista. Active Portfolio Management with Benchmarking: Adding a Value at Risk Constraint. *Working paper, University of Minnesota*, February 2005.

S. Alexander, Coleman T., and Y. Li. Minimizing CVaR and VaR for a Portfolio of Derivatives. *International Conference on Modeling, Optimization , and Risk Management in Finance*, http://gloriamundi.org, 2003.

P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent Measures of Risk. *Math. Finance*, 9(3):203–228, 1999.

C. Aserbi. Coherent Representation of Subjective Risk-Aversion. In *Risk Measures for 21st Century*, pages 147–207. John Wiley and Sons, Ltd., 2004.

C. Aserbi and D. Tashe. On the Coherence of Expected Shortfall. *Journal of Banking and Finance*, 26(7):1487–1503, 2002a.

C. Aserbi and D. Tashe. Expected Shortfall: A Natural Coherent Alternative to Value at Risk. *Economic Notes*, 31(2):379–388, 2002b.

A. Chabaane, J.-P. Laurent, Y. Malevergne, and Fr. Turpin. Alternative Risk Measures for Alternative Investments. *International Conference AFFI 2004*, http://u-cergy.fr/.

S. Cheng, Liu Y., and Wang S. Progress in Risk Measurement. *Advanced Modeling and Optimization*, 6:1–20, 2004.

G. Chow. Portfolio Selection Based on Return, Risk, and Relative Performance. *Financial Analysts Journal*, March-April:54–60, 1995.

E. De Giorgi. A Note on Portfolio Selection Under Various Risk Measures. *Working paper, University of Zurich, Institute for Empirical Research in Economics*, No. 122, 2002.

F. Delbaen. Coherent Risk Measures on General Probability Spaces. In *Advances in finance and stochastics*, pages 1–37. Springer, Berlin, 2002.

R. Dembo. Scenario Optimization. *Annals of Operations Research*, 30:63–80, 1991.

R. Dembo and D. Rosen. The Practice of Portfolio Replication. *ALGO Research Quartely*, 3:11–22, 2000.

H. Follmer and A. Schied. *Stochastic Finance*, volume 27 of *de Gruyter Studies in Mathematics*. Walter de Gruyter & Co., Berlin, 2002.

J. Gotoh and H. Konno. Third Degree Stochastic Dominance and Mean Risk Analysis. *Management Science*, 46:289–301, 2000.

J.C. Hull. *Options, Futures and Other Derivatives*. Prenice-Hall, Upper Saddle River, NJ, 2003.

P. Jorion. Portfolio Optimization with Tracking Error Constraints. *Financial Analysts Journal*, 59:70–82, 2003.

J.S. Kamdem. Value at Risk and Expected Shortfall for Linear Portfolios with Elliptically Distributed Risk Factors. *PhD Thesis, Universite de Reims, www.gloriamundi.org*, 2004.

P Krokhmal, Palmquist J., and S. Uryasev. Portfolio Optimization with Conditional Value at Risk. *The Journal of Risk*, 4(2):43–68, 2002.

J.-P. Laurent. Sensitivity Analysis of Risk Measures for Discrete Distributions. *http://laurent.jeanpaul.free.fr*, 2003.

H. Levy. *Stochastic Dominance: Investment Decision Making Under Uncertainty*. Kluwer Academic Publisher, 1998.

H. M. Markowitz. *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Basil Blackwell, Oxford, 1987.

R. Markus, Wolter H.-J., and H. Zimmermann. A Linear Model for Tracking Error Minimization. *Journal of Banking and Finance*, 23:85–103, 1999.

R.C. Merton. An Analytic Derivation of the Efficient Portfolio Frontier. *Journal of Financial and Quantitative Analysis*, 7:1851–1872, 1972.

K. Miettinen. *Nonlinear Multiobjective Optimization.* Kluwer Academic Publishers, 1998.

Wl. Ogryczak and A. Ruszczynski. From Stochastic Dominance to Mean-Risk Models. *European Journal of Operational Research*, 116:33–50, 1999.

Wl. Ogryczak and A. Ruszczynski. On Consistency of Stochastic Dominance and Mean-Semideviation Models. *Mathematical Programming, Ser.B*, 89(2):217–232, 2001.

G. Pflug. Some Remarks on the Value at Risk and the Conditional Value at Risk. In S. Uryasev, editor, *Probabilistic Constrained Optimization - Methodology and Applications*, pages 272–281. Kluwer Academic Publishers, 2000.

R. Rockafellar, S. Uryasev, and M. Zabarankin. Deviation Measures in Risk Analysis and Optimization. Technical Report 2002-7, ISE Dept., University of Florida, December 2002.

R.T. Rockafellar and S. Uryasev. Optimization of Conditional Value-at-Risk. *The Journal of Risk*, 2(3):21–41, 2000.

R.T. Rockafellar and S. Uryasev. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance*, 27(7):1443–1471, 2002.

R. Roll. A Mean/Variance Analysis of Tracking Error. *Journal of Portfolio Management*, 18:13–22, 1992.

S. Uryasev. Conditional Value-at-Risk: Optimization Algorithms and Applications. *Financial Engineering News*, (14):1–5, 2000.

S. Uryasev and R. T. Rockafellar. Conditional Value-at-Risk: Optimization Approach. In *Stochastic Optimization: Algorithms and Applications (Gainesville, FL, 2000)*, volume 54 of *Appl. Optim.*, pages 411–435. Kluwer Acad. Publ., Dordrecht, 2001.

N. Wagner. On a Model of Portfolio Selection with Benchmark. *Journal of Asset Management*, 3:55–64, 2002.

N. Wagner. Methods of Relative Portfolio Optimization. In *Advances in Portfolio Construction and Implementation*. Elsevier Ltd., 2003.

G. Watson and J. Mina. The Value of Relative VaR. *www.gloriamundi.org*, 2000.

L.H. Zhang. Global Asset Allocation with Multirisk Considerations. *Journal of Investing*, Fall:7–14, 1998.