TUTORIALS IN
OPERATIONS RESEARCH
INFORMS 2008

© 2008 INFORMS | ISBN 978-1-877640-23-0
DOI 10.1287/educ.1080.0052

Value-at-Risk vs. Conditional Value-at-Risk in Risk Management and Optimization

Sergey Sarykalin
Gobal Fraud Risk Management, American Express, Phoenix, Arizona 85021,
sergey.sarykalin@aexp.com

Gaia Serraino and Stan Uryasev
Department of Industrial and Systems Engineering, Risk Management and Financial Engineering Lab, University of Florida, Gainesville, Florida 32611
{serraino@ufl.edu, uryasev@ufl.edu}

Abstract From the mathematical perspective considered in this tutorial, risk management is a procedure for shaping a risk distribution. Popular functions managing risk are value-at-risk (VaR) and conditional value-at-risk (CVaR). The problem of choice between VaR and CVaR, especially in financial risk management, has been quite popular in academic literature. Reasons affecting the choice between VaR and CVaR are based on the differences in mathematical properties, stability of statistical estimation, simplicity of optimization procedures, acceptance by regulators, etc. This tutorial presents our personal experience working with these key percentile risk measures. We try to explain strong and weak features of these risk measures and illustrate them with several examples. We demonstrate risk management/optimization case studies conducted with the Portfolio Safeguard package.

Keywords VaR; CVaR; risk measures; deviation measures; risk management; optimization; Portfolio Safeguard package

1. Introduction

Risk management is a broad concept involving various perspectives. From the mathematical perspective considered in this tutorial, risk management is a procedure for shaping a loss distribution (for instance, an investor's risk profile). Among the vast majority of recent innovations, only a few have been widely accepted by practitioners, despite their active interest in this area. Conditional value-at-risk (CVaR), introduced by Rockafellar and Uryasev [19], is a popular tool for managing risk. CVaR approximately (or exactly, under certain conditions) equals the average of some percentage of the worst-case loss scenarios. CVaR risk measure is similar to the value-at-risk (VaR) risk measure, which is a percentile of a loss distribution. VaR is heavily used in various engineering applications, including financial ones. VaR risk constraints are equivalent to the so-called chance constraints on probabilities of losses. Some risk communities prefer VaR others prefer chance (or probabilistic) functions. There is a close correspondence between CVaR and VaR: with the same confidence level, VaR is a lower bound for CVaR. Rockafellar and Uryasev [19, 20] showed that CVaR is superior to VaR in optimization applications. The problem of the choice between VaR and CVaR, especially in financial risk management, has been quite popular in academic literature. Reasons affecting the choice between VaR and CVaR are based on the differences in mathematical properties, stability of statistical estimation, simplicity of optimization procedures, acceptance by regulators, etc. Conclusions made from this properties may often be quite contradictive. Plenty of relevant material on this subject can be found at the website http://www.gloriamundi.org. This tutorial should not be considered as a review on VaR and

CVaR: many important findings are beyond the scope of this tutorial. Here we present only our personal experience with these key percentile risk measures and try to explain strong and weak features of these two risk measures and illustrate them with several examples. We demonstrate risk management/optimization case studies with the Portfolio Safeguard (PSG) package by American Optimal Decisions (an evaluation copy of PSG can be requested at http://www.AOrDa.com).

Key observations presented in this tutorial are as follows:

• CVaR has superior mathematical properties versus VaR. CVaR is a so-called "coherent risk measure"; for instance, the CVaR of a portfolio is a continuous and convex function with respect to positions in instruments, whereas the VaR may be even a discontinuous function.

• CVaR deviation (and mixed CVaR deviation) is a strong "competitor" to the standard deviation. Virtually everywhere the standard deviation can be replaced by a CVaR deviation. For instance, in finance, a CVaR deviation can be used in the following concepts: the Sharpe ratio, portfolio beta, one-fund theorem (i.e., optimal portfolio is a mixture of a risk-free asset and a master fund), market equilibrium with one or multiple deviation measures, and so on.

• Risk management with CVaR functions can be done quite efficiently. CVaR can be optimized and constrained with convex and linear programming methods, whereas VaR is relatively difficult to optimize (although significant progress was made in this direction; for instance, PSG can optimize VaR for quite large problems involving thousands of variables and hundreds of thousands of scenarios).

• VaR risk measure does not control scenarios exceeding VaR (for instance you can significantly increase the largest loss exceeding VaR, but the VaR risk measure will not change). This property can be both good and bad, depending upon your objectives:

◦ The indifference of VaR risk measure to extreme tails may be a good property if poor models are used for building distributions. VaR disregards some part of the distribution for which only poor estimates are available. VaR estimates are statistically more stable than CVaR estimates. This actually may lead to a superior out-of-sample performance of VaR versus CVaR for some applications. For instance, for a portfolio involving instruments with strong mean reverting properties, VaR will not penalize instruments with extremely heavy losses. These instruments may perform very well at the next iteration. In statistics, it is well understood that estimators based on VaR are "robust" and may automatically disregard outliers and large losses, which may "confuse" the statistical estimation procedure.

◦ The indifference of VaR to extreme tails may be quite an undesirable property, allowing to take high uncontrollable risks. For instance, so-called "naked" option positions involve a very small chance of extremely high losses; these rare losses may not be picked by VaR.

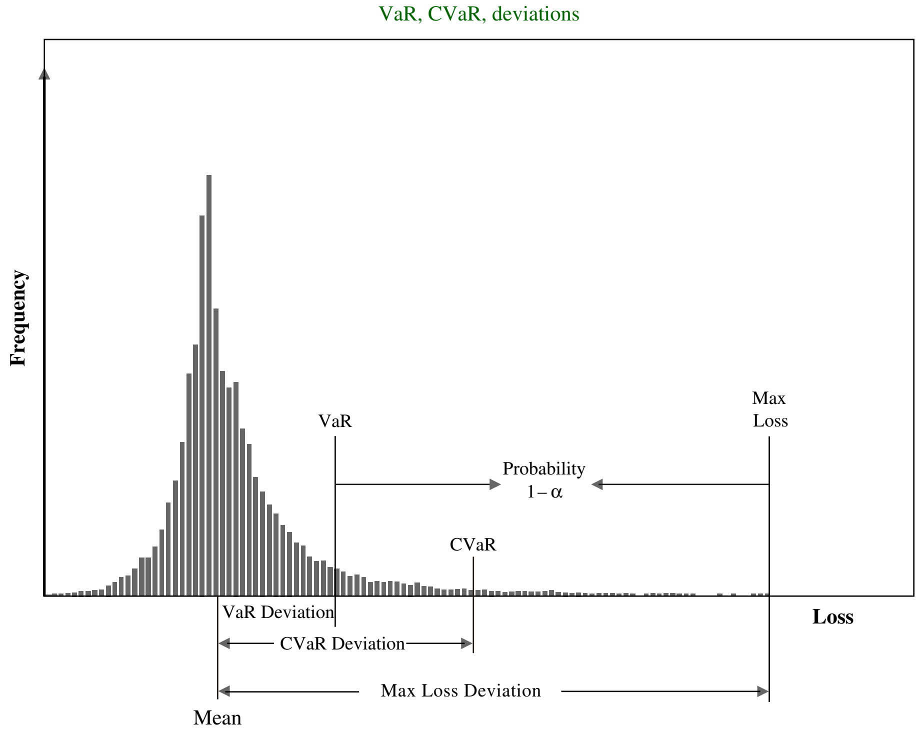
• CVaR accounts for losses exceeding VaR. This property may be good or bad, depending upon your objectives:

◦ CVaR provides an adequate picture of risks reflected in extreme tails. This is a very important property if the extreme tail losses are correctly estimated.

◦ CVaR may have a relatively poor out-of-sample performance compared with VaR if tails are not modelled correctly. In this case, mixed CVaR can be a good alternative that gives different weights for different parts of the distribution (rather than penalizing only extreme tail losses).

• Deviation and risk are quite different risk management concepts. A risk measure evaluates outcomes versus zero, whereas a deviation measure estimates wideness of a distribution. For instance, CVaR risk may be positive or negative, whereas CVaR deviation is always positive. Therefore, the Sharpe-like ratio (expected reward divided by risk measure) should involve CVaR deviation in the denominator rather than CVaR risk.

FIGURE 1. Risk functions: graphical representation of VaR, VaR Deviation, CVaR, CVaR Deviation, Max Loss, and Max Loss Deviation.



2. General Picture of VaR and CVaR

2.1. Definitions of VaR and CVaR

This section gives definitions of VaR and CVaR and discusses their use and basic properties. We refer to Figure 1 for their graphical representation.

Let X be a random variable with the cumulative distribution function $F_X(z) = P\{X \leq z\}$. X may have meaning of loss or gain. In this tutorial, X has meaning of loss and this impacts the sign of functions in definition of VaR and CVaR.

Definition 1 (Value-at-Risk). The VaR of X with confidence level $\alpha \in]0,1[$ is

$$\mathrm{VaR}_\alpha(X) = \min\{z \mid F_X(z) \geq \alpha\}. \tag{1}$$

By definition, $\mathrm{VaR}_\alpha(X)$ is a lower α-percentile of the random variable X. Value-at-risk is commonly used in many engineering areas involving uncertainties, such as military, nuclear, material, airspace, finance, etc. For instance, finance regulations, like Basel I and Basel II, use VaR deviation measuring the width of daily loss distribution of a portfolio.

For normally distributed random variables, VaR is proportional to the standard deviation. If $X \sim N(\mu, \sigma^2)$ and $F_X(z)$ is the cumulative distribution function of X, then (see Rockafellar and Uryasev [19]),

$$\mathrm{VaR}_\alpha(X) = F_X^{-1}(\alpha) = \mu + k(\alpha)\sigma, \tag{2}$$

where $k(\alpha) = \sqrt{2}\,\mathrm{erf}^{-1}(2\alpha - 1)$ and $\mathrm{erf}(z) = (2/\sqrt{\pi}) \int_0^z e^{-t^2}\,dt$.

The ease and intuitiveness of VaR are counterbalanced by its mathematical properties. As a function of the confidence level, for discrete distributions, $\mathrm{VaR}_\alpha(X)$ is a nonconvex, discontinuous function. For a discussion of numerical difficulties of VaR optimization, see, for example, Rockafellar [17] and Rockafellar and Uryasev [19].

Definition 2 (Conditional Value-at-Risk). An alternative percentile measure of risk is *conditional value-at-risk* (CVaR). For random variables with continuous distribution functions, $\text{CVaR}_\alpha(X)$ equals the conditional expectation of X subject to $X \geq \text{VaR}_\alpha(X)$. This definition is the basis for the name of conditional value-at-risk. The term conditional value-at-risk was introduced by Rockafellar and Uryasev [19]. The general definition of conditional value-at-risk (CVaR) for random variables with a possibly discontinuous distribution function is as follows (see Rockafellar and Uryasev [20]).

The CVaR of X with confidence level $\alpha \in]0,1[$ is the mean of the generalized α-tail distribution:

$$\text{CVaR}_\alpha(X) = \int_{-\infty}^{\infty} z\, dF_X^\alpha(z), \tag{3}$$

where

$$F_X^\alpha(z) = \begin{cases} 0, & \text{when } z < \text{VaR}_\alpha(X), \\ \dfrac{F_X(z)-\alpha}{1-\alpha}, & \text{when } z \geq \text{VaR}_\alpha(X). \end{cases}$$

Contrary to popular belief, in the general case, $\text{CVaR}_\alpha(X)$ is not equal to an average of outcomes greater than $\text{VaR}_\alpha(X)$. For general distributions, one may need to split a probability atom. For example, when the distribution is modelled by scenarios, CVaR may be obtained by averaging a fractional number of scenarios. To explain this idea in more detail, we further introduce alternative definitions of CVaR. Let $\text{CVaR}_\alpha^+(X)$, called "upper CVaR," be the conditional expectation of X subject to $X > \text{VaR}_\alpha(X)$:

$$\text{CVaR}_\alpha^+(X) = E[X \mid X > \text{VaR}_\alpha(X)].$$

$\text{CVaR}_\alpha(X)$ can be defined alternatively as the weighted average of $\text{VaR}_\alpha(X)$ and $\text{CVaR}_\alpha^+(X)$, as follows. If $F_X(\text{VaR}_\alpha(X)) < 1$, so there is a chance of a loss greater than $\text{VaR}_\alpha(X)$, then

$$\text{CVaR}_\alpha(X) = \lambda_\alpha(X)\text{VaR}_\alpha(X) + (1-\lambda_\alpha(X))\text{CVaR}_\alpha^+(X), \tag{4}$$

where

$$\lambda_\alpha(X) = \frac{F_X(\text{VaR}_\alpha(X)) - \alpha}{1-\alpha}, \tag{5}$$

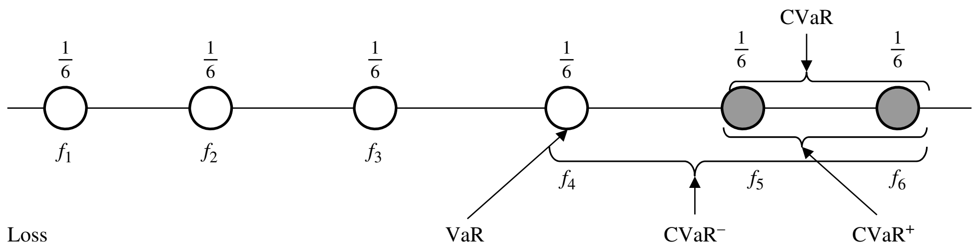
whereas if $F_X(\text{VaR}_\alpha(X)) = 1$, so that $\text{VaR}_\alpha(X)$ is the highest loss that can occur, then

$$\text{CVaR}_\alpha(x) = \text{VaR}_\alpha(x). \tag{6}$$

The definition of CVaR as in Equation (4) demonstrates that CVaR is not defined as a conditional expectation. The function $\text{CVaR}_\alpha^-(X) = E[X \mid X \geq \text{VaR}_\alpha(X)]$, called "lower CVaR," coincides with $\text{CVaR}_\alpha(X)$ for continuous distributions; however, for general distributions it is discontinuous with respect to α and is not convex. The construction of CVaR_α as a weighted average of VaR_α and $\text{CVaR}_\alpha^+(X)$ is a major innovation. Neither VaR nor $\text{CVaR}_\alpha^+(X)$ behaves well as a measure of risk for general loss distributions (both are discontinuous functions), but CVaR is a very attractive function. It is continuous with respect to α and jointly convex in (X,α). The unusual feature in the definition of CVaR is that VaR atom can be split. If $F_X(x)$ has a vertical discontinuity gap, then there is an interval of confidence level α having the same VaR. The lower and upper endpoints of that interval are $\alpha^- = F_X(\text{VaR}_\alpha^-(X))$ and $\alpha^+ = F_X(\text{VaR}_\alpha(X))$, where $F_X(\text{VaR}_\alpha^-(X)) = P\{X < \text{VaR}_\alpha(X)\}$. When $F_X(\text{VaR}_\alpha^-(X)) < \alpha < F_X(\text{VaR}_\alpha(X)) < 1$, the atom $\text{VaR}_\alpha(X)$ having total probability $\alpha^+ - \alpha^-$ is split by the confidence level α in two pieces with probabilities $\alpha^+ - \alpha$ and $\alpha - \alpha^-$. Equation (4) highlights this splitting.

The definition of CVaR is illustrated further with the following examples. Suppose we have six equally likely scenarios with losses $f_1 \cdots f_6$. Let $\alpha = \frac{2}{3}$ (see Figure 2). In this case, α does

FIGURE 2. CVaR Example 1: computation of CVaR when α does not spilt the atom.



not split any probability atom. Then $\mathrm{VaR}_\alpha(X) < \mathrm{CVaR}^-_\alpha(X) < \mathrm{CVaR}_\alpha(X) = \mathrm{CVaR}^+_\alpha(X)$, $\lambda_\alpha(X) = (F_X(\mathrm{VaR}_\alpha(X)) - \alpha)/(1-\alpha) = 0$ and $\mathrm{CVaR}_\alpha(X) = \mathrm{CVaR}^+_\alpha(X) = \frac{1}{2}f_5 + \frac{1}{2}f_6$, where f_5 and f_6 are losses number five and six, respectively. Now, let $\alpha = \frac{7}{12}$ (see Figure 3). In this case, α does split the $\mathrm{VaR}_\alpha(X)$ atom, $\lambda_\alpha(X) = (F_X(\mathrm{VaR}_\alpha(X)) - \alpha)/(1-\alpha) > 0$, and $\mathrm{CVaR}_\alpha(X)$ is given by

$$\mathrm{CVaR}_\alpha(X) = \tfrac{1}{5}\mathrm{VaR}_\alpha(X) + \tfrac{4}{5}\mathrm{CVaR}^+_\alpha(X) = \tfrac{1}{5}f_4 + \tfrac{2}{5}f_5 + \tfrac{2}{5}f_6.$$

In the last case, we consider four equally likely scenarios and $\alpha = \frac{7}{8}$ splits the last atom (see Figure 4). Now $\mathrm{VaR}_\alpha(X) = \mathrm{CVaR}^-_\alpha(X) = \mathrm{CVaR}_\alpha(X)$, upper CVaR, $\mathrm{CVaR}^+_\alpha(X)$ is not defined, $\lambda_\alpha(X) = (F_X(\mathrm{VaR}_\alpha(X)) - \alpha)/(1-\alpha) > 0$, and $\mathrm{CVaR}_\alpha(X) = \mathrm{VaR}(X) = f_4$. The PSG package defines the CVaR function for discrete distributions equivalently to (4) through the lower CVaR and upper CVaR. Suppose that $\mathrm{VaR}_\alpha(X)$ atom having total probability $\alpha^+ - \alpha^-$ is split by the confidence level α in two pieces with probabilities $\alpha^+ - \alpha$ and $\alpha - \alpha^-$. Then,

$$\mathrm{CVaR}_\alpha(X) = \frac{\alpha^+ - \alpha}{\alpha^+ - \alpha^-}\,\frac{1-\alpha^-}{1-\alpha}\mathrm{CVaR}^-_\alpha(X) + \frac{\alpha - \alpha^-}{\alpha^+ - \alpha^-}\,\frac{1-\alpha^+}{1-\alpha}\mathrm{CVaR}^+_\alpha(X), \tag{7}$$

where

$$\mathrm{CVaR}^-_\alpha(X) = E[X \mid X \ge \mathrm{VaR}_\alpha(X)], \qquad \mathrm{CVaR}^+_\alpha(X) = E[X \mid X > \mathrm{VaR}_\alpha(X)]. \tag{8}$$

Pflug [15] followed a different approach and suggested to define CVaR via an optimization problem, which he borrowed from Rockafellar and Uryasev [19]:

$$\mathrm{CVaR}_\alpha(X) = \min_C \left\{ C + \frac{1}{1-\alpha} E[X - C]^+ \right\}, \quad \text{where } [t]^+ = \max\{0, t\}. \tag{9}$$

One more equivalent representation of CVaR was given by Acerbi [1], who showed that CVaR is equal to "expected shortfall" defined by

$$\mathrm{CVaR}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\beta(X)\, d\beta.$$

FIGURE 3. CVaR Example 2: computation of CVaR when α splits the atom.

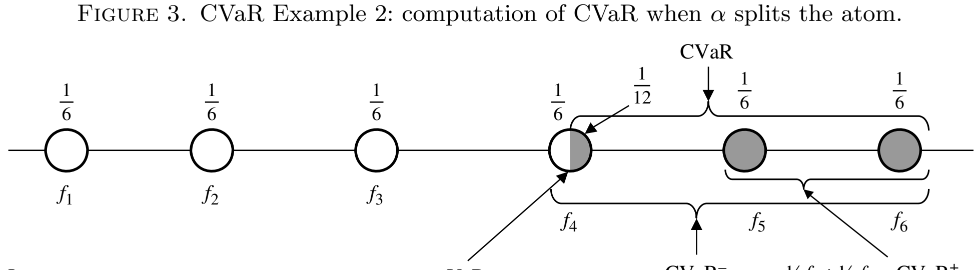
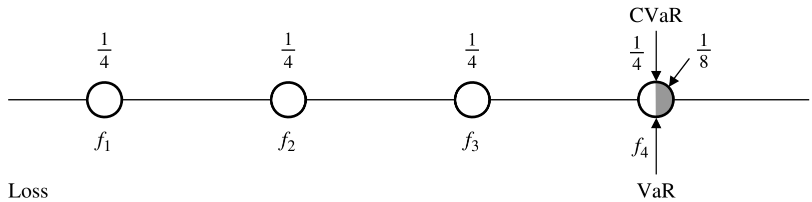


FIGURE 4. CVaR Example 3: computation of CVaR when α splits the last atom.



For normally distributed random variables, a CVaR deviation is proportional to the standard deviation. If $X \sim N(\mu, \sigma^2)$, then (see Rockafellar and Uryasev [19])

$$\mathrm{CVaR}_\alpha(X) = E[X \mid X \geq \mathrm{VaR}_\alpha(X)] = \mu + k_1(\alpha)\sigma, \tag{10}$$

where

$$k_1(\alpha) = \left(\sqrt{2\pi}\exp(\mathrm{erf}^{-1}(2\alpha - 1))^2(1-\alpha)\right)^{-1}$$

and $\mathrm{erf}(z) = (2/\sqrt{\pi}\,)\int_0^z e^{-t^2}\,dt$.

2.2. Risk Measures

Axiomatic investigation of risk measures was suggested by Artzner et al. [3]. Rockafellar [17] defined a functional $\mathcal{R}$: $\mathcal{L}^2 \to]-\infty, \infty]$ as a *coherent risk measure in the extended sense* if

R1: $\mathcal{R}(C) = C$ for all constant C;
R2: $\mathcal{R}((1-\lambda)X + \lambda X') \leq (1-\lambda)\mathcal{R}(X) + \lambda R(X')$ for $\lambda \in]0,1[$ (convexity);
R3: $\mathcal{R}(X) \leq \mathcal{R}(X')$ when $X \leq X'$ (monotonicity);
R4: $\mathcal{R}(X) \leq 0$ when $\|X^k - X\|_2 \to 0$ with $\mathcal{R}(X^k) \leq 0$ (closedness).

A functional $\mathcal{R}$: $\mathcal{L}^2 \to]-\infty, \infty]$ is called a *coherent risk measure in the basic sense* if it satisfies axioms R1, R2, R3, R4, and additionally the axiom

R5: $\mathcal{R}(\lambda X) = \lambda\mathcal{R}(X)$ for $\lambda > 0$ (positive homogeneity).

A functional $\mathcal{R}$: $\mathcal{L}^2 \to]-\infty, \infty]$ is called an *averse risk measure in the extended sense* if it satisfies axioms R1, R2, R4, and

R6: $\mathcal{R}(X) > EX$ for all nonconstant X (aversity).

Aversity has the interpretation that the risk of loss in a nonconstant random variable X cannot be acceptable; i.e., $\mathcal{R}(X) < 0$, unless $EX < 0$.

A functional $\mathcal{R}$: $\mathcal{L}^2 \to]-\infty, \infty]$ is called an *averse risk measure in the basic sense* if it satisfies R1, R2, R4, R6, and also R5.

Examples of coherent measures of risk are $\mathcal{R}(X) = \mu X = E[X]$ or $\mathcal{R}(X) = \sup X$. However, $R(X) = \mu(X) + \lambda\sigma(X)$ for some $\lambda > 0$ is not a coherent measure of risk because it does not satisfy the monotonicity axiom R3.

$R(X) = \mathrm{VaR}_\alpha(X)$ is not a coherent nor an averse risk measure. The problem lies in the convexity axiom $R2$, which is equivalent to the combination of positive homogeneity and subadditivity, this last defined as $\mathcal{R}(X + X') \leq \mathcal{R}(X) + \mathcal{R}(X')$. Although positive homogeneity is obeyed, the subadditivity is violated. It has been proved, for example, in Acerbi and Tasche [2], Pflug [15], and Rockafellar and Uryasev [20], that for any probability level $\alpha \in]0,1[$, $\mathcal{R}(X) = \mathrm{CVaR}_\alpha(X)$ is a coherent measure of risk in the basic sense. $\mathrm{CVaR}_\alpha(X)$ is also an averse measure of risk for $\alpha \in]0,1]$. An averse measure of risk might not be coherent; a coherent measure might not be averse.

2.3. Deviation Measures

In this section, we refer to Rockafellar [17] and Rockafellar et al. [24]. A functional $\mathcal{D}$: $\mathcal{L}^2 \to [0,\infty]$ is called a *a deviation measure in the extended sense* if it satisfies

D1: $\mathcal{D}(C)=0$ for constant C, but $\mathcal{D}(X)>0$ for nonconstant X;
D2: $\mathcal{D}((1-\lambda)X+\lambda X')\le(1-\lambda)\mathcal{D}(X)+\lambda D(X')$ for $\lambda\in]0,1[$ (convexity);
D3: $\mathcal{D}(X)\le d$ when $\|X^k-X\|_2\to 0$ with $\mathcal{D}(X^k)\le d$ (closedness).

A functional is called a *deviation measure in the basic sense* when it satisfies axioms D1, D2, D3, and, furthermore,

D4: $\mathcal{D}(\lambda X)=\lambda\mathcal{D}(X)$ for $\lambda>0$ (positive homogeneity).

A deviation measure in the extended or basic sense is called a *coherent* measure in the extended or basic sense if it additionally satisfies

D5: $\mathcal{D}(X)\le \sup X - E[X]$ for all X (upper range boundedness).

An immediate example of a deviation measure in the basic sense is the standard deviation

$$\sigma(X)=(E[X-EX]^2)^{1/2},$$

which satisfies axioms D1, D2, D3, D4, but not D5. In other words, standard deviation is not a coherent deviation measure. Here are more examples of deviation measures in the basic sense:

Standard semideviations:

$$\sigma_+(X)=(E[\max\{X-EX,0\}]^2)^{1/2},$$
$$\sigma_-(X)=(E[\max\{EX-X,0\}]^2)^{1/2};$$

Mean Absolute Deviation:

$$\mathrm{MAD}(X)=E[|X-EX|].$$

Moreover, we define the *α-value-at-risk deviation measure* and the *α-conditional value-at-risk deviation measure* as

$$\alpha-\mathrm{VaR}_\alpha^\Delta(X)=\mathrm{VaR}_\alpha(X-EX) \tag{11}$$

and

$$\alpha-\mathrm{CVaR}_\alpha^\Delta(X)=\mathrm{CVaR}_\alpha(X-EX), \tag{12}$$

respectively. The VaR deviation measure $\mathrm{VaR}_\alpha^\Delta(X)$ is not a deviation measure in the general or basic sense because the convexity axiom D2 is not satisfied. The CVaR deviation measure $\mathrm{CVaR}_\alpha^\Delta(X)$ is a coherent deviation measure in the basic sense.

2.4. Risk Measures vs. Deviation Measures

Rockafellar et al. (originally in [24], then in [17]), obtained the following result:

Theorem 1. *A one-to-one correspondence between deviation measures $\mathcal{D}$ in the extended sense and averse risk measures $\mathcal{R}$ in the extended sense is expressed by the relations*

$$\mathcal{R}(X)=\mathcal{D}(X)+EX,$$
$$\mathcal{D}(X)=\mathcal{R}(X-EX);$$

additionally,

$$\mathcal{R}\text{ is coherent }\leftrightarrow\mathcal{D}\text{ is coherent.}$$

Moreover, the positive homogeneity is preserved:

$$\mathcal{R} \text{ is positively homogeneous } \leftrightarrow \mathcal{D} \text{ is positively homogeneous.}$$

In other words, for an averse risk measures $\mathcal{R}$ in the basic sense and a deviation measures $\mathcal{D}$ in the basic sense the one-to-one correspondence is valid, and, additionally, coherent $\mathcal{R}$ $\leftrightarrow$ coherent $\mathcal{D}$.

With this theorem we obtain that for the standard deviation, $\sigma(X)$, which is a deviation measure in the basic sense, the counterpart is the standard risk $EX + \sigma(X)$, which is a risk averse measure in the basic sense. For CVaR deviation, $\text{CVaR}^{\Delta}_{\alpha}(X)$, which is a coherent deviation measure in the basic sense, the counterpart is CVaR risk, $\text{CVaR}_{\alpha}(X)$, which is a risk-averse coherent measure in the basic sense.

Another coherent deviation measure in the basic sense is the so-called Mixed Deviation CVaR, which we think is the most promising for risk management purposes. Mixed Deviation CVaR is defined as

$$\text{Mixed-CVaR}^{\Delta}_{\alpha}(X) = \sum_{k=1}^{K} \lambda_k \text{CVaR}^{\Delta}_{\alpha_k}(X)$$

for $\lambda_k \geq 0$, $\sum_{k=1}^{K} \lambda_k = 1$, and α_k in $]0,1[$. The counterpart to the Mixed Deviation CVaR is the Mixed CVaR, which is the coherent averse risk measure in the basic sense, defined by

$$\text{Mixed-CVaR}_{\alpha}(X) = \sum_{k=1}^{K} \lambda_k \text{CVaR}_{\alpha_k}(X).$$

3. VaR and CVaR in Optimization and Statistics

3.1. Equivalence of Chance and VaR Constraints

For this section, we refer to Rockafellar [18]. Several engineering applications deal with probabilistic constraints such as the reliability of a system or the system's ability to meet demand; in portfolio management, often it is required that portfolio loss at a certain future time is, with high reliability, at most equal to a certain value. In these cases an optimization model can be set up so that constraints are required to be satisfied with some probability level rather than almost surely. Let $x \in \Re^n$ and let $\omega \in \Omega$ be a random event ranging over the set Ω of all random events. For a given x, we may require that most of the time some random functions $f_i(x,\omega)$, $i = 1,\dots,m$, satisfy the inequalities $f_i(x,\omega) \leq 0$, $i = 1,\dots,m$; that is, we may want that

$$\text{Prob}\{f_i(x,\omega) \leq 0\} \geq p_i \quad \text{for } i = 1,\dots,m, \tag{13}$$

where $0 \leq p_i \leq 1$. Requiring this probability to be equal to 1 is the same as requiring that $f_i(x,\omega) \leq 0$ almost surely. In most applications, this approach can lead to modelling and technical problems. In modelling, there is little guidance on what level of p_i to set; moreover, one has to deal with the issue of constraint interactions and decide whether, for instance, it is better to require $\text{Prob}\{f_1(x,\omega) \leq 0\} \geq p_1 = 0.99$ and $\text{Prob}\{f_2(x,\omega) \leq 0\} \geq p_2 = 0.95$ or to work with a joint condition like $\text{Prob}\{f_i(x,\omega) \leq 0\} \geq p$. Dealing numerically with the functions $F_i(x) = \text{Prob}\{f_i(x,\omega) \leq 0\}$ leads to the task of finding the relevant properties of F_i; a common difficulty is that the convexity of $f_i(x,\omega)$ with respect to x may not carry over to the convexity of $F_i(x)$ with respect to x.

Chance constraints and percentiles of a distribution are closely related. Let $\text{VaR}_{\alpha}(x)$ be the VaR_{α} of a loss function $f(x,\omega)$; that is,

$$\text{VaR}_{\alpha}(x) = \min\{\epsilon\colon \text{Prob}\{f(x,\omega) \leq \epsilon\} \geq \alpha\}. \tag{14}$$

Then, the following constraints are equivalent:

$$\text{Prob}\{f(x,\omega) \le \epsilon\} \ge \alpha \leftrightarrow \text{Prob}\{f(x,\omega) > \epsilon\} \le 1-\alpha \leftrightarrow \text{VaR}_\alpha(x) \le \epsilon. \tag{15}$$

Generally, $\text{VaR}_\alpha(x)$ is nonconvex with respect to x; therefore, $\text{VaR}_\alpha(x) \le \epsilon$ and $\text{Prob}\{f(x,\omega) \le \epsilon\} \ge \alpha$ may be nonconvex constraints.

3.2. CVaR Optimization

For this section, we refer to Rockafellar and Uryasev [19] and Uryasev [29]. Nowadays, VaR has achieved the high status of being written into industry regulations (for instance, in regulations for finance companies). It is difficult to optimize VaR numerically when losses are not normally distributed. Only recently VaR optimization was included in commercial packages such as PSG. As a tool in optimization modeling, CVaR has superior properties in many respects. CVaR optimization is consistent with VaR optimization and yields the same results for normal or elliptical distributions (see definition of elliptical distribution in Embrechts et al. [6]); for models with such distributions, working with VaR, CVaR, or minimum variance (Markowitz [11]) is equivalent (see Rockafellar and Uryasev [19]). Most importantly, CVaR can be expressed by a minimization formula suggested by Rockafellar and Uryasev [19]. This formula can be incorporated into the optimization problem with respect to decision variables $x \in X \in \Re^n$, which are designed to minimize risk or shape it within bounds. Significant shortcuts are thereby achieved while preserving the crucial problem features like convexity. Let us consider that a random loss function $f(x,y)$ depends upon the decision vector x and a random vector y of risk factors. For instance, $f(x,y) = -(y_1x_1 + y_2x_2)$ is the negative return of a portfolio involving two instruments. Here, x_1, x_2 are positions and y_1, y_2 are rates of returns of two instruments in the portfolio. The main idea in Rockafellar and Uryasev [19] is to define a function that can be used instead of CVaR:

$$F_\alpha(x,\zeta) = \zeta + \frac{1}{1-\alpha} E\{[f(x,y)-\zeta]^+\}. \tag{16}$$

The authors proved that

1. $F_\alpha(x,\zeta)$ is convex with respect to (w.r.t.) α;
2. $\text{VaR}_\alpha(x)$ is a minimum point of function $F_\alpha(x,\zeta)$ w.r.t. ζ;
3. minimizing $F_\alpha(x,\zeta)$ w.r.t. ζ gives $\text{CVaR}_\alpha(x)$:

$$\text{CVaR}_\alpha(x) = \min_\alpha F_\alpha(x,\zeta). \tag{17}$$

In optimization problems, CVaR can enter into the objective or constraints or both. A big advantage of CVaR over VaR in that context is the preservation of convexity; i.e., if $f(x,y)$ is convex in x, then $\text{CVaR}_\alpha(x)$ is convex in x. Moreover, if $f(x,y)$ is convex in x, then the function $F_\alpha(x,\zeta)$ is convex in both x and ζ. This convexity is very valuable because minimizing $F_\alpha(x,\zeta)$ over $(x,\zeta) \in X \times \Re$ results in minimizing $\text{CVaR}_\alpha(x)$:

$$\min_{x\in X} \text{CVaR}_\alpha(x) = \min_{(x,\,\zeta)\in X\times\Re} F_\alpha(x,\zeta). \tag{18}$$

In addition, if (x^*,ζ^*) minimizes F_α over $X \times \Re$, then not only does x^* minimize $\text{CVaR}_\alpha(x)$ over X but also

$$\text{CVaR}_\alpha(x^*) = F_\alpha(x^*,\zeta^*).$$

In risk management, CVaR can be utilized to "shape" the risk in an optimization model. For that purpose, several confidence levels can be specified. Rockafellar and Uryasev [19] showed that for any selection of confidence levels α_i and loss tolerances ω_i, $i=1,\ldots,l$, the problem

$$\begin{aligned}&\min_{x\in X}\ g(x)\\ &\text{s.t. } \operatorname{CVaR}_{\alpha_i}(x)\le\omega_i,\quad i=1,\ldots,l\end{aligned}\tag{19}$$

is equivalent to the problem

$$\begin{aligned}&\min_{x,\zeta_1,\ldots,\zeta_l,\in X\times\Re\times\cdots\Re}\ g(x)\\ &\text{s.t. } F_{\alpha_i}(x,\zeta_i)\le\omega_i,\quad i=1,\ldots,l.\end{aligned}\tag{20}$$

When X and g are convex and $f(x,y)$ is convex in x, the optimization problems (18) and (19) are ones of convex programming and, thus, especially favorable for computation. When Y is a discrete probability space with elements y_k, $k=1,\ldots,N$ having probabilities p_k, $k=1,\ldots,N$, we have

$$F_{\alpha_i}(x,\zeta_i)=\zeta_i+\frac{1}{1-\alpha_i}\sum_{k=1}^{N}p_k[f(x,y_k)-\zeta_i]^+.\tag{21}$$

The constraint $F_\alpha(x,\zeta)\le\omega$ can be replaced by a system of inequalities by introducing additional variables η_k:

$$\begin{gathered}\eta_k\ge 0,\quad f(x,y_k)-\zeta-\eta_k\le 0,\quad k=1,\ldots,N,\\ \zeta+\frac{1}{1-\alpha}\sum_{k=1}^{N}p_k\eta_k\le\omega.\end{gathered}\tag{22}$$

The minimization problem in (19) can be converted into the minimization of $g(x)$ with the constraints $F_{\alpha_i}(x,\zeta_i)\le\omega_i$ being replaced as presented in (22). When f is linear in x, constraints (22) are linear.

3.3. Generalized Regression Problem

In linear regression, a random variable Y is approximated in terms of random variables $X_1,X_2,\ldots,X_n$ by an expression $c_0+c_1X_1+\cdots+c_nX_n$. The coefficients are chosen by minimizing the mean square error:

$$\min_{c_0,c_1,\ldots,c_n} E(Y-[c_0+c_1X_1+\cdots+c_nX_n])^2.\tag{23}$$

Mean square error minimization is equivalent to minimizing the standard deviation with the unbiasedness constraint (see, Rockafellar et al. [21, 26]):

$$\begin{aligned}&\min\ \sigma(Y-[c_0+c_1X_1+\cdots+c_nX_n])\\ &\text{s.t. } E[c_0+c_1X_1+\cdots+c_nX_n]=EY.\end{aligned}\tag{24}$$

Rockafellar et al. [21, 26] considered a general axiomatic setting for error measures and corresponding deviation measures. They defined an error measure as a functional $\mathcal{E}\colon\mathcal{L}^2(\Omega)\to[0,\infty]$ satisfying the following axioms:

E1: $\mathcal{E}(0)=0$, $\mathcal{E}(X)>0$ for $X\ne 0$, $\mathcal{E}(C)<\infty$ for constant C;
E2: $\mathcal{E}(\lambda X)=\lambda\mathcal{E}(X)$ for $\lambda>0$ (positive homogeneity);
E3: $\mathcal{E}(X+X')\le\mathcal{E}(X)+\mathcal{E}(X')$ for all X and X' (subadditivity);
E4: $\{X\in\mathcal{L}^2(\Omega)\mid\mathcal{E}(X)\le c\}$ is closed for all $c<\infty$ (lower semicontinuity).

For an error measure $\mathcal{E}$, the projected deviation measure $\mathcal{D}$ is defined by the equation $\mathcal{D}(X)=\min_C\mathcal{E}(X-C)$, and the statistic $S(X)$ is defined by $\mathcal{S}(X)=\arg\min_C\mathcal{E}(X-C)$.

Their main finding is that the general regression problem

$$\min_{c_0, c_1, \ldots, c_n} \mathcal{E}(Y - [c_0 + c_1 X_1 + \cdots + c_n X_n]) \tag{25}$$

is equivalent to

$$\begin{aligned} \min_{c_1, \ldots, c_n} \quad & \mathcal{D}(Y - [c_1 X_1 + \cdots + c_n X_n]) \\ \text{s.t.} \quad & c_0 \in \mathcal{S}(Y - [c_1 X_1 + \cdots + c_n X_n]). \end{aligned}$$

The equivalence of optimization problems (23) and (24) is a special case of this theorem. This leads to the identification of a link between statistical work on percentile regression (see Koenker and Basset [7]) and CVaR deviation measure: minimization of the Koenker and Bassett [7] error measure is equivalent to minimization of CVaR deviation. Rockafellar et al. [26] show that when the error measure is the Koenker and Basset [7] function, $\mathcal{E}^{\alpha}_{KB}(X) = E[\max\{0, X\} + (\alpha^{-1} - 1)\max\{0, -X\}]$, the projected measure of deviation is $\mathcal{D}(X) = \mathrm{CVaR}^{\Delta}_{\alpha}(X) = \mathrm{CVaR}_{\alpha}(X - EX)$, with the corresponding averse measure of risk and associated statistic given by

$$\begin{aligned} \mathcal{R}(X) &= \mathrm{CVaR}_{\alpha}(X), \\ \mathcal{S}(X) &= \mathrm{VaR}_{\alpha}(X). \end{aligned}$$

Then,

$$\begin{aligned} &\min_{C \in \Re} (E[X - C]_+ + (\alpha^{-1} - 1) E[X - C]_-) = \mathrm{CVaR}^{\Delta}_{\alpha}(X), \\ &\arg\min_{C \in \Re} (E[X - C]_+ + (\alpha^{-1} - 1) E[X - C]_-) = \mathrm{VaR}_{\alpha}(X). \end{aligned}$$

A similar result is available for the "mixed Koenker and Bassett error measure" and the corresponding mixed deviation CVaR; see Rockafellar et al. [26].

3.4. Stability of Estimation

The need to estimate VaR and CVaR arises typically when we are interested in estimating tails of distributions. It is of interest, in this respect, to compare the stability of estimates of VaR and CVaR based on a finite number of observations. The common flaw in such comparisons is that some confidence level is assumed and estimations of $\mathrm{VaR}_{\alpha}(X)$ and $\mathrm{CVaR}_{\alpha}(X)$ are compared with the common value of confidence level α, usually, 90%, 95%, and 99%. The problem with such comparisons is that VaR and CVaR with the same confidence level measure "different parts" of the distribution. In reality, for a specific distribution, the confidence levels α_1 and α_2 for comparison of VaR and CVaR should be found from the equation

$$\mathrm{VaR}_{\alpha_1}(X) = \mathrm{CVaR}_{\alpha_2}(X). \tag{26}$$

For instance, in the credit risk example in Serraino et al. [27], we find that CVaR with confidence level $\alpha = 0.95$ is equal to VaR with confidence level $\alpha = 0.99$. The paper by Yamai and Yoshiba [30] can be considered a "good" example of a "flawed" comparison of VaR and CVaR estimates. Yamai and Yoshiba [30] examine VaR and CVaR estimations for the parametrical family of stable distributions. The authors ran 10,000 simulations of size 1,000 and compared standard deviations of VaR and CVaR estimates normalized by their mean values. Their main findings are as follows. VaR estimators are generally more stable than CVaR estimators with the same confidence level. The difference is most prominent for fat-tailed distributions and is negligible when the distributions are close to normal. A larger sample size increases the accuracy of CVaR estimation. We provide here two illustrations of Yamai and Yoshiba's [30] results of these estimators' performances.

In the first case, the distribution of an equity option portfolio is modelled. The portfolio consists of call options based on three stocks with joint log-normal distribution. VaR and CVaR are estimated at the 95% confidence level on 10,000 sets of Monte Carlo simulations with a sample size of 1,000. The resulting loss distribution for the portfolio of at-the-money options is quite close to normal; estimation errors of VaR and CVaR are similar. The resulting loss distribution for the portfolio of deep out-of-the-money options is fat tailed; in this case, the CVaR estimator performed significantly worse than the VaR estimator.

In the second case, estimators are compared on the distribution of a loan portfolio, consisting of 1,000 loans with homogeneous default rates of 1% through 0.1%. Individual loan amounts obey the exponential distribution with an average of $100 million. Correlation coefficients between default events are homogeneous at levels 0.00, 0.03, and 0.05. Results show that estimation errors of CVaR and VaR estimators are similar when the default rate is higher; for lower default rates, the CVaR estimator gives higher errors. Also, the higher the correlation between default events, the more the loan portfolio distribution becomes fat tailed, and the higher is the error of CVaR estimator relative to VaR estimator.

As we pointed out, these numerical experiments compare VaR and CVaR with the same confidence level, and some other research needs to be done to compare stability of estimators for the same part of the distribution.

3.5. Decomposition According to Contributions of Risk Factors

This subsection discusses decomposition of VaR and CVaR risk according to risk factor contributions; see, for instance, Tasche [28] and Yamai and Yoshiba [30]. Consider a portfolio loss X, which can be decomposed as

$$X = \sum_{i=1}^{n} X_i z_i,$$

where X_i are losses of individual risk factors and z_i are sensitivities to the risk factors, $i = 1, \ldots, n$. The following decompositions of VaR and CVaR hold for continuous ditributions:

$$\mathrm{VaR}_\alpha(X) = \sum_{i=1}^{n} \frac{\partial \mathrm{VaR}_\alpha(X)}{\partial z_i} z_i = E[X_i \mid X = \mathrm{VaR}_\alpha(X)] z_i, \tag{27}$$

$$\mathrm{CVaR}_\alpha(X) = \sum_{i=1}^{n} \frac{\partial \mathrm{CVaR}_\alpha(X)}{\partial z_i} z_i = E[X_i \mid X \geq \mathrm{VaR}_\alpha(X)] z_i. \tag{28}$$

When a distribution is modelled by scenarios, it is much easier to estimate quantities $E[X_i \mid X \geq \mathrm{VaR}_\alpha(X)]$ in the CVaR decomposition than quantities $E[X_i \mid X = \mathrm{VaR}_\alpha(X)]$ in the VaR decomposition. Estimators of $\partial \mathrm{VaR}_\alpha(X)/\partial z_i$ are less stable than estimators of $\partial \mathrm{CVaR}_\alpha(X)/\partial z_i$.

3.6. Generalized Master-Fund Theorem and CAPM

The one-fund theorem is a fundamental result of the classical portfolio theory. It establishes that any mean-variance-efficient portfolio can be constructed as a combination of a single master-fund portfolio and a risk-free instrument. Rockafellar et al. [22] investigated in detail the consequences of substituting the standard deviation in the setting of classical theory with general deviation measures, in cases where rates of return may have discrete distributions, mixed discrete-continuous distributions (which can arise from derivatives, such as options) or continuous distributions. Their main result is that the one-fund theorem holds regardless of the particular choice of the deviation measure. The optimal risky portfolio needs not always be unique and it might not always be expressible by a master fund as traditionally conceived, even when only the standard deviation is involved. The authors introduce the concepts of a basic fund as the portfolio providing the minimum portfolio deviation δ for a gain of exactly 1 over the risk-free rate, with the corresponding basic deviation, defined as the minimum deviation amount. Then they establish the generalized one-fund theorem, according to which for any level of risk premium Δ over the risk-free rate, the solution of the portfolio optimization problem is given by investing an amount Δ in the basic fund and an amount $1-\Delta$ in the risk-free rate. In particular, the price of the basic fund can be positive, negative, or equal to zero, leading respectively to a long position, short position, or no investment in the basic fund. Rockafellar et al., first in Rockafellar et al. [23] and further in Rockafellar et al. [25], developed these concepts to obtain the generalized capital asset pricing (CAPM) relations. They find that the generalized CAPM equilibrium holds under the following assumptions: there are several groups of investors, each with utility function $U_j(ER_j, \mathcal{D}_j(R_j))$ based on deviation measure $\mathcal{D}_j$; the utility functions depend on mean and deviation and they are concave w.r.t. mean and deviation, increasing w.r.t. mean, and decreasing w.r.t. deviation; investors maximize their utility functions subject to the budget constraint. The main finding is that an equilibrium exists w.r.t. $\mathcal{D}_i$ in which each group of investors has its own master fund and investors invest in the risk-free asset and their own master funds. A generalized CAPM holds and takes the form

$$\bar{r}_{ij} - r_0 = \beta_{ij}(\bar{r}_{jM} - r_0),$$
$$\beta_{ij} = \frac{\mathrm{cov}(G_j, r_{ij})}{\mathcal{D}(-r_{jM})},$$

where

$\bar{r}_{ij}$ is expected return of asset i in group j,
r_0 is risk-free rate,
$\bar{r}_{jM}$ is expected return of market fund for investor group j,
G_j is the risk identifier for the market fund j.

From this general statement we can obtain that in classical framework when all investors are interested in standard deviation, β_i is defined as

$$\beta_i = \frac{\mathrm{cov}(r_i, r_M)}{\sigma^2(r_M)}.$$

Similarly, when all investors are interested in semideviation $\mathcal{D}(X) = \sigma_-(X)$, then

$$\beta_i = \frac{\mathrm{cov}(r_i, r_M)}{\sigma_-^2(-r_M)},$$

and when $\mathcal{D}(X) = \mathrm{CVaR}_\alpha^\Delta(X)$ with continuously distributed random values, then

$$\beta_i = \frac{E[r_i - \bar{r}_i \mid -r_M \geq \mathrm{VaR}_\alpha(-r_M)]}{\mathrm{CVaR}_\alpha^\Delta(-r_M)}.$$

It is interesting to observe in the last case that "beta" picks up only events when market is in $\alpha \cdot 100\%$ of its highest losses; i.e., $-r_M \geq \mathrm{VaR}_\alpha(-r_M)$.

4. Comparative Analysis of VaR and CVaR

4.1. VaR Pros and Cons

4.1.1. Pros. VaR is a relatively simple risk management notion. Intuition behind α-percentile of a distributions is easily understood and VaR has a clear interpretation: how much you may lose with certain confidence level. VaR is a single number measuring risk, defined by some specified confidence level, e.g., $\alpha = 0.95$. Two distributions can be ranked by comparing their VaRs for the same confidence level. Specifying VaR for all confidence levels completely defines the distribution. In this sense, VaR is superior to the standard deviation. Unlike the standard deviation, VaR focuses on a specific part of the distribution specified by the confidence level. This is what is often needed, which made VaR popular in risk management, including finance, nuclear, airspace, material science, and various military applications.

One of important properties of VaR is stability of estimation procedures. Because VaR disregards the tail, it is not affected by very high tail losses, which are usually difficult to measure. VaR is estimated with parametric models; for instance, covariance VaR based on the normal distribution assumption is very well known in finance, with simulation models such as historical or Monte Carlo or by using approximations based on second-order Taylor expansion.

4.1.2. Cons. VaR does not account for properties of the distribution beyond the confidence level. This implies that $\mathrm{VaR}_\alpha(X)$ may increase dramatically with a small increase in α. To adequately estimate risk in the tail, one may need to calculate several VaRs with different confidence levels. The fact that VaR disregards the tail of the distribution may lead to unintentional bearing of high risks. In a financial setting, for instance, let us consider the strategy of "naked" shorting deep out-of-the-money options. Most of the time, this will result in receiving an option premium without any loss at expiration. However, there is a chance of a big adverse market movement leading to an extremely high loss. VaR cannot capture this risk.

Risk control using VaR may lead to undesirable results for skewed distributions. Later on we will demonstrate this phenomenon with the case study comparing risk profiles of VaR and CVaR optimization. In this case, the VaR optimal portfolio has about 20% longer tail than the CVaR optimal portfolio, as measured by the max loss of those portfolios.

VaR is a nonconvex and discontinuous function for discrete distributions. For instance, in financial setting, VaR is a nonconvex and discontinuous function w.r.t. portfolio positions when returns have discrete distributions. This makes VaR optimization a challenging computational problem. Nowadays there are codes, such as PSG, that can work with VaR functions very efficiently. PSG can optimize portfolios with a VaR performance function and also shape distributions of the portfolio with multiple VaR constraints. For instance, in portfolio optimization it is possible to maximize expected return with several VaR constraints at different confidence levels.

4.2. CVaR Pros and Cons

4.2.1. Pros. CVaR has a clear engineering interpretation. It measures outcomes that hurt the most. For example, if L is a loss then the constraint $\mathrm{CVaR}_\alpha(L) \le \bar{L}$ ensures that the average of $(1-\alpha)\%$ highest losses does not exceed $\bar{L}$.

Defining $\mathrm{CVaR}_\alpha(X)$ for all confidence levels α in $(0,1)$ completely specifies the distribution of X. In this sense, it is superior to standard deviation.

Conditional value at risk has several attractive mathematical properties. CVaR is a coherent risk measure. $\mathrm{CVaR}_\alpha(X)$ is continuous with respect to α. CVaR of a convex combination of random variables $\mathrm{CVaR}_\alpha(w_1X_1 + \cdots + w_nX_n)$ is a convex function with respect to $(w_1, \ldots, w_n)$. In financial setting, CVaR of a portfolio is a convex function of portfolio positions. CVaR optimization can be reduced to convex programming, in some cases to linear programming (i.e., for discrete distributions).

4.2.2. Cons. CVaR is more sensitive than VaR to estimation errors. If there is no good model for the tail of the distribution, CVaR value may be quite misleading; accuracy of CVaR estimation is heavily affected by accuracy of tail modelling. For instance, historical scenarios often do not provide enough information about tails; hence, we should assume a certain model for the tail to be calibrated on historical data. In the absence of a good tail model, one should not count on CVaR. In financial setting, equally weighted portfolios may outperform CVaR-optimal portfolios out of sample when historical data have mean reverting characteristics.

4.3. What Should You Use, VaR or CVaR?

VaR and CVaR measure different parts of the distribution. Depending on what is needed, one may be preferred over the other.

Let us illustrate this topic with financial applications of VaR and CVaR and examine the question of which measure is better for portfolio optimization. A trader may prefer VaR to CVaR, because he may like high uncontrolled risks; VaR is not as restrictive as CVaR with the same confidence level. Nothing dramatic happens to a trader in case of high losses. He will not pay losses from his pocket; if fired, he may move to some other company. A company owner will probably prefer CVaR; he has to cover large losses if they occur; hence, he "really" needs to control tail events. A board of directors of a company may prefer to provide VaR-based reports to shareholders and regulators because it is less than CVaR with the same confidence level. However, CVaR may be used internally, thus creating asymmetry of information between different parties.

VaR may be better for optimizing portfolios when good models for tails are not available. VaR disregards the hardest to measure events. CVaR may not perform well out of sample when portfolio optimization is run with poorly constructed set of scenarios. Historical data may not give right predictions of future tail events because of mean-reverting characteristics of assets. High returns typically are followed by low returns; hence, CVaR based on history may be quite misleading in risk estimation.

If a good model of tail is available, then CVaR can be accurately estimated and CVaR should be used. CVaR has superior mathematical properties and can be easily handled in optimization and statistics.

When comparing stability of estimation of VaR and CVaR, appropriate confidence levels for VaR and CVaR must be chosen, avoiding comparison of VaR and CVaR for the same level of α because they refer to different parts of the distribution.

5. Numerical Results

This section reports results of several case studies exemplifying how to use VaR and CVaR in optimization settings.

5.1. Portfolio Safeguard

We used PSG to do the case studies. We posted MATLAB files to run these case studies in a MATLAB environment on the MathWorks website (http://www.mathworks.com), in the file exchange-optimization area. PSG is designed to solve a wide range of risk management and stochastic optimization problems. PSG has built-in algorithms for evaluating and optimizing various statistical and deterministic functions. For instance, PSG includes the statistical functions Mean, Variance, Standard Deviation, VaR, CVaR, CDaR, MAD, Maximum Loss, Partial Moment, Probability of Exceeding a Threshold, and the deterministic functions Cardinality, Fixed Charge, and Buyin. For a complete list of functions, see Table A.1 in Appendix I. Required data inputs are matrices of Scenarios or Covariance Matrices on which statistical functions are evaluated. PSG uses a new design for defining optimization problems. A function is defined by just providing an underlying Matrix to a Function. Figure 5 illustrates PSG procedure for function definition (Mean and Max Loss Functions are defined: Matrix of Scenarios ⇒ Loss Vector ⇒ Functions). With PSG design all needed data, including names of variables, are taken from a Matrix of Scenarios or Covariance Matrix. PSG can calculate values and sensitivities of defined functions on decision vectors. Also, you can place functions to objective and constraint and define an optimization problem. Linear combinations of functions can be placed in the constraints. PSG includes many case

FIGURE 5. Example of evaluation of two PSG functions: Mean and Max Loss.

INPUT DATA: Matrix of scenarios			
x	y	Scenario benchmark	Scenario probability
θ_{11}	θ_{12}	θ_{01}	p_1
θ_{21}	θ_{22}	θ_{02}	p_2

Loss vector	
Loss function	Scenario probability
$\theta_{01} - (\theta_{11} * X + \theta_{21} * Y)$	p_1
$\theta_{02} - (\theta_{12} * X + \theta_{22} * Y)$	p_2

Function evaluation	
	Function value
Mean	$p_1 * [\theta_{01} - (\theta_{11} * X + \theta_{21} * Y)] + p_2 * [\theta_{02} - (\theta_{12} * X + \theta_{22} * Y)]$
Max Loss	$\max[\theta_{01} - (\theta_{11} * X + \theta_{21} * Y)], [\theta_{02} - (\theta_{12} * X + \theta_{22} * Y)]$

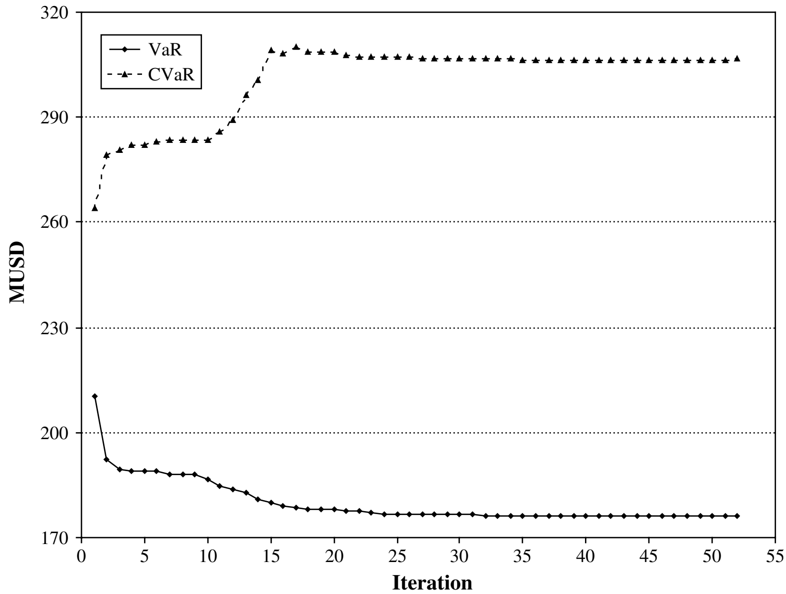
Notes. These functions are completely defined by the matrix of scenarios. Both names of variables and function values are calculated using matrix of scenarios.

studies, most of them motivated by financial applications (such as collateralized debt obligation structuring, portfolio management, portfolio replication, and selection of insurance contracts) that are helpful in learning the software. To build a new optimization problems, you can simply make a copy of the case study and change input matrices and parameters. PSG can solve problems with simultaneous constraints on various risks at different time intervals (e.g., multiple constraints on standard deviation obtained by resampling, combined with VaR and drawdown constraints), thus allowing robust decision making. It has built-in efficient algorithms for solving large-scale optimization problems (up to 1,000,000 scenarios and up to 10,000 decision variables in callable MATLAB and C++ modules). PSG offers several tools for analyzing solutions, generated by optimization or through other procedures. Among these tools are sensitivities of risk measures to changes in decision variables, and the incremental impact of decision variables on risk measures and various functions of risk measures. Analysis can reveal decision variables having the biggest impact on risk and other functions, such as expected portfolio return. For visualization of these and other characteristics, PSG provides tools for building and plotting various characteristics combining different functions, points, and variables. It is user friendly and callable from MATLAB and C/C++ environment.

5.2. Risk Control Using VaR (PSG MATLAB Environment)

Risk control using VaR may lead to paradoxical results for skewed distributions. In this case, minimization of VaR may lead to a stretch of the tail of the distribution exceeding VaR. The purpose of VaR minimization is to reduce extreme losses. However, VaR minimization may lead to an increase in the extreme losses that we try to control. This is an undesirable feature of VaR optimization. Larsen et al. [9] showed this phenomenon on a credit portfolio with the loss distribution created with a Monte Carlo simulation of 20,000 scenarios of joint credit

Figure 6. VaR minimization.



Notes. 99%-VaR and 99%-CVaR for algorithm A2 in Larsen et al. [9]. The algorithm lowered VaR at the expense of an increase in CVaR.

states. The distribution is skewed with a long fat right tail. A more detailed description of this portfolio can be found in Bucay and Rosen [4] and Mausser and Rosen [13, 14]. Larsen et al. [9] suggested two heuristic algorithms for optimization of VaR. These algorithms are based on the minimization of CVaR. The minimization of VaR leads to about 16% increase of the average loss for the worst 1% scenarios, compared with the worst 1% scenarios in CVaR minimum solution. These numerical results are consistent with the theoretical results: CVaR is a coherent, whereas VaR is not a coherent, measure of risk. Figure 6 reproduces results from Larsen et al. [9] showing how iteratively VaR is decreasing and CVaR is increasing. We observed a similar performance of VaR on a small portfolio consisting of 10 bonds and modelled with 1,000 scenarios. We solved two optimization problems with PSG. In the first one, we minimized 99%-CVaR deviation of losses subject to constraints on budget and required return. In the second one, we minimized 99%-VaR deviation of losses subject to the same constraints.

$$\begin{aligned}\text{Problem 1:}\quad \min\ & \mathrm{CVaR}_\alpha^\Delta(x)\\ \text{s.t.}\ & \sum_{i=1}^n r_i x_i \ge \bar{r},\\ & \sum_{i=1}^n x_i = 1.\end{aligned} \tag{29}$$

$$\begin{aligned}\text{Problem 2:}\quad \min\ & \mathrm{VaR}_\alpha^\Delta(x)\\ \text{s.t.}\ & \sum_{i=1}^n r_i x_i \ge \bar{r},\\ & \sum_{i=1}^n x_i = 1.\end{aligned} \tag{30}$$

Table 1. Value of risk functions.

	$\min \text{CVaR}^{\Delta}_{0.99}$	$\min \text{VaR}^{\Delta}_{0.99}$	Ratio
$\text{CVaR}_{0.99}$	0.0073	0.0083	1.130
$\text{CVaR}^{\Delta}_{0.99}$	0.0363	0.0373	1.026
$\text{VaR}_{0.99}$	0.0023	0.0005	0.231
$\text{VaR}^{\Delta}_{0.99}$	0.0313	0.0295	0.944
Max loss $= \text{CVaR}_1$	0.0133	0.0148	1.116
Max loss deviation $= \text{CVaR}^{\Delta}_1$	0.0423	0.0438	1.036

Notes. Column "$\min \text{CVaR}^{0.99}_{\alpha}$" reports the value of risk functions for the portfolio obtained by minimizing 99%-CVaR deviation, column "$\min \text{VaR}^{\Delta}_{\alpha}$" reports the value of risk functions for the portfolio obtained by minimizing 99%-VaR deviation, and column "Ratio" contains the ratio of every cell of column "$\min \text{VaR}^{\Delta}_{\alpha}$" to every cell of column "$\min \text{CVaR}^{\Delta}_{\alpha}$."

where x is the vector of portfolio weights, r_i is the rate of return of asset i, $\bar{r}$ is the lower bound on estimated portfolio return. We then evaluated different risk functions at the optimal points for Problems 1 and 2. Results are shown in Table 1.

Suppose that we start with the portfolio having minimal 99%-CVaR deviation. Minimization of 99%-VaR deviation leads to 13% increase in 99%-CVaR, compared with 99%-CVaR in the optimal 99%-CVaR deviation portfolio. We found that even in a problem with a relatively small number of scenarios, if the distribution is skewed, minimization of VaR deviation may lead to a stretch of the tail compared with the CVaR optimal portfolio. This result is quite important when we look at financial risk management regulations like Basel II that are based on minimization of VaR deviation.

5.3. Linear Regression-Hedging: VaR, CVaR, Mean Absolute, and Standard Deviations (PSG MATLAB Environment)

This case study investigates performance of optimal hedging strategies based on different deviation measures measuring quality of hedging. The objective is to build a portfolio of financial instruments that mimics the benchmark portfolio. Weights in the replicating portfolio are chosen such that the deviation between the value of this replicating portfolio and the value of the benchmark is minimized. Benchmark value and replicating financial instruments values are random valves. Determining the optimal hedging strategy is a linear regression problem where the response is the benchmark portfolio valve, the predictors are the replicating financial instrument valves, and the coefficients of the predictors to be determined are the portfolio weights. Let $\hat{\theta}$ be the replicating portfolio value, θ_0 be the benchmark portfolio value, $\theta_1, \ldots, \theta_I$ be replicating instrument values, and $x_1, \ldots, x_I$ be their weights. The replicating portfolio value can be expressed as follows:

$$\hat{\theta} = x_1\theta_1 + \cdots + x_I\theta_I. \tag{31}$$

The coefficients $x_1, \ldots, x_I$ should be chosen to minimize a replication error function depending upon the residual $\theta_0 - \hat{\theta}$. The intercept in this case is absent. According to the equivalence of (25) and (26), an error minimization problem is equivalent to the minimization of the appropriate deviation; see Rockafellar et al. [21] and [26]. This case study considers hedging pipeline risk in the mortgage underwriting process. Hedging instruments are 5% Mortgage-Backed Securities (MBS) forward, 5.5% MBS, and call options on 10-year treasury note futures. Changes of values of the benchmark and the hedging instruments are driven by changes in the mortgage rate. We minimize five different deviation measures: Standard Deviation, Mean Absolute Deviation, CVaR Deviation, Two-Tailed 75%-VaR, and Two-Tailed 90%-VaR. We tested in-sample and out-of-sample performance of the hedging strategies. On our set of scenarios, we found that Two-Tailed 90%-VaR has the best out-of-sample performance, whereas the standard deviation has the worst out-of-sample performance.

We think that the out-of-sample performance of hedging strategies based on different deviation measures significantly depends on the skewness of the distribution. In this case, the distribution of residuals is quite skewed.

We use here PSG definitions of Loss and Deviation functions. The Loss Function is defined as follows:

$$\text{Loss Function} = L(x,\theta) = L(x_1,\ldots,x_I,\theta_0,\ldots,\theta_I) = \theta_0 - \sum_{i=1}^{I} \theta_i x_i. \tag{32}$$

The loss function has j scenarios, $L(x,\theta^1),\ldots,L(x,\theta^J)$, each with probability p_j, $j=1,\ldots,J$. Here are deviation measures considered in this case study:

$$\text{Mean Absolute Deviation} = \text{MAD}(L(x,\theta)), \tag{33}$$

$$\text{Standard Deviation} = \sigma(L(x,\theta)), \tag{34}$$

$$\alpha\%\text{-VaR Deviation} = \text{VaR}_\alpha^\Delta(L(x,\theta)), \tag{35}$$

$$\alpha\%\text{-CVaR Deviation} = \text{CVaR}_\alpha^\Delta(L(x,\theta)), \tag{36}$$

$$\begin{aligned}\text{Two-Tail } \alpha\%\text{-VaR Deviation} &= \text{TwoTailVaR}_\alpha^\Delta(L(x,\theta)) \\ &= \text{VaR}_\alpha(L(x,\theta)) + \text{VaR}_\alpha(-L(x,\theta)).\end{aligned} \tag{37}$$

We solved the following minimization problems:

Minimize 90%-CVaR Deviation

$$\min_x \text{CVaR}_{0.9}^\Delta(L(x,\theta)) \tag{38}$$

Minimize Mean Absolute Deviation

$$\min_x \text{MAD}(L(X,\theta)) \tag{39}$$

Minimize Standard Deviation

$$\min_x \sigma(L(x,\theta)) \tag{40}$$

Minimize Two-Tail 75%-VaR Deviation

$$\min_x \text{TwoTailVaR}_{0.75}^\Delta(L(X,\theta)) \tag{41}$$

Minimize Two-Tail 90%-VaR Deviation

$$\min_x \text{TwoTailVaR}_{0.9}^\Delta(L(X,\theta)). \tag{42}$$

The data set for the case study includes 1,000 scenarios of value changes for each hedging instrument and for the benchmark. For the out-of-sample testing, we partitioned the 1,000 scenarios into 10 groups with 100 scenarios in each group. Each time, we selected one group for the out-of-sample test and we calculated optimal hedging positions based on the remaining nine groups containing 900 scenarios. For each group of 100 scenarios, we calculated the ex ante losses (i.e., underperformances of hedging portfolio versus target) with the optimal hedging positions obtained from 900 scenarios. We repeated the procedure 10 times, once for every out-of-sample group with 100 scenarios. To estimate the out-of-sample performance, we aggregated the out-of-sample losses from the 10 runs and obtained a combined set including 1,000 out-of-sample losses. Then, we calculated five deviation measures on the out-of-sample 1,000 losses: Standard Deviation, Mean Absolute Deviation, CVaR Deviation, Two-Tail 75%-VaR Deviation, and Two-Tail 90%-VaR Deviation. In addition, we calculated three downside risk measures: 90%-CVaR, 90%-VaR, and Max Loss = 100%-CVaR on the out-of-sample losses. Tables 2 and 3 show the results.

By minimizing the Two-Tail 90%-VaR Deviation, we obtained the best values for all three considered downside risk measures (negative loss indicates gain). Minimization of CVaR deviation lead to good results, whereas minimization of standard deviation gave the worst level for three downside risk measures.

Table 2. Out-of-sample performance of various deviations on optimal hedging portfolios.

Optimal points	$\text{CVaR}^{\Delta}_{0.9}$	MAD	σ	$\text{TwoTailVaR}^{\Delta}_{0.75}$	$\text{TwoTailVaR}^{\Delta}_{0.9}$
$\text{CVaR}^{\Delta}_{0.9}$	0.690	0.815	1.961	0.275	1.122
MAD	1.137	0.714	1.641	0.379	1.880
σ	1.405	0.644	1.110	0.979	1.829
$\text{TwoTailVaR}^{\Delta}_{0.75}$	1.316	0.956	1.955	0.999	1.557
$\text{TwoTailVaR}^{\Delta}_{0.9}$	0.922	0.743	1.821	0.643	1.256

Notes. Each row reports values of five different deviation measures evaluated at optimal hedging points obtained with five hedging strategies.

5.4. Example of Equivalence of Chance and VaR Constraints (PSG MATLAB Environment)

This case study illustrates the equivalence between chance constraints and VaR constraints, as explained in §3.1. We will illustrate numerically the equivalence

$$\text{Prob}\{L(x,\theta) > \epsilon\} \le 1-\alpha \leftrightarrow \text{VaR}_\alpha(L(x,\theta)) \le \epsilon, \tag{43}$$

where

$$\text{Loss Function } L(x,\theta) = L(x_1,\ldots,x_I,\theta_1,\ldots,\theta_I) = -\sum_{i=1}^{I}\theta_i x_i. \tag{44}$$

The case study is based on a data set including 1,000 return scenarios for 10 clusters of loans. Here, $I = 10$ is the number of instruments, $\theta_1,\ldots,\theta_I$ are rates of returns of instruments, $x_1,\ldots,x_I$ are instrument weights, and $L(x,\theta)$ is a portfolio loss. We solved two portfolio optimization problems. In both cases we maximized the estimated return of the portfolio. In the first problem, we imposed a constraint on probability; in the second problem, we imposed an equivalent constraint on VaR. In particular, in the first problem we require the 95%-VaR of the optimal portfolio to be at most equal to the constant ϵ, whereas in the second problem we require the probability of losses greater than ϵ to be lower than $1-\alpha = 1-0.95 = 0.05$. We expected to obtain at optimality the same objective function value and similar optimal portfolios for the two problems. Problem formulations are as follows:

$$\begin{aligned} \text{Problem 1:}\quad \max\ & E[-L(x,\theta)] \\ \text{s.t.}\ & \text{Prob}\{L(x,\theta) > \epsilon\} \le 1-\alpha = 0.05, \\ & v_i \le x_i \le u_i, \quad i = 1,\ldots,I, \\ & \sum_{i=1}^{I} x_i = 1. \end{aligned} \tag{45}$$

Table 3. Out-of-sample performance of various downside risks on optimal hedging portfolios.

Optimal points	Max Loss	$\text{CVaR}_{0.9}$	$\text{VaR}_{0.9}$
$\text{CVaR}^{\Delta}_{0.9}$	−18.01	−18.05	−18.08
MAD	−16.49	−17.44	−17.88
σ	−13.31	−15.29	−15.60
$\text{TwoTailVaR}^{\Delta}_{0.75}$	−15.31	−16.19	−16.71
$\text{TwoTailVaR}^{\Delta}_{0.9}$	−18.02	−18.51	−18.66

Notes. Each row reports values of three different risk measures evaluated at optimal hedging points obtained with five hedging strategies (negative loss indicates gain).

Table 4. Chance vs. VaR constraints.

Optimal weights	Prob ≤ 0.05	VaR ≤ ϵ
x_1	0.051	0.051
x_2	0.055	0.055
x_3	0.071	0.071
x_4	0.053	0.053
x_5	0.079	0.079
x_6	0.289	0.289
x_7	0.020	0.020
x_8	0.300	0.300
x_9	0.063	0.063
x_{10}	0.020	0.020

Notes. Optimal portfolios obtained on the same data set when we maximized return with the chance constraint (the first column) and with the VaR constraint (the second column).

$$\begin{aligned}\text{Problem 2:}\quad \max\ & E[-L(x,\theta)]\\ \text{s.t.}\ & \mathrm{VaR}_\alpha(L(x,\theta)) \le \epsilon,\\ & v_i \le x_i \le u_i, \quad i = 1,\ldots,I,\\ & \sum_{i=1}^{I} x_i = 1.\end{aligned} \tag{46}$$

In the problems, v_i is the lower bound on position for asset i, and u_i is the upper bound on position for asset i. We also have the budget constraint: sum of weights is equal to 1. The two problems at optimality selected the same portfolios and have the same objective function value of 120.19. Table 4 shows the optimal points.

5.5. Portfolio Rebalancing Strategies: Risk vs. Deviation (PSG MATLAB Environment)

In this case study we consider a portfolio rebalancing problem. A portfolio manager allocates his wealth to different funds periodically solving an optimization problem and in each time period building the portfolio that minimizes a certain risk function, given budget constraints and bounds on each exposure. We solved the following problem:

$$\min\ R(x,\theta) - k * E[-L(x,\theta)] \tag{47}$$

$$\text{s.t.}\ \sum_{i=1}^{I} x_i = 1 \tag{48}$$

$$v_i \le (x_i) \le u_i \quad i = 1\ldots,I, \tag{49}$$

where we denote by $R(x,\theta)$ a risk function, $E[-L(x,\theta)]$ is the expected portfolio return, $\theta_1,\ldots,\theta_I$ are rates of returns of instruments, and $x_1,\ldots,x_I$ are instrument weights. The scenario data set is composed of 46 monthly return observations for seven managed funds; we solved the first optimization problem using the first 10 scenarios, then we rebalanced the portfolio monthly. We used as risk functions VaR, CVaR, VaR Deviation, CVaR Deviation, and Standard Deviation.

We then evaluated Sharpe ratio and mean value of each sequence of portfolios obtained by succesively solving the optimization problem with a given objective function. Results are reported in Tables 5 and 6 for different values of the parameter k. Both in terms of Sharpe ratio and mean portfolio value we found a good performance of VaR and VaR deviation

Table 5. Out-of-sample Sharpe ratio.

k	VaR	CVaR	VaR Deviation	CVaR Deviation	Standard Deviation
−1	1.2710	1.2609	1.2588	1.2693	1.2380
−3	1.2711	1.2667	1.2762	1.2652	1.2672
−5	1.2712	1.2666	1.2721	1.2743	1.2628

Notes. Sharpe ratio for the rebalancing strategy when different risk functions are used in the objective function with several values of parameter k.

Table 6. Out-of-sample portfolio mean return.

k	VaR	CVaR	VaR Deviation	CVaR Deviation	Standard Deviation
−1	0.2508	0.2556	0.2445	0.2567	0.2425
−3	0.2645	0.2575	0.2631	0.2598	0.2576
−5	0.2663	0.2612	0.2662	0.2617	0.2532

Notes. Portfolio mean return for the rebalancing strategy when different risk functions are used in the objective function with several values of parameter k.

minimization, whereas standard deviation minimization gives inferior results. Overall, we rebalanced the portfolios 37 times for each objective function. Results depend on the scenario data set and on the parameter k; thus, we cannot conclude that minimization of a certain risk function is always the best choice. However, we observe that in the presence of mean reversion, the tails of historical distribution are not good predictors of the tails in the future. In this case, VaR disregarding the tails may lead to a good out-of-sample portfolio performance. In fact, VaR disregards the unstable part of the distribution.

Appendix I

Table A.1. PSG functions.

Function group	Full name	PSG version
Deterministic Functions		
Linear group	Variable Function	1.1
	Linear Function	1.1
	Linear Multiple	1.1
Nonlinear group	Polynomial Absolute	1.1
	Relative Entropy	1.1
	Maximum Positive	1.2 Beta
	Maximum Negative	1.2 Beta
	CVaR Positive	1.2 Beta
	CVaR Negative	1.2 Beta
	VaR Positive	1.2 Beta
	VaR Negative	1.2 Beta
Cardinality group	Cardinality Positive	1.2 Beta
	Cardinality Negative	1.2 Beta
	Cardinality	1.2 Beta
	Buyin Positive	1.2 Beta
	Buyin Negative	1.2 Beta
	Buyin	1.2 Beta
	Fixed Charge Positive	1.2 Beta
	Fixed Charge Negative	1.2 Beta
	Fixed Charge	1.2 Beta

Table A.1. Continued.

Function group	Full name	PSG version
Risk functions		
Average group	Average Gain	1.1
	Average Loss	1.1
VaR group	VaR Deviation for Gain	1.1
	VaR Deviation for Loss	1.1
	VaR Risk for Gain	1.1
	VaR Risk for Loss	1.1
CVaR group	CVaR Ceviation for Gain	1.1
	CVaR Deviation for Loss	1.1
	CVaR Risk for Gain	1.1
	CVaR Risk for Loss	1.1
CDaR group	CDaR Deviation for Gain	1.1
	CDaR Deviation for Gain Multiple	1.1
	CDaR Deviation	1.1
	CDaR Deviation Multiple	1.1
	Drawdown Deviation Average for Gain	1.1
	Drawdown Deviation Average for Gain Multiple	1.1
	Drawdown Deviation Average	1.1
	Drawdown Deviation Average Multiple	1.1
	Drawdown Deviation Maximum for Gain	1.1
	Drawdown Deviation Maximum for Gain Multiple	1.1
	Drawdown Deviation Maximum	1.1
	Drawdown Deviation maximum Multiple	1.1
Maximum group	Maximum Deviation for Gain	1.1
	Maximum Deviation for Loss	1.1
	Maximum Risk for Gain	1.1
	Maximum Risk for Loss	1.1
Mean Abs group	Mean Absolute Deviation	1.1
	Mean Absolute Penalty	1.1
	Mean Absolute Risk for Gain	1.1
	Mean Absolute Risk for Loss	1.1
Partial moment group	Partial Moment Gain Deviation	1.1
	Partial Moment Loss Deviation	1.1
	Partial Moment Penalty for Gain	1.1
	Partial Moment Penalty for Loss	1.1
Probability group	Probability Exceeding Deviation for Gain	1.1
	Probability Exceeding Deviation for Gain Multiple	1.1
	Probability Exceeding Deviation for Loss	1.1
	Probability Exceeding Deviation for Loss Multiple	1.1
	Probability Exceeding Penalty for Gain	1.1
	Probability Exceeding Penalty for Gain Multiple	1.1
	Probability Exceeding Penalty for Loss	1.1
	Probability Exceeding Penalty for Loss Multiple	1.1
Standard group	Standard Deviation	1.1
	Standard Gain	1.1
	Standard Penalty	1.1
	Standard Risk	1.1
	Mean Square Penalty	1.1
	Variance	1.1
Utility group	Exponential Utility	1.2 Beta
	Logarithmic Utility	1.2 Beta
	Power Utility	1.2 Beta

References

[1] C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance* 26:1505–1518, 2002.

[2] C. Acerbi and D. Tasche. On the coherence of expected shortfall. *Journal of Banking and Finance* 26:1487–1503, 2002.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance* 9:203–227, 1999.

[4] N. Bucay and D. Rosen. Credit risk of an international bond portfolio: A case study. *ALGO Research Quarterly* 2(1):9–29.

[5] A. Chekhlov, S. Uryasev, and M. Zabarankin. Portfolio optimization with drawdown constraints. B. Scherer, ed. *Asset and Liability Management Tools*. Risk Books, London, 263–278, 2003.

[6] P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependence in risk management: Properties and pitfalls. M. Dempster, ed. *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, UK, 2002.

[7] R. Koenker and G. W. Bassett. Regression quantiles. *Econometrica* 46:33–50, 1978.

[8] H. Konno and H. Yamazaki. Mean absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science* 37:519–531, 1991.

[9] N. Larsen, H. Mausser, and S. Uryasev. Algorithms for optimization of value-at-risk. P. Pardalos and V. K. Tsitsiringos, eds. *Financial Engineering, e-Commerce and Supply Chain*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 129–157, 2000.

[10] A. Lo. The three P's of total risk management. *Financial Analysts Journal* 55:13–26, 1999.

[11] H. M. Markowitz. Portfolio selection. *Journal of Finance* 7(1):77–91, 1952.

[12] H. Mausser and D. Rosen. Beyond VaR: From measuring risk to managing risk. *ALGO Research Quarterly* 1(2):5–20, 1999.

[13] H. Mausser and D. Rosen. Applying scenario optimization to portfolio credit risk. *ALGO Research Quarterly* 2(2):19–33, 1999.

[14] H. Mausser and D. Rosen. Efficient risk/return frontiers for credit risk. *ALGO Research Quarterly* 2(4):35–48, 1999.

[15] G. C. Pflug. Some remarks on the value-at-risk and the conditional value-at-risk. S. P. Uryasev, ed. *Probabilistic Constrainted Optimization: Methodology and Applications*. Kluwer, Norwell, MA, 278–287, 2000.

[16] R. T. Rockafellar. *Convex Analysis*. Princeton University Press, Princeton, NJ, 1970.

[17] R. T. Rockafellar. Coherent approaches to risk in optimization under uncertainty. *INFORMS Tutorials in Operations Research*. Institute for Operations Research and the Management Sciences, Hanover, MD, 38–61, 2007.

[18] R. T. Rockafellar. *Optimization Under Uncertainty, Lectures Notes*. http://www.math.washington.edu/~rtr/uncertainty.pdf, 2001.

[19] R. T. Rockafellar and S. P. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk* 2:21–42, 2000.

[20] R. T. Rockafellar and S. P. Uryasev. Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance* 26:1443–1471, 2002.

[21] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Deviation measures in generalized linear regression. Research Report 2002-9, Department of Industrial and Systems Engineering, University of Florida, Gainesville, 2002.

[22] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Master funds in portfolio analysis with general deviation measures. *Journal of Banking and Finance* 30:743–778, 2005.

[23] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Optimality conditions in portfolio analysis with generalized deviation measures. *Mathematical Programming* 108(2–3):515–540, 2006.

[24] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Generalized deviations in risk analysis. *Finance and Stochastics* 10:51–74, 2006.

[25] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Equilibrium with investors using a diversity of deviation measures. *Journal of Banking and Finance* 31:1443–1471, 2007.

[26] R. T. Rockafellar, S. Uryasev, and M. Zabarankin. Risk tuning with generalized linear regression. *Mathematics of Operations Research*, forthcoming.

[27] G. Serraino, U. Theiler, and S. Uryasev. Risk return optimization with different risk aggregation strategies. G. N. Gregoriov, ed. *Groundbreaking resarch and developments in VaR*, forthcoming. Bloomberg Press, New York, 2009.

[28] D. Tasche. Risk contribution and performance measurement. Working paper, Technical University of Munich, Munich, Germany.

[29] S. Uryasev. Conditional value-at-risk: Optimization algorithms and applications. *Financial Engineering News* (14 February):1–5, 2000.

[30] Y. Yamai and T. Yoshiba. Comparative analysis of expected shortfall and value-at-risk: Their estimation error, decomposition and optimization. *Monetary and Economic Studies* 20(1):57–86, 2002.