Optimal Financial Portfolios

Stoyan V. Stoyanov
Chief Financial Researcher
FinAnalytica, Inc., Seattle, USA
e-mail: stoyan.stoyanov@finanalytica.com

Svetlozar T. Rachev
Department of Econometrics and Statistics,
University of Karlsruhe, D-76128 Karlsruhe, Germany and
Department of Statistics and Applied Probability,
University of California Santa Barbara
e-mail: rachev@pstat.ucsb.edu

Frank J. Fabozzi
Adjunct Professor of Finance and Becton Fellow
Yale University, School of Management
e-mail: frank.fabozzi@yale.edu

November 30, 2006

Abstract

We consider classes of reward-risk optimization problems that arise from different choices of reward and risk measures. In certain examples the generic problem reduces to linear or quadratic programming problems. We state an algorithm based on a sequence of convex feasibility problems for the general quasi-concave ratio problem. We also consider reward-risk ratios that are appropriate in particular for non-normal assets return distributions and are not quasi-concave.

Acknowledgement We would like to thank Milen Ivanov for the valuable discussions that helped us establish some of the results. We are also grateful for the helpful comments of R. T. Rockafellar and W. J. Ziemba. Rachev gratefully acknowledges research support by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara, the Deutschen Forschungsgemeinschaft and the Deutscher Akademischer Austausch Dienst.

1 Introduction

An accepted theory of asset choice under uncertainty is expected utility. Expected utility representation of agent's preferences is present if there exists a function $u \in U$ such that a random wealth X is preferred to a random wealth Y if and only if $Eu(X) \geq Eu(Y)$. The class U is characterized by investors' preferences; for instance, if they are assumed to be non-satiable and risk-averse, then U contains non-decreasing, concave functions. The problem which an investor solves is selecting the random wealth maximizing his expected utility. There are two basic approaches to the problem of portfolio selection under uncertainty stemming from this theory. One of them is the stochastic dominance approach, based on the axiomatic model of risk-averse preferences. Unfortunately, the optimization problems that arise are not easy to solve in practice. The other is reward-risk analysis. According to it, the portfolio choice is made with respect to two criteria – the expected portfolio return and portfolio risk. A portfolio is preferred to another one if it has higher expected return and lower risk. There are convenient computational recipes and geometric interpretations of the trade-off between the two criteria. A disadvantage of the reward-risk analysis is that it cannot capture the richness of the stochastic dominance approach. As a matter of fact, the relationship between the two approaches is an ongoing research topic (see Hanoch and Levy (1969) and Ogryczak and Ruszczynski (2001) and the references therein).

Related to the reward-risk analysis is the reward-risk ratio optimization. Since the publication of the Sharpe ratio, see Sharpe (1966, 1994) and also Ziemba et al. (1974), which is based on the mean-variance analysis, some new performance measures such as the STARR ratio, the Minimax measure, Sortino-Satchell ratio, Farinelli-Tibiletti ratio and most recently the Rachev ratio, the Generalized Rachev ratio and the modified Sharpe ratio have been proposed (for an empirical comparison, see Biglova et al. (2004), Rachev et al. (2005) and the references therein). For additional information on the modified Sharpe ratio, see Ziemba (2005). The new ratios take into account empirically observed phenomena, that assets returns distributions are fat-tailed and skewed, by incorporating proper reward and risk measures. For example, in Ziemba (1974) the mean-dispersion ratio, which is an analogue of the Sharpe ratio, is derived as the proper ratio assuming stock returns follow the heavy-tailed stable-Paretian distribution.

In this paper, we focus on general reward-risk ratio optimization. We begin with a brief description of the Markowitz problem, the related Sharpe ratio optimization and a formulation of a general reward-risk ratio problem. In Section 4, we develop a simplification of the generic problem in the case of positive homogeneous concave and convex, reward and risk measures respectively under some technical conditions. We show how some ratio optimization problems reduce to more simple ones in Sections 5 and 6. In Section 7, an algorithm for a general quasi-concave ratio problem is considered. Finally, in the last two sections, we discuss the Rachev and Generalized Rachev ratio.

2 The mean-variance analysis and the Sharpe ratio

The classical mean-variance framework introduced by Markowitz (1952, 1959) and developed further in Markowitz (1987) is the first proposed model of the reward-risk type and we shall briefly describe it. Suppose that at time $t_0 = 0$ we have an investor who can choose to invest among a universe of n assets. Having made the decision, he keeps the allocation unchanged until the moment t_1 when he can make another investment decision based on the new information accumulated up to t_1. The vector of assets returns $r = (r_1, \dots, r_n)^T$ is stochastic with expected value $Er = (Er_1, \dots, Er_n)^T$. The result of the investment decision is a portfolio with composition $w = (w_1, \dots, w_n)^T$ where w_i is the portfolio weight corresponding to the i-th item. The weights of all portfolio items sum up to 1, $w^T e = \sum_{i=1}^n w_i = 1$ where $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$. The expected portfolio return is

$$\mu_p = \sum_{i=1}^{n} w_i E r_i = w^T E r.$$

A key point in Markowitz's approach is that the standard deviation of portfolio return σ_p is assumed to be the measure of risk. Let $\Sigma = \{\text{cov}(r_i, r_j)\}_{i,j=1}^n$ be the covariance matrix of the random asset returns, then

$$\sigma_p^2 = \sum_{i,j} w_i w_j \text{cov}(r_i, r_j) = w^T \Sigma w.$$

Sometimes the investor faces certain exogenous constraints. For instance, a certain subset of the assets is not allowed to constitute more than a given fraction of total portfolio value. A portfolio that satisfies all constraints in the selection problem will be called *admissible* or *feasible*. Where appropriate, we shall denote the set of all feasible portfolios by $\mathcal{X}$.

The main principle behind the mean-variance analysis[1] can be summarized briefly in two ways:

1. From all feasible portfolios with a given upper bound on σ_p, find the ones that have maximum expected return μ_p;

2. From all feasible portfolios with a given lower bound on μ_p, find the ones that have minimum risk σ_p;

Behind the two formulations of the principle, we can find two optimization problems

$$\begin{array}{ll} \max\limits_{w} & w^T E r \\ \text{subject to} & w^T e = 1 \\ & w^T \Sigma w \leq R^* \\ & Lb \leq Aw \leq Ub \end{array} \tag{1}$$

[1] For a description, see for example Ziemba and Board (1996)

and

$$
\begin{array}{rl}
\min\limits_{w} & w^T \Sigma w \\
\text{subject to} & w^T e = 1 \\
& w^T Er \geq R_* \\
& Lb \leq Aw \leq Ub
\end{array}
\tag{2}
$$

where R^* is the upper bound on portfolio risk, R_* is the lower bound on portfolio return, $A \in \mathbb{R}^{k \times n}$ is a matrix, $Lb \in \mathbb{R}^k$ is a vector of lower bounds and $Ub \in \mathbb{R}^k$ is a vector of upper bounds. The set of k double linear inequalities $Lb \leq Aw \leq Ub$ generalizes all exogenous constraints. The solution of Problem (1) or Problem (2) represents the optimal portfolio or the portfolio that is most preferable among the set of all feasible portfolios. The originally proposed problem by Markowitz is Problem (2) and it is known as the Markowitz problem. For the general dualty theorems of the type of Problems (1) and (2), see Rachev and Rüschendorf (1998). See also, Ziemba and Vickson (2006). The optimal portfolio w^o found in this way is a function of the imposed bounds R^* or R_* depending on whether we consider Problem (1) or Problem (2). Let us choose Problem (2) for the sake of being unambiguous. Then as we have explained $w^o = w^o(R_*)$. Changing the parameter R_* we obtain the set of all optimal portfolios, or the *mean-variance efficient set*. We denote it with $\mathcal{E}_o$. The curve $(w^{oT}Er, w^{oT}\Sigma w^o)$ where $w^o \in \mathcal{E}_o$ is called the *efficient frontier*.

There is a third way to arrive at the mean-variance efficient set by considering the optimization problem

$$
\begin{array}{rl}
\max\limits_{w} & w^T Er - \lambda w^T \Sigma w \\
\text{subject to} & w^T e = 1 \\
& Lb \leq Aw \leq Ub
\end{array}
\tag{3}
$$

where $\lambda > 0$ is a parameter. In this representation, the objective function $w^T Er - \lambda w^T \Sigma w$ is interpreted as a utility function and λ is called *the risk-aversion* parameter. Since it is possible to show that the three problems are equivalent (see, for example, Ziemba and Vickson (2006), Rockafellar and Uryasev (2002) and Palmquist et al. (2002)), the mean-variance efficient set can be obtained by the problem above via varying the risk-aversion parameter.

Suppose that we have received the portfolios from the mean-variance efficient set and that we can compare and choose among all of them. Are we indifferent towards all these portfolios? We can compare them in terms of their expected return for a unit of risk, that is we can compare the ratios

$$
SR(w^o) = \frac{w^{oT} Er}{\sqrt{w^{oT} \Sigma w^o}}
\tag{4}
$$

for all portfolios $w^o \in \mathcal{E}_o$. We would prefer the portfolio with the highest ratio as it provides the highest expected return for a unit of risk. That is we solve the problem

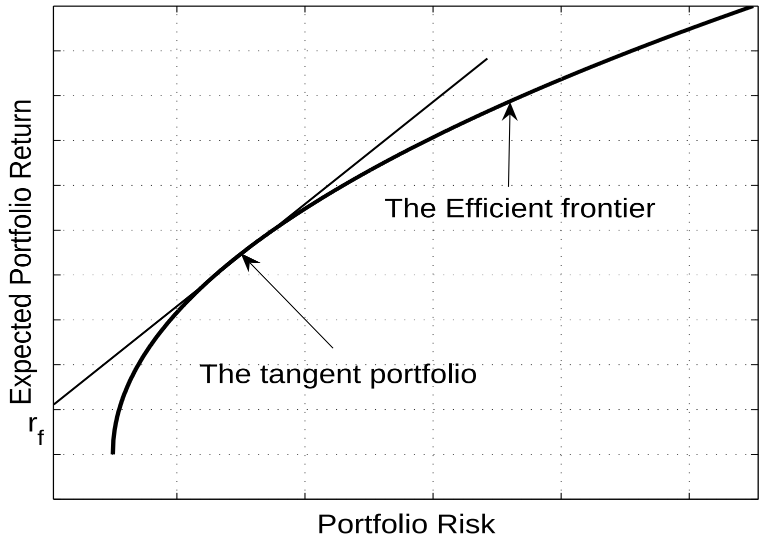

Figure 1: The efficient frontier and the tangent portfolio. Here r_f is a positive risk-free rate.

$$\max_{w \in \mathcal{E}_o} \quad \frac{w^T Er}{\sqrt{w^T \Sigma w}} \tag{5}$$

Geometrically, the point on the efficient frontier that corresponds to the solution of Problem (5) is where a straight line passing through the origin is tangent to the efficient frontier (see Figure 1). The optimal portfolio received is called the *tangent portfolio*. The optimal portfolio is known as the Markowitz market portfolio, or can also be called *tangent portfolio with zero risk-free rate* (see Black (1972)). If the risk-free rate is non-zero, then we maximize the excess return for a unit of risk, $(w^T Er - r_f)/\sqrt{w^T \Sigma w}$. The solution is illustrated on Figure 1.

The ratio defined in equation (4) is a version of the reward-to-variability ratio called the Sharpe ratio, hence the notation. It was first introduced to measure the performance of mutual funds and was originally proposed as the ratio between the expected excess return (the expected return of the fund above a benchmark portfolio return) and the standard deviation of the returns of the fund, see Sharpe (1966, 1994).

3 The ρ-efficient frontier and the ρ-tangent portfolio

In recent years, significant efforts have been dedicated to building extensions to the classical mean-variance analysis. The principal reason is that it leads to optimal decisions only when the vector of assets returns follows the multivariate normal distribution, i.e. $r \in N(Er, \Sigma)$, and there is ample empirical evidence against that assumption[2].

[2]See Ziemba (1974) and also Rachev and Mittnik (2000) and the references therein for extensions compatible with heavy-tailed models.

The extensions involve including different risk measures in the optimization problems. The deficiencies of the standard deviation as a risk measure were acknowledged by Markowitz who was the first to suggest the semi-standard deviation as a substitute, Markowitz (1959). A common criticism is that standard deviation symmetrically penalizes loss as well as profit.

An example of a class of risk measures proposed is the class of the coherent risk measures, see Artzner et al. (1998). A functional $\rho(\cdot)$ on the space of real-valued random variables is called a *coherent risk measure* if it is:

1. monotonous: $X, Y \in V; Y \geq X \quad \Longrightarrow \quad \rho(Y) \leq \rho(X)$
2. sub-additive: $X, Y, X+Y \in V \quad \Longrightarrow \quad \rho(X+Y) \leq \rho(Y) + \rho(X)$
3. positively homogeneous: $X \in V, h > 0, hX \in V \quad \Longrightarrow \quad \rho(hX) = h\rho(X)$
4. translation invariant: $X \in V, a \in \mathbb{R} \quad \Longrightarrow \quad \rho(X+a) = \rho(X) - a$

These axioms are desirable for the purpose of risk management. It is possible to find a representation of coherent risk measures in terms of a functional defined on classes of probability measures, for details see Artzner et al. (1998). Relaxing the positive homogeneity assumption, we obtain the more general class of *convex risk measures*. Representation theorems for convex risk measures can be found in Föllmer and Schied (2002*a*), see also Rockafellar and Ziemba (2000), and for a more extensive treatment, see Föllmer and Schied (2002*b*). In this paper, we do not rely on the strong representation results. We use only the defining axioms.

The mean-variance principle can be extended for a general risk measure $\rho(\cdot)$, not necessarily coherent, and the corresponding optimization problems can be re-stated. For example, Problem (2) becomes:

$$
\begin{array}{rl}
\min\limits_{w} & \rho(w^T r) \\
\text{subject to} & w^T e = 1 \\
& w^T E r \geq R_* \\
& Lb \leq Aw \leq Ub
\end{array}
\tag{6}
$$

The related mean-risk efficient set $\mathcal{E}_o$ is obtained by varying the bound R_* and the efficient frontier is the curve $(w^T Er, \rho(w^T r))$, where $w \in \mathcal{E}_o$. The same reasoning as in the case of the mean-variance model shows that the most preferable portfolio in the set $\mathcal{E}_o$ is among the solutions of

$$
\max_{w \in \mathcal{E}_o} \quad \frac{w^T E r}{\rho(w^T r)}
\tag{7}
$$

The ratio in the objective should be well-defined, that is the denominator should not turn into zero for any feasible portfolio. Otherwise the objective becomes unbounded and Problem (7) makes no sense. Throughout the paper, we will always assume that the risk functional is strictly positive for all feasible portfolios.

If a solution of Problem (7) exists, it is called the *ρ-market portfolio* or *ρ-tangent portfolio with zero risk-free rate*. If a solution exists, it is not necessarily unique. The following simple result holds.

Proposition 1. *The solutions of Problem (7) and*

$$\begin{array}{rl} \max\limits_{w} & \dfrac{w^T Er}{\rho(w^T r)} \\ \textit{subject to} & w^T e = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{8}$$

coincide as well as the objective function values at the solution points.

Proof. First since $\mathcal{E}_o \subset \mathcal{X} = \{w : w^T e = 1, Lb \leq Aw \leq Ub\}$, it follows that

$$\max_{w \in \mathcal{E}_o} \frac{w^T Er}{\rho(w^T r)} \leq \max_{w \in \mathcal{X}} \frac{w^T Er}{\rho(w^T r)} \tag{9}$$

Suppose that the solution of Problem (8) is attained at $\overline{w} \in \mathcal{X}/\mathcal{E}_o$. Then solving Problem (6) with $R_* = \overline{w}^T Er$ we obtain $\overline{w}^o \in \mathcal{E}_0$ such that $\rho(\overline{w}^{oT} r) \leq \rho(\overline{w}^T r)$. Therefore

$$\frac{\overline{w}^T Er}{\rho(\overline{w}^T Er)} \leq \frac{\overline{w}^{oT} Er}{\rho(\overline{w}^{oT} Er)} \tag{10}$$

since $\overline{w}^T Er \leq \overline{w}^{oT} Er$ is also true because of the expected return constraint in the optimization problem. We also know that

$$\max_{w \in \mathcal{X}} \frac{w^T Er}{\rho(w^T Er)} = \frac{\overline{w}^T Er}{\rho(\overline{w}^T Er)} \leq \frac{\overline{w}^{oT} Er}{\rho(\overline{w}^{oT} Er)} \leq \max_{w \in \mathcal{E}_o} \frac{w^T Er}{\rho(w^T Er)}$$

The equality holds because $\overline{w}$ is assumed to be a solution of Problem (8) and the last inequality holds because it is not guaranteed that the global maximum of Problem (7) is attained at $\overline{w}^o$. Combining equation (10) above with equation (9), we can see that the inequalities are satisfied as equalities. Therefore $\overline{w} \in \mathcal{E}_o$ and the corresponding objective function values coincide. □

We can find ρ-tangent portfolios by considering directly Problem (8) and maximizing the corresponding performance measure without finding first the set $\mathcal{E}_o$.

3.1 General reward-risk ratio optimization

Optimization of reward-risk measures is not a new topic. A multitude of performance measures have been proposed in the literature, some of them with non-linear reward measures, see Biglova et al. (2004), Rachev et al. (2005) and the references therein. Here we discuss the general reward-risk ratio optimization problem

$$
\begin{array}{ll}
\max\limits_{w} & \dfrac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \\
\text{subject to} & w^T e = 1 \\
& Lb \leq Aw \leq Ub
\end{array}
\tag{11}
$$

where $\mu(\cdot)$ is a general non-linear reward measure, $\rho(\cdot)$ is a risk measure and r_b is stochastic and denotes the returns of a benchmark portfolio. We assume different properties of $\mu(\cdot)$ and $\rho(\cdot)$ and explore the possible simplification of the problem that would facilitate the numerical solution. In the sequel we treat Problem (11) as an alternative to Problem (8) in which we use *active* portfolio returns, i.e. portfolio returns relative to a benchmark portfolio. At times we will omit r_b in the notation for the sake of brevity.

Our analysis will be based on (quasi-) convex functions, their optimal properties and extensions of some techniques for programming with fractional objectives (see Charnes and Cooper (1962)). We briefly recall that a function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is convex if

$$
f(\alpha x_1 + (1-\alpha)x_2) \leq \alpha f(x_1) + (1-\alpha) f(x_2), \quad \alpha \in [0,1]
$$

where $x_1, x_2 \in D$ and the domain D is a convex set. A function $f : D \subset \mathbb{R}^n \to \mathbb{R}$ is quasi-convex if all sub-level sets $\{x : f(x) \leq t\}$ for t fixed are convex. A function f is concave (quasi-concave) if $-f$ is convex (quasi-convex). Every convex (concave) function is quasi-convex (quasi-concave). The converse is not true. We denote by $\mathbb{R}^{++}$ all strictly positive real numbers, i.e. $\mathbb{R}^{++} = \{x \in \mathbb{R} : x > 0\}$.

First we establish the next simple

Proposition 2. *If $\mu : D_1 \subset \mathbb{R}^n \to \mathbb{R}^{++}$ is a concave function and $\rho : D_2 \subset \mathbb{R}^n \to \mathbb{R}^{++}$ is a convex function then*

a) the ratio $\mu/\rho : D_1 \cap D_2 \to \mathbb{R}^{++}$ is quasi-concave;

b) the ratio $\rho/\mu : D_1 \cap D_2 \to \mathbb{R}^{++}$ is quasi-convex;

c) the following relationship holds

$$
\arg\max_x \frac{\mu(x)}{\rho(x)} = \arg\min_x \frac{\rho(x)}{\mu(x)}, \quad x \in D_1 \cap D_2
$$

Proof. The proofs of a) and b) follow from a more general result in Mangasarian (1970).

c) Let $r(x) = \rho(x)/\mu(x)$. Suppose that $x^o = \arg\min\limits_x r(x)$, $x \in D_1 \cap D_2$. Since by assumption $r(x) > 0$, the difference

$$
\frac{1}{r(x^o)} - \frac{1}{r(x)} = \frac{r(x) - r(x^o)}{r(x^o) r(x)} > 0, \quad x \in D_1 \cap D_2
$$

and therefore $x^o = \arg\max r^{-1}(x)$, $x \in D_1 \cap D_2$. The converse is also true.

□

Part c) in Proposition 2 is correct even if we do not have the assumptions of convexity. The only important property is that the ratio is positive. Unfortunately parts a) and b) cannot be made stronger under the assumed general properties of the numerator and the denominator. For some specific choices of $\mu(x)$ and $\rho(x)$, it could be proved that the corresponding ratios in a) or b) are concave or convex respectively. Nevertheless the quasi-convex (quasi-concave) functions preserve some of the nice properties of the convex (concave) functions with respect to their extrema, see Boyd and Vandenberghe (2004). If the quasi-convex (quasi-concave) function has a univariate argument, it is easy to give a geometric description of its properties. It has one global minimum (maximum) and it is composed of two monotonic parts. The difference from the class of convex (concave) functions is that the two monotonic parts are not strictly monotonic, that is the graph of a quasi-convex (quasi-concave) function might have some "flat" sections, which makes its optimization a more involved affair. Generally an optimization problem with quasi-convex (quasi-concave) objective can be solved by decomposing it into a sequence of convex optimization problems.

In case the domains D_1 and D_2 are open sets and both μ and ρ are differentiable, then instead of a quasi-concave (-convex) ratio we get a pseudo-concave (-convex) ratio. For more details on pseudo- and quasi-convexity, see Mangasarian (1970). Note that pseudo-convex functions are also quasi-convex and, therefore, all results in our paper hold for the smaller class as well. The property which distinguishes the two classes is that the local extrema of a pseudo-convex function are also global, which is not necessarily true for a quasi-convex function. This property makes the differentiable pseudo-convex functions better positioned for numerical optimization with general methods for convex programming because the minima satisfy the Kuhn-Tucker conditions. In the paper, we do not draw special attention to the class of pseudo-convex functions because our approach is constructed for the more general class of quasi-convex functions.

4 Reward-risk ratio of the form $\mu(w^T r)/\rho(w^T r)$

In this section we consider performance measures that have the general form $\mu(w^T r)/\rho(w^T r)$ where $\mu(\cdot)$ is a reward measure and $\rho(\cdot)$ is a risk measure. The reward measure is assumed to be a functional on the space of real-valued random variables that is

1. positive homogeneous: $\mu(tX) = t\mu(X), \quad t > 0$
2. concave: $\mu(\alpha X_1 + (1-\alpha)X_2) \geq \alpha\mu(X_1) + (1-\alpha)\mu(X_2), \quad \alpha \in [0,1]$

The risk-measure is a functional on the space of real-valued random variables which is assumed to be:

1. positive homogeneous: $\rho(tX) = t\rho(X), \quad t > 0$
2. sub-additive: $\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2)$

These two properties guarantee that $\rho(\cdot)$ is a convex functional. Throughout Section 4, the reward and the risk measure are presumed to satisfy the above properties.

Both the reward and the risk of a portfolio with composition $(w_1, \ldots, w_n)$ will be regarded as functions on the set of admissible compositions. In the general situation, when we consider portfolio performance with respect to a benchmark, it is easy to show that both functions are concave and convex accordingly. We are not providing the proof as it is straightforward.

Proposition 3. *Suppose that $\mu(\cdot)$ and $\rho(\cdot)$ are functionals satisfying the above conditions. Then the reward function $\mu(w^T r - r_b) : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}^{++}$ is concave and the risk function $\rho(w^T r - r_b) : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}^{++}$ is convex, provided that the domain $\mathcal{X}$ is a convex set.*

Combining the above result with Proposition 2 we have the next

Corollary 1. *Suppose that $\mu(\cdot)$ and $\rho(\cdot)$ are functionals as in Proposition 3. Then the general performance ratio optimization problem (11) is a quasi-concave problem.*

Because of the assumed positive homogeneity of the reward and the risk measure, Problem (11) can be simplified and reduced to convex optimization problems.

Proposition 4. *The general performance measure optimization Problem (11) is equivalent to the following two problems*

$$
(A)\quad
\begin{array}{ll}
\max\limits_{(x,t)} & \mu(x^T r - t r_b) \\
\text{subject to} & \rho(x^T r - t r_b) \leq 1 \\
 & x^T e = t \\
 & tLb \leq Ax \leq tUb \\
 & t \geq 0
\end{array}
\qquad\qquad
(B)\quad
\begin{array}{ll}
\min\limits_{(x,t)} & \rho(x^T r - t r_b) \\
\text{subject to} & \mu(x^T r - t r_b) \geq 1 \\
 & x^T e = t \\
 & tLb \leq Ax \leq tUb \\
 & t \geq 0
\end{array}
$$

in the following sense. If the pair (x_A^o, t_A^o) is a solution to Problem (A) and (x_B^o, t_B^o) is a solution to Problem (B), then $w^o = x_A^o/t_A^o = x_B^o/t_B^o$ solves Problem (11). Moreover $\mu(x_A^{oT} r - t_A^o r_b) = \mu(w^{oT} r - r_b)/\rho(w^{oT} r - r_b) = \rho^{-1}(x_B^{oT} r - t_B^o r_b)$. Conversely, if w^o is a solution to Problem (11) and $t^o = \rho^{-1}(w^{oT} r - r_b)$, then the pair $(t^o w^o, t^o)$ is a solution to Problem (A). If $t^o = \mu^{-1}(w^{oT} r - r_b)$, then $(t^o w^o, t^o)$ is a solution to Problem (B).

Problem (B) is a convex problem and Problem (A) can be easily transformed into a convex problem by changing the sign of the objective and considering minimization.

The proof of Proposition 4 is given in the Appendix. One of the key assumptions is that the reward measure is strictly positive for the feasible portfolios. Certainly this is very restrictive as one might argue that there could be feasible portfolios with negative excess reward, $\mu(w^T r - r_b) \leq 0$. In case the set $\mathcal{X}$ contains portfolios of both types — with positive and negative excess reward — then obviously the optimal solution does not belong to the set of the latter. Thus we might consider the optimization over $\mathcal{X} \cap \{w : \mu(w^T r - r_b) \geq \epsilon > 0\}$ which is a convex set and does not change the nature of the problem. The parameter ϵ is such that the intersection is non-empty. Therefore the proper less restrictive assumption for $\mu(\cdot)$ is that *there should exist at least one $w \in \mathcal{X}$ such that $\mu(w^T r - r_b) > 0$.*

Another important assumption is that the risk measure is a positive functional for all feasible portfolios. If $\rho(\cdot)$ is not bounded away from zero for all feasible portfolios, i.e. there exists $w \in \mathcal{X}$ such that $\rho(w^T r - r_b) = 0$, then Proposition 4 does not hold. In this case the variable t in Problem (A) explodes since $t = \rho^{-1}(w^T r - r_b)$, which makes the problem unbounded. The solution of Problem B will be a portfolio $w \in \mathcal{X}$ such that $\rho(w^T r - r_b) = 0$ because for the other portfolios the objective is strictly positive. Certainly, due to the assumed convexity, existence of a feasible portfolio w with $\rho(w^T r - r_b) < 0$ means that there exists a feasible portfolio v, such that $\rho(v^T r - r_b) = 0$. As a result, we can conclude that strict positivity of the risk measure for all feasible portfolios is crucial for the stated equivalence between the three problems.

Concerning the risk functional, our analysis is primarily based on the properties of positive homogeneity and sub-additivity. Therefore, from mathematical viewpoint, it is possible to replace the risk functional by any positive homogeneous, sub-additive functional which is strictly positive for the admissible portfolios and Proposition (4) will still hold. For example, one can use the class of deviation measures, see Rockafellar et al. (2005) and Rockafellar et al. (2006). They are by definition non-negative for all random variables. Deviation measures are not risk measures and thus they capture different aspects of the portfolio return distribution. In effect, the optimization problems involving deviation measures are, generally, different in nature than those involving risk measures. Nevertheless all results in the paper remain valid when $\rho(\cdot)$ is a deviation measure. One example is the Sharpe ratio problem in which the denominator is a symmetric deviation measure. The Sharpe ratio problem is considered in Section 6.

5 Reward-risk ratio of the form $w^T Er/\rho(w^T r)$

If we choose as a reward measure the mathematical expectation, then $\mu(\cdot)$ becomes a linear functional[3].

The performance measures with a linear reward functional arise naturally from general optimal portfolio problems like Problem (6). Assuming the same properties for the risk measure as in the previous section, Proposition 4 transforms into

Proposition 5. *The general performance measure optimization Problem (11) is equivalent to the following two problems*

$$
\text{(A1)} \quad
\begin{array}{rl}
\max\limits_{(x,t)} & x^T Er - tEr_b \\
\textit{subject to} & \rho(x^T r - tr_b) \le 1 \\
& x^T e = t \\
& tLb \le Ax \le tUb \\
& t \ge 0
\end{array}
\qquad\qquad
\text{(B1)} \quad
\begin{array}{rl}
\min\limits_{(x,t)} & \rho(x^T r - tr_b) \\
\textit{subject to} & x^T Er - tEr_b = 1 \\
& x^T e = t \\
& tLb \le Ax \le tUb \\
& t \ge 0
\end{array}
$$

in the sense that if the pair (x^o_A, t^o_A) is a solution to Problem (A1) and (x^o_B, t^o_B) is a solution to Problem (B1), then $w^o = x^o_A/t^o_A = x^o_B/t^o_B$ solves Problem (11). Moreover

[3]Actually a much stronger result holds. If the reward functional is assumed to be linear, then it is necessarily the mathematical expectation. This is the Riesz representation theorem, see DeGiorgi (2004).

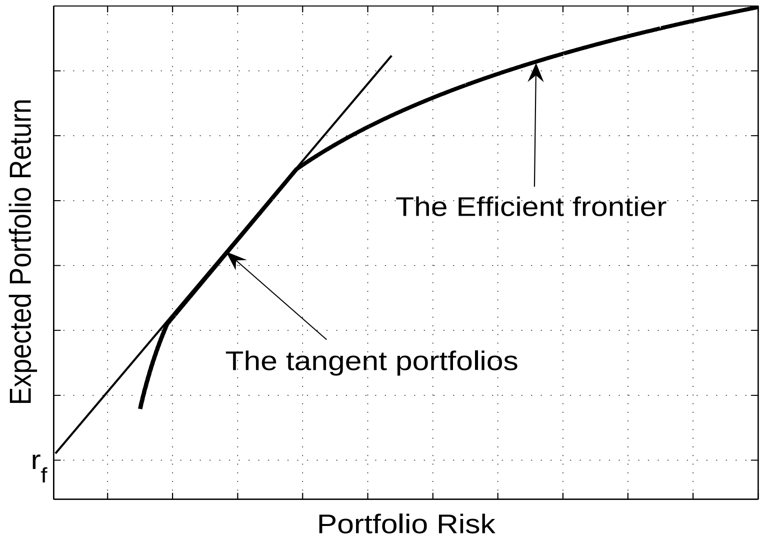

Figure 2: The efficient frontier, a case with a non-unique tangent portolio. Here r_f is a positive risk-free rate.

$x_A^{oT}Er - t_A^o Er_b = (w^{oT}Er - Er_b)/\rho(w^{oT}r - r_b) = \rho^{-1}(x_B^{oT}r - t_B^o r_b)$. *Conversely, if* w^o *is a solution to Problem (11) and* $t^o = \rho^{-1}(w^{oT}r - r_b)$*, then the pair* $(t^o w^o, t^o)$ *is a solution to Problem (A1). If* $t^o = (w^{oT}Er - Er_b)^{-1}$*, then* $(t^o w^o, t^o)$ *is a solution to Problem (B1).*

Proof. The claim follows trivially from Proposition 4 replacing the general reward measure $\mu(\cdot)$ with the mathematical expectation. For more details, see the proof of Proposition 4 in the Appendix. □

We take advantage of the linearity property and in Problem (B1) the additional constraint is equality in contrast to the additional constraint in Problem (B). From numerical viewpoint, it makes Problem (B1) more attractive than Problem (A1).

The connection with the optimal portfolio problem makes interesting the question of the interplay between Problem (A1) and Problem (B1) and how that is related to the efficient frontier and its graphical representation. It is possible to show that if both problems have two different solutions, then any convex combination of them is again a solution.

Proposition 6. *If* (x^1, t^1) *and* (x^2, t^2) *are two distinct solutions of Problem (A1) or Problem (B1), then* $(x^\lambda, t^\lambda) = \lambda(x^1, t^1) + (1-\lambda)(x^2, t^2)$*,* $\lambda \in [0, 1]$ *is also a solution to Problem (A1) or Problem (B1).*

The proof is given in the Appendix. Two distinct solutions are possible if the risk measure is not strictly convex. In terms of the efficient frontier, it means that there are linear sections and a portfolio that maximizes the performance measure happens to be in the linear section (see Figure 2). Because of the established equivalence of both problems in Proposition 5, from practical viewpoint we can choose to solve one or the other. If we solve both and we arrive at two distinct solutions then it follows that we have such a case.

Actually Figure 2 represents the general view of the efficient frontier, it is not a special case. This is proved in a more general situation in Section 7.

6 Particular examples

We can easily relate the general problems above to some specific reward-risk ratios considered by practitioners. In those specific cases, the optimization problems can be further simplified.

6.1 The Sharpe ratio

The Sharpe ratio optimization problem can be reduced to two problems — one of type A and one of type B in which the functional $\rho(\cdot)$ is the standard deviation of portfolio returns. This is just one example in which $\rho(\cdot)$ is not a risk measure but a deviation measure. In the general case, the variance of the excess return is

$$D(w^T r - r_b) = \sum_{i=1}^{n} w_i^2 Dr_i + \sum_{i \neq j} w_i w_j \text{cov}(r_i, r_j) + 2 \sum_{i=1}^{n} w_i \text{cov}(r_i, -r_b) + Dr_b$$

where DX stands for the variance of the random variable X. In matrix form

$$D(w^T r - r_b) = (-1, w)^T \Sigma_1 (-1, w), \qquad \Sigma_1 = \left(\begin{array}{c|c} \sigma_b^2 & \sigma_{br} \\ \hline \sigma_{br}^T & \Sigma \end{array} \right),$$

σ_b^2 is the variance of the benchmark asset r_b,

$$\sigma_{br} = (\text{cov}(r_b, r_1), \text{cov}(r_b, r_2), \dots, \text{cov}(r_b, r_n))$$

is the vector of covariances of the returns of the benchmark asset with the returns of the other assets $r = (r_1, r_2, \dots, r_n)^T$ and $\Sigma = \{\text{cov}(r_i, r_j)\}_{i,j=1}^n$ is the covariance matrix of the assets' returns. The notation $(-1, w)$ stands for the vector column $(-1, w_1, \dots, w_n)^T$ and $(-1, w)^T$ is the corresponding vector row.

The performance measure optimization problem is obtained from Problem (8)

$$\begin{array}{rl} \max\limits_{w} & \dfrac{w^T Er - Er_b}{\sqrt{(-1, w)^T \Sigma_1 (-1, w)}} \\ \text{subject to} & w^T e = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{12}$$

The result is contained in the next

Proposition 7. *The performance measure optimization Problem (12) is equivalent to the following two problems*

$$\text{(SR A)} \qquad \begin{array}{rl} \max\limits_{(x,t)} & x^T Er - tEr_b \\ \textit{subject to} & (-t,x)^T \Sigma_1 (-t,x) \leq 1 \\ & x^T e = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

and

$$\text{(SR B)} \qquad \begin{array}{rl} \min\limits_{(x,t)} & (-t,x)^T \Sigma_1 (-t,x) \\ \textit{subject to} & x^T Er - tEr_b = 1 \\ & x^T e = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

in the sense that if the pair (x_A^o, t_A^o) is a solution to Problem (SR A) and (x_B^o, t_B^o) is a solution to Problem (SR B), then $w^o = x_A^o/t_A^o = x_B^o/t_B^o$ solves Problem (12). Moreover

$$\begin{aligned} x_A^{oT} Er - t_A^o Er_b &= (w^{oT} Er - Er_b)/\sqrt{(-1,w^o)^T \Sigma_1 (-1,w^o)} \\ &= \left(\sqrt{(-t_B^o, x_B^o)^T \Sigma_1 (-t_B^o, x_B^o)}\right)^{-1}. \end{aligned}$$

Conversely, if w^o is a solution to Problem (12) and $t^o = \left(\sqrt{(-1,w^o)^T \Sigma_1 (-1,w^o)}\right)^{-1}$, then the pair $(t^o w^o, t^o)$ is a solution to Problem (SR A). If $t^o = (w^{oT} Er - Er_b)^{-1}$, then $(t^o w^o, t^o)$ is a solution to Problem (SR B).

Proof. When $\rho((x,t)) = \sqrt{(-t,x)^T \Sigma (-t,x)}$ in the problem of type A in Proposition 5 we have the constraint $(-t,x)^T \Sigma (-t,x) \leq 1$. Raising both sides of the inequality to the second power does not change the optimization problem since any (x,t) satisfying the constraint $\sqrt{(-t,x)^T \Sigma (-t,x)} \leq 1$ also satisfies $(-t,x)^T \Sigma (-t,x) \leq 1$ and vice versa.

In the problem of type B in Proposition 5, the objective is

$$\sqrt{(-t,x)^T \Sigma (-t,x)} \to \min_{(x,t)}$$

Raising the objective to the second power does not change the optimal solution points. This is a direct consequence of the fact that Σ is a covariance matrix and is positive semi-definite, i.e. $(-t,x)^T \Sigma (-t,x) \geq 0$ for any (x,t), and that the function $f(u) = u^2$ is strictly monotonic if $u \geq 0$. $\square$

Problem (SR B) is quadratic and can be solved as a quadratic programming problem. The reciprocal of the objective function value obtained at an optimal solution point does not equal the maximal Sharpe ratio but the squared maximal Sharpe ratio. If the pair (x_B, t_B) solves Problem (SR B), then the reciprocal of the objective function is

$$\begin{aligned}
\left((-t_B, x_B)^T \Sigma_1 (-t_B, x_B)\right)^{-1} &= \left(t_B^2 (-1, w_B)^T \Sigma_1 (-1, w_B)\right)^{-1} \\
&= \left(\frac{(-1, w_B)^T \Sigma_1 (-1, w_B)}{(w_B^T Er - Er_b)^2}\right)^{-1} \\
&= \left(\frac{w_B^T Er - Er_b}{\sqrt{(-1, w_B)^T \Sigma_1 (-1, w_B)}}\right)^2
\end{aligned}$$

For the Markowitz problem, the functional $\rho(\cdot)$, which is represented by the standard deviation, is strictly convex if the covariance matrix is non-singular and therefore, assuming non-singular covariance matrix, Figure 1 presents the general case for the view of the graph of the efficient frontier.

6.2 The STARR ratio

The expected tail loss (ETL), also known as expected shortfall, is proposed in the literature as a superior alternative to the industry standard Value-at-Risk (VaR), for a discussion see Yamai and Yoshiba (2002*b*,*a*). If the portfolio return distribution is continuous, then the ETL coincides with the Conditional Value at Risk (CVaR) introduced in Rockafellar and Uryasev (2002). In this case, the ETL (or CVaR) of portfolio return X at the $100(1-\alpha)$ percent confidence level is defined as

$$ETL_\alpha(X) = CVaR_\alpha(X) = E(-X | -X > VaR_\alpha(X)) \tag{13}$$

where $VaR_\alpha(X)$ is defined as $VaR_\alpha(X) = -\inf\{x \in \mathbb{R} : P(X \leq x) \geq \alpha\}$ and is the VaR measure. If the portfolio return distribution is discontinuous and there is a jump at $VaR_\alpha(X)$, the definition of the CVaR should be modified in order for the risk measure to be coherent. For a discussion, see Acerbi and Tasche (2002) and Rockafellar and Uryasev (2002). The definition of the CVaR changes to

$$CVaR_\alpha(X) = \int_{-\infty}^{\infty} t dG_X(t) \tag{14}$$

where

$$G_X(t) = \begin{cases} 0, & t < VaR_\alpha(X) \\ (P(-X \leq t) - (1-\alpha))/\alpha, & t \geq VaR_\alpha(X) \end{cases}$$

Clearly (14) transforms to (13) when there is no jump of the distribution function at $VaR_\alpha(X)$. In the sequel, we will use the notation CVaR and the more general definition given in (14).

Considering the optimal portfolio Problem 6, we arrive at the STARR ratio as the natural reward-risk measure associated with the problem. The STARR ratio is defined as

$$STARR(w) = \frac{w^T Er - Er_b}{CVaR_\alpha(w^T r - r_b)} \tag{15}$$

and is proposed in Martin et al. (2003). Since the CVaR is a coherent risk measure, it satisfies the conditions that we impose on the risk function $\rho(\cdot)$ in Section 4 and we can use Proposition 5 with $\rho(w^T r - r_b) = CVaR_\alpha(w^T r - r_b)$. For both problems of type A and B, we can apply the linearization technique developed in Rockafellar and Uryasev (2002) when there are scenarios for portfolio items returns.

Suppose that we have N vectors $r^k \in \mathbb{R}^n$, $k = 1, \dots, N$ with scenarios for the returns of all portfolio items. They could be N random draws from the joint multivariate distribution of the assets returns. The portfolio returns scenarios are $s_k = w^T r^k$, $k = 1, 2, \dots, N$ where w is a concrete vector of weights. Using the scenarios, we can compute an estimate $\widehat{CVaR}$ of portfolio CVaR:

$$\widehat{CVaR}_\alpha(w^T r) = \frac{1}{[N\alpha]} \sum_{k=1}^{[N\alpha]} (-1) s_{(k)} \tag{16}$$

where $s_{(k)}$, $k = 1, \dots, N$ denote the sorted portfolio scenarios in increasing order and $[a]$ is the largest integer smaller than a. It is shown in Rockafellar and Uryasev (2002) that the same estimate can be obtained via minimization

$$\widehat{CVaR}_\alpha(w^T r) = \min_{\theta \in \mathbb{R}} \left(\theta + \frac{1}{N\alpha} \sum_{k=1}^{N} \left[-w^T r^k - \theta \right]^+ \right) \tag{17}$$

where $[a]^+ = \max(a, 0)$. Taking advantage of the latter representation and applying the approach in Rockafellar and Uryasev (2002) to the corresponding optimization problems from Proposition 5, we obtain the following linearizations:

$$
\text{(STARR A)} \qquad
\begin{array}{ll}
\max\limits_{(x,t,d,\theta)} & x^T Er - tEr_b \\
\text{subject to} & \theta + \dfrac{1}{N\alpha} \displaystyle\sum_{k=1}^{N} d_k \leq 1 \\
 & -x^T r^k + tr_b^k - \theta \leq d_k, \quad k = 1, 2, \dots, N \\
 & x^T e = t \\
 & tLb \leq Ax \leq tUb \\
 & t \geq 0, \quad d_k \geq 0, \quad k = 1, 2 \dots N
\end{array}
$$

and

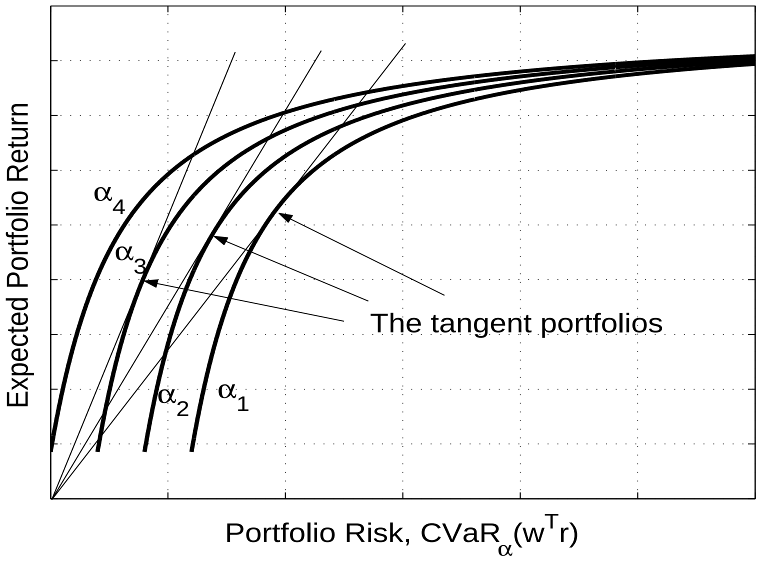

Figure 3: The efficient frontier changes as the parameter α of the CVaR increases, $\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$; $r_b = 0$

$$
\text{(STARR B)} \qquad
\begin{array}{rl}
\min\limits_{(x,t,d,\theta)} & \theta + \dfrac{1}{N\alpha}\displaystyle\sum_{k=1}^{N} d_k \\
\text{subject to} & x^T Er - tEr_b = 1 \\
& -x^T r^k + tr_b^k - \theta \leq d_k, \quad k = 1, 2, \ldots, N \\
& x^T e = t \\
& tLb \leq Ax \leq tUb \\
& t \geq 0, \quad d_k \geq 0, \quad k = 1, 2 \ldots N
\end{array}
$$

In both problems r_b^k, $k = 1, \ldots, N$ denote the scenarios for the benchmark portfolio and $d = (d_1, \ldots, d_N)$ is a vector of additional variables, one for each scenario.

We have assumed in Section 2 that the risk measure $\rho(\cdot)$ is a positive functional for all feasible portfolios. Due to the translation invariance property, the CVaR measure can violate this assumption, as any other coherent risk measure. The CVaR of a portfolio is

$$CVaR_\alpha(w^T r) = CVaR_\alpha(w^T(r - \mu_p) + \mu_p) = CVaR_\alpha(w^T(r - \mu_p)) - \mu_p$$

In this expression, $CVaR_\alpha(w^T(r - \mu_p))$ is the CVaR of the centralized portfolio returns. The defining equation (14) implies the inequality

$$CVaR_\alpha(w^T r) \leq CVaR_\beta(w^T r), \quad \beta < \alpha$$

meaning that we can always choose α to be sufficiently high, such that $CVaR_\alpha(w^T r) = 0$ because μ_p does not depend on α. Geometrically increasing α continuously will shift the entire efficient frontier to the left until the "tangent" line becomes vertical (see Figure 3).

The generic ratio optimization problem also reduces to a set of two linear problems if $\rho(\cdot)$ belongs to the larger family of spectral risk measures. The CVaR measure belongs to this class — it is a spectral measure with constant spectral function. The linearizations can be received following the same reasoning as in the current sub-section. For more details on spectral risk measures and the linearization of the optimal portfolio problem, see Acerbi (2002), Acerbi and Simonetti (2002) and Acerbi (2004).

7 Reward-risk ratio of the form $\mu(w^T r)/\rho(w^T r)$ as a quasi-concave problem

It is possible to relax some of the assumed properties of $\mu(\cdot)$ and $\rho(\cdot)$ in Section 4 and still have a quasi-concave ratio optimization problem. Actually, the only properties that we need are that $\rho(\cdot)$ be a convex functional, positive for the admissible portfolios and $\mu(\cdot)$ — a concave functional also positive for the admissible portfolios. Then Proposition 3 remains valid and using Proposition 2 we can conclude that we have a quasi-concave problem.

Unfortunately, without the positive homogeneity property we cannot establish the nice equivalent problems in Proposition 4. Nevertheless we can use that the problem is quasi-concave and find the global maximum as the generic optimization problem reduces to a sequence of convex problems (see Boyd and Vandenberghe (2004)). According to Proposition 2, the ratio optimization Problem (11) can be equivalently re-stated

$$\begin{array}{ll} \min\limits_{w} & \dfrac{\rho(w^T r - r_b)}{\mu(w^T r - r_b)} \\ \text{subject to} & w^T e = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{18}$$

The objective function in the above problem is quasi-convex. Now consider the following feasibility problem

$$\begin{array}{l} \rho(w^T r - r_b) - t\mu(w^T r - r_b) \leq 0 \\ w^T e = 1 \\ Lb \leq Aw \leq Ub \end{array} \tag{19}$$

where t is a fixed positive number. For every fixed $t \geq t_{\min} = \min_{w \in \mathcal{X}} \rho(w^T r - r_b)/\mu(w^T r - r_b)$, the set $\{w : \rho(w^T r) - t\mu(w^T r) \leq 0\}$ is convex and therefore we have a convex feasibility problem. Also for every $t \geq t_{\min}$, Problem (19) is feasible because the set $\{w : \rho(w^T r - r_b) - t\mu(w^T r - r_b) \leq 0\}$ is non-empty.

A simple algorithm using bisection can be devised on the basis of the above fact. Suppose that we know an interval $[a, b]$ such that $t_{\min} \in [a, b]$ and we have a tolerance level ϵ. Then

1. Set $t = (a + b)/2$.

2. Solve the feasibility Problem (19).

3. If the problem is feasible, then set $b = t$. If the problem is infeasible then set $a = t$ and go to step 1.

4. Repeat the above steps until $b - a < \epsilon$.

The final interval $[a, b]$ is guaranteed to contain the solution $t_{\min}$ since $t_{\min} \in [a, b]$ on each iteration. In k iterations, the length of the final interval is $2^{-k}(b-a)$ where $[a, b]$ is the length of the initial interval. Therefore, a total of $\log_2((b-a)/\epsilon)$ steps are required until convergence (see Boyd and Vandenberghe (2004) for a general description of the quasi-convex problem).

Other methods for numerical solution of (18) can be constructed under the additional assumption of differentiability of the numerator and the denominator. Then, as we noted, the ratio is a differentiable pseudo-convex function, the minima of which satisfy the Kuhn-Tucker optimality conditions. Therefore, problem (18) can directly be solved using numerical methods for convex programming.

In the first section, we related the Sharpe ratio to the Markowitz problem. Next we related the ratio of the form $w^T Er/\rho(w^T r)$ to the more general optimal portfolio Problem (6). In a similar way now we can associate an optimization problem with the reward-risk ratio considered in the current section. Let us examine the optimization problem

$$\begin{array}{ll} \max\limits_{w} & \mu(w^T r - r_b) \\ \text{subject to} & \rho(w^T r - r_b) \leq R^* \\ & w^T e = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{20}$$

The optimal vector of weights w^o is dependent on the imposed upper bound R^*, $w^o = w^o(R^*)$. Just as in the Markowitz setting, we can call the set of all optimal allocations the efficient set, denote it with $\mathcal{E}_o$, and the curve $(\mu(w^T r - r_b), \rho(w^T r - r_b))$, $w \in \mathcal{E}_o$ the efficient frontier. It appears that the corresponding version of Proposition 1 remains valid

Proposition 8. *The solutions of Problem (11) and*

$$\max_{w \in \mathcal{E}_o} \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)}$$

coincide as well as the objective function values at the solution points.

Proof. The same arguments as in Proposition 1 can be applied without modification. □

The geometry of the efficient frontier under the assumptions of the current section is described in

Proposition 9. *The efficient frontier generated by Problem (20) — in which $\mu(\cdot)$ is a concave reward measure and $\rho(\cdot)$ is a convex risk measure — is a concave monotonically increasing function. That is if R_1^* and R_2^*, $R_1^* < R_2^*$ are two upper bounds of the risk constraint and $R_\lambda^* = \lambda R_1^* + (1-\lambda)R_2^*$, $\lambda \in [0,1]$, then*

$$\mu\left((w^o(R_\lambda^*))^T r - r_b\right) \geq \lambda\mu\left((w^o(R_1^*))^T r - r_b\right) + (1-\lambda)\mu\left((w^o(R_2^*))^T r - r_b\right)$$

and

$$\mu\left((w^o(R_2^*))^T r - r_b\right) \geq \mu\left((w^o(R_1^*))^T r - r_b\right)$$

where $w^o(R^*)$ denotes the optimal allocation obtained with R^* as a risk constraint and the risk constraint is assumed binding for $R^* = R_1^*$ and $R^* = R_2^*$.

The proof is given in the Appendix. Using these two propositions, we can conclude that the general view of the graph of the efficient frontier is like the one on Figure 2. Thus the geometric intuition supports the analytic result that in the case of concave reward measure and convex risk measure, the set of all globally optimal portfolios is convex.

8 Non-quasi-concave reward-risk ratios

There are reward-risk ratios suggested in literature that are not in the class of the quasi-concave functions because both the numerator and the denominator are convex. Such are for instance the Farinelli-Tibiletti ratio and the Generalized Rachev ratio (see Biglova et al. (2004) and Rachev et al. (2005)). Under certain conditions, it appears that simplification is possible in some particular examples.

8.1 The Rachev ratio

In this section we consider the problem of the Rachev ratio optimization which is a special case of the Generalized Rachev ratio. The definition is

$$RR(w) = \frac{CVaR_\alpha(r_b - w^T r)}{CVaR_\beta(w^T r - r_b)} \tag{21}$$

The reward measure is the CVaR function which is convex and it follows that the Rachev ratio does not belong to the class of quasi-concave ratios because both the numerator and the denominator are convex and Proposition 2 does not hold. Nevertheless we can still use the positive homogeneity property of the numerator and the denominator and simplify the generic problem

$$
\begin{array}{ll}
\max\limits_{w} & \dfrac{CVaR_\alpha(r_b - w^T r)}{CVaR_\beta(w^T r - r_b)} \\
\text{subject to} & w^T e = 1 \\
& Lb \leq Aw \leq Ub
\end{array}
\tag{22}
$$

where both the numerator and the denominator are assumed to be strictly positive functionals for all admissible portfolios. The result is contained in the next proposition.

Proposition 10. *Problem (22) is equivalent to the following two optimization problems*

$$
\text{(RR A)} \quad
\begin{array}{ll}
\max\limits_{(x,t)} & CVaR_\alpha(tr_b - x^T r) \\
\textit{subject to} & CVaR_\beta(x^T r - tr_b) \leq 1 \\
& x^T e = t \\
& tLb \leq Ax \leq tUb \\
& t \geq 0
\end{array}
$$

and

$$
\text{(RR B)} \quad
\begin{array}{ll}
\min\limits_{(x,t)} & CVaR_\beta(x^T r - tr_b) \\
\textit{subject to} & CVaR_\alpha(tr_b - x^T r) \geq 1 \\
& x^T e = t \\
& tLb \leq Ax \leq tUb \\
& t \geq 0
\end{array}
$$

in the sense that if the pair (x^o_A, t^o_A) is a solution to Problem (RR A) and (x^o_B, t^o_B) is a solution to Problem (RR B), then $w^o = x^o_A/t^o_A = x^o_B/t^o_B$ solves Problem (22). Moreover $CVaR_\alpha(t^o_A r_b - x^{oT}_A r) = CVaR_\alpha(r_b - w^{oT} r)/CVaR_\beta(w^{oT} r - r_b) = (CVaR_\beta(x^{oT}_B r - t^o_B r_b))^{-1}$. Conversely, if w^o is a solution to Problem (22) and $t^o = (CVaR_\beta(x^{oT}_B r - r_b))^{-1}$, then the pair $(t^o w^o, t^o)$ is a solution to Problem (RR A). If $t^o = (CVaR_\alpha(r_b - w^{oT} r))^{-1}$, then $(t^o w^o, t^o)$ is a solution to Problem (RR B).

Proof. The same arguments from the proof of Proposition 4 can be applied without modification. □

Both problems in the proposition are not convex. In Problem (RR A), the set of all admissible portfolios is convex because the sub-level set $\{(x,t) : CVaR_\beta(x^T r - tr_b) \leq 1\}$ is convex. The objective involves maximization of a convex function and for this reason the problem is non-convex. In Problem (RR B), the objective is the minimization of a convex function but the set $\{(x,t) : CVaR_\alpha(tr_b - x^T r) \geq 1\}$ is non-convex and so is the entire problem.

Basically the lack of convexity makes it impossible to linearize both CVaR functions simultaneously in the problems using the approach in Rockafellar and Uryasev (2002) when we have scenarios for the returns of the assets. Suppose that we try with Problem (RR A) with $r_b = 0$ to simplify the expressions. The corresponding linearization is

$$
\begin{array}{lll}
\max\limits_{(x,t,g,\delta,d,\theta)} & \delta + \dfrac{1}{N\alpha}\displaystyle\sum_{k=1}^{N} g_k & \\
\text{subject to} & x^T r^k - \delta \le g_k, \quad k = 1,2,\ldots,N & \\
& g_k \ge 0, \quad k = 1,2\ldots N & \\
& \theta + \dfrac{1}{N\alpha}\displaystyle\sum_{k=1}^{N} d_k \le 1 & (23) \\
& -x^T r^k - \theta \le d_k, \quad k = 1,2,\ldots,N & \\
& d_k \ge 0, \quad k = 1,2\ldots N & \\
& x^T e = t & \\
& tLb \le Ax \le tUb & \\
& t \ge 0 &
\end{array}
$$

Inevitably the objective of the above problem explodes because we have maximization and there is nothing to bound the variables g_k from above. All constraints concerning those variables are of the type $x^T r^k - \delta \le g_k$ and therefore all g_k increase indefinitely. There is no such restriction for δ as well and it also increases indefinitely. Now suppose that we try to "fix" the problem by changing the constraints to be $x^T r^k - \delta \ge g_k$. The new problem could turn out infeasible if for one k, $0 > x^T r^k - \delta \ge g_k$, because we require all g_k to be positive.

We propose a linearization introducing binary variables into the optimization. Thus the resulting problem is a mixed-integer linear programming (MIP) problem. The linearization is based on equation (16) and on the consideration that if the CVaR estimate is computed as the average of the largest $[N\alpha]$ observations in the sample multiplied by negative one, then the average of any other $[N\alpha]$ observations will be smaller than the CVaR. The above statement transforms into the optimization problem

$$
\begin{array}{lll}
\max\limits_{(g,\lambda)} & \dfrac{1}{[N\alpha]}\displaystyle\sum_{k=1}^{N} g_k & \\
\text{subject to} & g_k \le B\lambda_k, \quad k = 1,2,\ldots,N & \\
& g_k \ge -w^T r^k - B(1-\lambda_k), \quad k = 1,2,\ldots,N & (24) \\
& g_k \le -w^T r^k + B(1-\lambda_k), \quad k = 1,2,\ldots,N & \\
& \lambda^T e = [N\alpha] & \\
& \lambda_k \in \{0,1\}, \quad g_k \ge 0, \quad k = 1,2,\ldots,N &
\end{array}
$$

where $r^k \in \mathbb{R}^n$, $k = 1,\ldots,N$ are vectors of assets returns scenarios, α is the parameter of the CVaR function, B is a very large number, such that $|w^T r^k| \le B$ for all scenarios, and $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N)$ is a vector of binary variables. Let us verify if Problem (24) is equivalent to the above statement. For this reason, we have to examine two cases:

1. Suppose that $\lambda_k = 0$. Then

$$
\begin{aligned}
&g_k \le 0\\
&g_k \ge -w^T r^k - B\\
&g_k \le -w^T r^k + B\\
&g_k \ge 0
\end{aligned}
$$

From the first and the fourth inequality it follows that $g_k = 0$. The second and the third are redundant since B is chosen to be very large.

2. Suppose that $\lambda_k = 1$.

$$
\begin{aligned}
&g_k \le B\\
&g_k \ge -w^T r^k\\
&g_k \le -w^T r^k\\
&g_k \ge 0
\end{aligned}
$$

From the last three equalities it follows that $g_k = -w^T r^k \ge 0$. The first inequality is redundant.

Thus the binary variable λ_k indicates whether the variable g_k is non-zero. If $\lambda_k = 1$ then the auxiliary variable is positive and equals the *k-th* scenario of the portfolio return multiplied by negative one. If it is negative, then it forces the binary variable to become equal to zero, since g_k cannot be negative. Ultimately in the objective function we have a sum of positive numbers equal to the corresponding portfolio return scenarios or zeros. The number of the non-zero summands, or respectively the non-zero λ_k, is controlled by the constraint $\lambda^T e = [N\alpha]$ according to which the sum of all λ_k should equal $[N\lambda]$. Certainly $\widehat{CVaR}_\alpha(w^T r)$ matches the solution of Problem (24).

We have demonstrated that for each w fixed by solving Problem (24) we obtain $\widehat{CVaR}_\alpha(w^T r)$. We can combine Problem (RR A) and Problem (24) as the objective of the former is also maximization. As a result we have

$$
\begin{array}{ll}
\max\limits_{(x,t,g,\lambda,d,\theta)} & \dfrac{1}{[N\alpha]} \displaystyle\sum_{k=1}^{N} g_k \\
\text{subject to} & g_k \le B\lambda_k, \quad k = 1, 2, \dots, N \\
& g_k \ge x^T r^k - B(1-\lambda_k), \quad k = 1, 2, \dots, N \\
& g_k \le x^T r^k + B(1-\lambda_k), \quad k = 1, 2, \dots, N \\
& \lambda^T e = [N\alpha] \\
& \lambda_k \in \{0,1\}, \quad g_k \ge 0, \quad k = 1, 2, \dots, N \\
& \theta + \dfrac{1}{N\alpha} \displaystyle\sum_{k=1}^{N} d_k \le 1 \\
& -x^T r^k - \theta \le d_k, \quad k = 1, 2, \dots, N \\
& d_k \ge 0, \quad k = 1, 2 \dots N \\
& x^T e = t \\
& tLb \le Ax \le tUb \\
& t \ge 0.
\end{array}
\tag{25}
$$

Solving Problem (25) we obtain the global maximum of the Rachev ratio. The number of binary variables λ_k equals the number of scenarios and the computational burden is much more related to the number of scenarios than to the portfolio size. Actually, the computational cost of a MIP problem is much higher than that of a linear problem with continuous variables only. Certainly this approach to the CVaR linearization is applicable in the STARR ratio optimization problem and in general, in the optimal portfolio problem with the CVaR as a risk function, but the method of Rockafellar and Uryasev (2002) is more efficient as it involves continuous variables only.

8.2 The Generalized Rachev ratio

The Generalized Rachev ratio is defined as

$$GRR(w) = \frac{CVaR_{(\gamma,\alpha)}(r_b - w^T r)}{CVaR_{(\delta,\beta)}(w^T r - r_b)} \tag{26}$$

where $CVaR_{(\gamma,\alpha)}(X) = E\left((\max(-X,0))^{\gamma} | -X > VaR_{\alpha}(X)\right)$ and $VaR_{\alpha}(X)$ is the VaR measure. For simplicity we will assume that X has a continuous distribution. The Rachev ratio is a special case with $\gamma = \delta = 1$. The power function in the conditional expectation does not allow for a straightforward linearization as in the case of the STARR ratio or the Rachev ratio. Like the Rachev ratio, in the Generalized Rachev ratio both the numerator and the denominator are convex if $\gamma \geq 1$ and $\delta \geq 1$. This result is contained in the next

Proposition 11. *The function $CVaR_{(\gamma,\alpha)}(r_b - w^T r) : \mathcal{X} \subset \mathbb{R}^n \to \mathbb{R}^{++}$ is convex for $\gamma \geq 1$, provided that $\mathcal{X}$ is a convex set.*

The proof is given in the Appendix. As a consequence, the optimization reward-risk ratio problem for $\gamma, \delta \geq 1$ is not quasi-concave. Still it is possible to obtain a pair of problems using the homogeneity of the numerator and the denominator. The basic fact is included in the next

Proposition 12. *The functional $CVaR_{(\gamma,\alpha)}(X)$ as defined in equation (26) is positive homogeneous of degree γ, that is*

$$CVaR_{(\gamma,\alpha)}(tX) = t^{\gamma} CVaR_{(\gamma,\alpha)}(X)$$

where $t > 0$.

The proof is given in the Appendix. The generic reward-risk ratio optimization problem is

$$\begin{array}{ll} \max\limits_{w} & \dfrac{CVaR_{(\gamma,\alpha)}(r_b - w^T r)}{CVaR_{(\delta,\beta)}(w^T r - r_b)} \\ \text{subject to} & w^T e = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{27}$$

Both the numerator and the denominator are positive, this is clear from the definition. Now we can state a similar pair of problems as in the case of the Rachev ratio. Such simplification appears possible only if $\gamma = \delta$.

Proposition 13. *Problem (27) with $\gamma = \delta$ is equivalent to the following two optimization problems*

$$\text{(GRR A)} \quad \begin{array}{rl} \max\limits_{(x,t)} & CVaR_{(\delta,\alpha)}\left(tr_b - x^T r\right) \\ \text{subject to} & CVaR_{(\delta,\beta)}\left(x^T r - tr_b\right) \leq 1 \\ & x^T e = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

and

$$\text{(GRR B)} \quad \begin{array}{rl} \min\limits_{(x,t)} & CVaR_{(\delta,\beta)}\left(x^T r - tr_b\right) \\ \text{subject to} & CVaR_{(\delta,\alpha)}\left(tr_b - x^T r\right) \geq 1 \\ & x^T e = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

in the sense that if the pair (x_A^o, t_A^o) is a solution to Problem (GRR A) and (x_B^o, t_B^o) is a solution to Problem (GRR B), then $w^o = x_A^o/t_A^o = x_B^o/t_B^o$ solves Problem (22). Moreover $CVaR_{(\delta,\alpha)}(t_A^o r_b - x_A^{oT} r) = CVaR_{(\delta,\alpha)}(r_b - w^{oT} r)/CVaR_{(\delta,\beta)}(w^{oT} r - r_b) = (CVaR_{(\delta,\beta)}(x_B^{oT} r - t_B^o r_b))^{-1}$. Conversely, if w^o is a solution to Problem (22) and $t^o = (CVaR_{(\delta,\beta)}(w^{oT} r - r_b))^{-1}$, then the pair $((t^o)^{1/\delta} w^o, (t^o)^{1/\delta})$ is a solution to Problem (GRR A). If $t^o = (CVaR_{(\delta,\alpha)}(r_b - w^{oT} r))^{-1}$, then $((t^o)^{1/\delta} w^o, (t^o)^{1/\delta})$ is a solution to Problem (GRR B).

The proof is given in the Appendix. Neither Problem (GRR A) nor Problem (GRR B) are convex programming problems for $\gamma \geq 1$. We reach this conclusion using the same analysis as in the case of the Rachev ratio. Nevertheless they could prove to be beneficial.

If $\gamma \neq \delta$, the numerator and the denominator are in different units. For the purpose of GRR optimization, we can standardize both functions and modify the ratio in equation (26) as

$$MGRR(w) = \frac{\left(CVaR_{(\gamma,\alpha)}\left(r_b - w^T r\right)\right)^{1/\gamma}}{\left(CVaR_{(\delta,\beta)}\left(w^T r - r_b\right)\right)^{1/\delta}}. \tag{28}$$

Proposition 12 implies that the numerator and the denominator are both positive homogeneous of degree 1. It is easy to obtain another pair of problems using the reasoning in Proposition 4. Actually, Problems (GRR A) and (GRR B) appear as special cases.

Proposition 14. *Problem (27) with equation (28) as objective is equivalent to the following two optimization problems*

$$
\text{(MGRR A)} \qquad \begin{array}{rl} \max\limits_{(x,t)} & CVaR_{(\gamma,\alpha)}\left(tr_b - x^T r\right) \\ \textit{subject to} & CVaR_{(\delta,\beta)}\left(x^T r - tr_b\right) \le 1 \\ & x^T e = t \\ & tLb \le Ax \le tUb \\ & t \ge 0 \end{array}
$$

and

$$
\text{(MGRR B)} \qquad \begin{array}{rl} \min\limits_{(x,t)} & CVaR_{(\delta,\beta)}\left(x^T r - tr_b\right) \\ \textit{subject to} & CVaR_{(\gamma,\alpha)}\left(tr_b - x^T r\right) \ge 1 \\ & x^T e = t \\ & tLb \le Ax \le tUb \\ & t \ge 0 \end{array}
$$

in the sense that if the pair (x^o_A, t^o_A) is a solution to Problem (MGRR A) and (x^o_B, t^o_B) is a solution to Problem (MGRR B), then $w^o = x^o_A/t^o_A = x^o_B/t^o_B$ solves the corresponding reward-risk ratio problem. Moreover $(CVaR_{(\gamma,\alpha)}(t^o_A r_b - x^{oT}_A r))^{1/\gamma} = (CVaR_{(\gamma,\alpha)}(r_b - w^{oT} r))^{1/\gamma}/(CVaR_{(\delta,\beta)}(w^{oT} r - r_b))^{1/\delta} = (CVaR_{(\delta,\beta)}(x^{oT}_B r - t^o_B r_b))^{-1/\delta}$. Conversely, if w^o is a solution to the reward-risk ratio problem and $t^o = (CVaR_{(\delta,\beta)}(w^{oT} r - r_b))^{-1/\delta}$, then the pair $(t^o w^o, t^o)$ is a solution to Problem (GRR A). If $t^o = (CVaR_{(\gamma,\alpha)}(r_b - w^{oT} r))^{-1/\gamma}$, then $(t^o w^o, t^o)$ is a solution to Problem (GRR B).

Proof. We show how to deal with the equivalence with Problem (MGRR A). The same arguments as in Proposition 4 and the substitution $t = (CVaR_{(\delta,\beta)}(w^T r - r_b))^{-1/\delta}$ lead to a maximization problem with objective

$$
(CVaR_{(\gamma,\alpha)}(tr_b - x^T r))^{1/\gamma} \to \max_{(x,t)}
$$

and risk constraint $(CVaR_{(\delta,\beta)}(x^T r - tr_b))^{1/\delta} \le 1$. Raising the objective to the power γ does not change the solution points as the transformation is strictly increasing. Raising the inequality to power δ does not change the feasibility either. Thus we realize Problem (MGRR A). □

The objectives of Problems (MGRR A) and (MGRR B) at the optimal points do not equal the optimal reward-risk ratio values because of the increasing transform that we apply to change them.

The reward-risk ratio MGRR(w) is very similar to the Farinelli-Tibiletti ratio. If we consider the excess portfolio returns with a non-zero benchmark, the Farinelli-Tibiletti ratio is defined as (see Farinelli, Tibiletti (2003) and Biglova et. al. (2004)):

$$
FT(w) = \frac{\left(E\left(w^T r - r_b\right)^p_+\right)^{1/p}}{\left(E\left(w^T r - r_b\right)^q_-\right)^{1/q}}
$$

where $(w^T r - r_b)_+^p = (\max(w^T r - r_b, 0))^p$ and $(w^T r - r_b)_-^q = (\max(r_b - w^T r, 0))^q$. Clearly $FT(w)$ is a special case of $MGRR(w)$ as the conditional expectation can reduce to the unconditional one.

9 Conclusion

We consider the problem of reward-risk ratio optimization as an optimal portfolio selection problem, imposing different properties on reward and risk measures. When the mathematical expectation is the reward measure and the risk measure is convex and linearizable, it is possible to reduce the generic performance measure optimization to linear programming problems. If the risk measure is not linearizeable, the general problem reduces to convex programming problems and in the special case of the Sharpe ratio, to a quadratic problems. Simplification is also possible if the reward functional is concave. In that case, the quasi-concave ratio optimization problem reduces to a sequence of convex feasibility problems. In the special case of the Rachev ratio, we propose a mixed-integer linear program for finding the global maximum. The Generalized Rachev ratio and a modified version of it are also considered.

Appendix

In the Appendix, we give the proofs of some of the Propositions.

- **Proof of Proposition 4.**

Proof. First we show that Problem (11) is equivalent to Problem (A). The proof is in two steps. In the first step, we show that Problem (11) is equivalent to

$$\begin{array}{ll} \max\limits_{(x,t)} & \mu(x^T r - t r_b) \\ \text{subject to} & \rho(x^T r - t r_b) = 1 \\ & x^T e = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array} \tag{29}$$

Indeed, substituting $t = \rho^{-1}(w^T r - r_b) > 0$ and using the assumed positive homogeneity, we arrive at the problem above. For any feasible point w in Problem (11), we have that $x = tw$ is feasible in Problem (29). Thus

$$\max_{w \in \mathcal{X}} \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \leq \max_{(x,t) \in \mathcal{X}_1} \mu(w^T r - t r_b)$$

where $\mathcal{X}$ denotes the set of feasible portfolios in Problem (11) and $\mathcal{X}_1$ denotes the set of feasible (x, t) pairs in Problem (29). Conversely, if (x, t) is feasible in Problem (29), then $w = x/t$ is feasible in Problem (11) and

$$\max_{w \in \mathcal{X}} \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \geq \max_{(x,t) \in \mathcal{X}_1} \mu(w^T r - t r_b)$$

Combining both inequalities, we see that the function values coincide at the solutions. In addition if w is a solution to Problem (11), then (tw, t) is a solution to Problem (29) and if (x, t) is a solution to Problem (29) then $w = x/t$ is a solution to Problem (11).

In the second step, let us consider Problem (29) and Problem (A) where the risk function constraint has been relaxed. We claim that if (x^o, t^o) is a solution to Problem (29), then $\rho(x^{oT} r - t^o r_b) = 1$ and therefore the pair is a solution to Problem (29). Let us assume that the risk constraint is non-binding at the solution point, that is $\rho(x^{oT} r - t^o r_b) = t^o \rho(w^{oT} r - r_b) < 1$. Problem A can be restated as

$$\begin{array}{ll} \max\limits_{(w,t)} & t\mu(w^T r - r_b) \\ \text{subject to} & t\rho(w^T r - r_b) \leq 1 \\ & tw^T e = t \\ & tLb \leq Atw \leq tUb \\ & t \geq 0 \end{array}$$

Obviously, multiplying t by any constant a such that $1 < a \leq t^o \rho^{-1}(w^{oT} r - r_b)$ does not change the feasibility of $(at^o w^o, at^o)$ and

$$at^o\mu(w^{oT}r - r_b) > t^o\mu(w^{oT}r - r_b)$$

since according to the assumed properties of $\mu(\cdot)$, it is strictly positive. Therefore we arrive at a contradiction that (x^o, t^o) is a solution. It follows that at a solution point, the risk constraint is binding.

The equivalence of Problem (B) and Problem (11) can be established first by using Proposition 2, part (c) and then by applying the same reasoning as above to the problems:

$$\begin{array}{rl} \min\limits_{w} & \dfrac{\rho(w^Tr - r_b)}{\mu(w^Tr - r_b)} \\ \text{subject to} & \\ & w^Te = a \\ & Lb \leq Aw \leq Ub \end{array} \quad \text{and} \quad \begin{array}{rl} \min\limits_{(x,t)} & \rho(x^Tr - tr_b) \\ \text{subject to} & \mu(x^Tr - tr_b) = 1 \\ & x^Te = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

We arrive at Problem (B) by changing the reward constraint to $\mu(x^Tr - tr_b) \geq 1$. Using the same arguments as above, it is easy to show that it is binding at the solution. ☐

- **Proof of Proposition 6.**

Proof. We prove the claim for Problem (B1). Then the result holds for Problem (A1) because of the established equivalence between them in Proposition (5).

It is straightforward to verify that (x^λ, t^λ) satisfies all constraints in the problem. We show how to deal with all of them:

- the first one: $x^TEr - tEr_b = 1$.

$$x^{\lambda T}Er - t^\lambda Er_b = \lambda(x^{1T}Er - t^1Er_b) + (1-\lambda)(x^{2T}Er - t^2Er_b) = 1$$

  The first equality follows since x^λ and t^λ are convex combinations of x^1, x^2 and t^1, t^2 respectively. The second equality holds because of the assumption that (x^1, t^1) and (x^2, t^2) are optimal solutions and therefore are feasible points and satisfy all constraints.

- the second one: $x^Te = t$.
  Since both points (x^1, t^1) and (x^2, t^2) are feasible, then $x^{1T}e = t^1$ and $x^{2T}e = t^2$. Therefore

$$\lambda x^{1T}e + (1-\lambda)x^{2T}e = \lambda t^1 + (1-\lambda)t^2$$

  and it follows that $x^{\lambda T}e = t^\lambda$.

- the third one: $tLb \leq Ax \leq tUb$.
  Since both points (x^1, t^1) and (x^2, t^2) are feasible, then $t^1Lb \leq Ax^1 \leq t^1Ub$ and $t^2Lb \leq Ax^2 \leq t^2Ub$. Multiplying by λ and $(1-\lambda)$ respectively and summing parts by parts we obtain

$$(\lambda t^1 + (1-\lambda)t^2)Lb \leq A(\lambda x^1 + (1-\lambda)x^2) \leq (\lambda t^1 + (1-\lambda)t^2)Ub$$

which is the same as $t^\lambda Lb \leq Ax^\lambda \leq t^\lambda Ub$.

Using once again the assumption that both points are solutions, it follows that

$$\rho(x^{1T}r - t^1 r_b) = \rho(x^{2T}r - t^2 r_b) = \rho_{\min}$$

The assumed convexity of the objective (see Proposition 3) implies that

$$\begin{aligned}\rho(x^{\lambda T}r - t^\lambda r_b) &= \rho(\lambda(x^{1T}r - t^1 r_b) + (1-\lambda)(x^{2T}r - t^2 r_b)) \\ &\leq \lambda\rho(x^{1T}r - t^1 r_b) + (1-\lambda)\rho(x^{2T}r - t^2 r_b) = \rho_{\min}\end{aligned}$$

Strict inequality is not possible since $\rho_{\min}$ is the minimum of the objective over all feasible vectors. Therefore $\rho(x^{\lambda T}r - t^\lambda r_b) = \rho_{\min}$ and we have proved the claim. ☐

- **Proof of Proposition 9.**

Proof. The monotonic property is easiest to establish. Since $R_1^* < R_2^*$ and $\rho(w^T r - r_b)$ is a convex function of w (see Proposition 3), then $\mathcal{X}_1 = \{w : \rho(w^T r - r_b) \leq R_1^*, w^T e = 1, Lb \leq Aw \leq Ub\} \subset \mathcal{X}_2 = \{w : \rho(w^T r - r_b) \leq R_2^*, w^T e = 1, Lb \leq Aw \leq Ub\}$. Thus

$$\max_{x \in \mathcal{X}_1} \mu\left(w^T r - r_b\right) \leq \max_{x \in \mathcal{X}_2} \mu\left(w^T r - r_b\right)$$

The risk constraint is binding at the solution points which means that $\rho((w^o(R_i^*))^T r - r_b) = R_i^*$, $i = 1, 2$. Let us construct the portfolio $w^\lambda = \lambda w^o(R_1^*) + (1-\lambda)w^o(R_2^*)$, $\lambda = [0, 1]$. The two portfolios $w^o(R_1^*)$ and $w^o(R_2^*)$ satisfy all constraints because they are optimal solutions and are feasible points. All linear constraints in the problem form a convex set, therefore the convex combination w^λ satisfies all linear constraints. In addition,

$$\begin{aligned}\rho(w^{\lambda T}r - r_b) &\leq \lambda\rho\left((w^o(R_1^*))^T r - r_b\right) + (1-\lambda)\rho\left((w^o(R_2^*))^T r - r_b\right) \\ &= \lambda R_1^* + (1-\lambda)R_2^*\end{aligned}$$

and

$$\mu(w^{\lambda T}r - r_b) \geq \lambda\mu\left((w^o(R_1^*))^T r - r_b\right) + (1-\lambda)\mu\left((w^o(R_2^*))^T r - r_b\right)$$

The inequalities follow because $\rho(w^T r - r_b)$ and $\mu(w^T r - r_b)$ are respectively convex and concave functions of w. If we solve the optimization problem with $R^* = R_\lambda^* = \lambda R_1^* + (1-\lambda)R_2^*$, we obtain the optimal portfolio $w^o(R_\lambda^*)$ with reward

$$\mu\left((w^o(R_\lambda^*))^T r - r_b\right) \geq \mu(w^{\lambda T}r - r_b)$$

which, combined with the above inequality of the reward proves the claim. Moreover, the risk constraint is binding at $w^o(R^*_\lambda)$. Assume the contrary, that

$$\rho\left((w^o(R^*_\lambda))^T r - r_b\right) < R^*_\lambda$$

and let $\mathcal{X}_\lambda = \{w : \rho(w^T r - r_b) \leq R^*_\lambda, w^T e = 1, Lb \leq Aw \leq Ub\}$. Since $R^*_1 \leq R^*_\lambda \leq R^*_2$ and $\rho(w^T r - r_b)$ is a convex function of w, $\mathcal{X}_\lambda \subset \mathcal{X}_2$. But then

$$\rho\left((w^o(R^*_\lambda))^T r - r_b\right) = \rho\left((w^o(R^*_2))^T r - r_b\right) < R^*_2$$

and we have a contradiction with the assumed equality $\rho((w^o(R^*_2))^T r - r_b) = R^*_2$.

□

- **Proof of Proposition 11.**

  We will use the following result in the proof of Proposition 11.

  Lemma 1. *Suppose that the functional $G_{H,\epsilon}(X)$ is defined as*

$$G_{H,\epsilon}(X) = E(H(\max(X,0))|X > q_\epsilon(X)) \tag{30}$$

where X is a real-valued random variable with a continuous distribution, $q_\epsilon(X)$ is the ϵ-quantile of X, $H(\cdot)$ is a non-decreasing, convex function from $[0,\infty]$ to $[0,\infty]$ and vanishes at the origin. Then $G_{H,\epsilon}(X)$ is convex, in the sense that if $Z = aX + (1-a)Y,\ a \in [0,1]$, then $G_{H,\epsilon}(Z) \leq aG_{H,\epsilon}(X) + (1-a)G_{H,\epsilon}(Y)$.

Proof. We will assume that X and Y are real-valued random variables defined on a common probability space $(\Omega, \mathcal{F}, P)$. Suppose that $Z = aX + (1-a)Y$, where $a \in [0,1]$. Then

$$\begin{aligned}
L &:= -aG_{H,\epsilon}(X) - (1-a)G_{H,\epsilon}(Y) + G_{H,\epsilon}(Z) \\
&= -aG_{H,\epsilon}(X) - (1-a)G_{H,\epsilon}(Y) + G_{H,\epsilon}(aX + (1-a)Y) \\
&= -aE(H(\max(X,0))|X > q_\epsilon(X)) - (1-a)E(H(\max(Y,0))|Y > q_\epsilon(Y)) \\
&+ E(H(\max(Z,0))|Z > q_\epsilon(Z)) \\
&\leq -aE(H(\max(X,0))|X > q_\epsilon(X)) - (1-a)E(H(\max(Y,0))|Y > q_\epsilon(Y)) \\
&+ aE(H(\max(X,0))|Z > q_\epsilon(Z)) + (1-a)E(H(\max(Y,0))|Z > q_\epsilon(Z)) \\
&= \frac{a}{1-\alpha}\left[E(H(\max(X,0))\mathbf{1}_{\{Z>q_\epsilon(Z)\}}) - E(H(\max(X,0))\mathbf{1}_{\{X>q_\epsilon(X)\}})\right] \\
&+ \frac{1-a}{1-\alpha}\left[E(H(\max(Y,0))\mathbf{1}_{\{Z>q_\epsilon(Z)\}}) - E(H(\max(Y,0))\mathbf{1}_{\{Y>q_\epsilon(Y)\}})\right] \\
&= \frac{a}{1-\epsilon}\left[E\left(H(\max(X,0))\left(\mathbf{1}_{\{Z>q_\epsilon(Z)\}} - \mathbf{1}_{\{X>q_\epsilon(X)\}}\right)\right)\right] \\
&+ \frac{1-a}{1-\epsilon}\left[E\left(H(\max(Y,0))\left(\mathbf{1}_{\{Z>q_\epsilon(Z)\}} - \mathbf{1}_{\{Y>q_\epsilon(Y)\}}\right)\right)\right] = \frac{a}{1-\epsilon}A + \frac{1-a}{1-\epsilon}B
\end{aligned}$$

The inequality holds because, (1) $\max(x,0)$ is convex and thus $\max(aX + (1-a)Y, 0) \leq a\max(X,0) + (1-a)\max(Y,0)$ and (2) H is assumed non-decreasing and convex, therefore

$$
\begin{aligned}
H(\max(aX+(1-a)Y,0)) &\leq H(a\max(X,0)+(1-a)\max(Y,0)) \\
&\leq aH(\max(X,0))+(1-a)H(\max(Y,0))
\end{aligned}
$$

The expressions A and B will be handled separately.

$$
\begin{aligned}
A &= \int_{\Omega} H(\max(X(w),0)) \left(\mathbf{1}_{\{Z(w)>q_\epsilon(Z)\}} - \mathbf{1}_{\{X(w)>q_\epsilon(X)\}}\right) dP \\
&= \int_{S_1} H(\max(X(w),0)) f(w) dP + \int_{\overline{S}_1} H(\max(X(w),0)) f(w) dP
\end{aligned}
$$

where $f(w) = \mathbf{1}_{\{Z(w)>q_\epsilon(Z)\}} - \mathbf{1}_{\{X(w)>q_\epsilon(X)\}}$, $S_1 = \{w : X(w) > q_\epsilon(X)\}$ and also $S_1 \bigcup \overline{S}_1 = \Omega$. It is easy to see that $f(w) \leq 0,\ w \in S_1$ and $f(w) \geq 0,\ w \notin S_1$. Therefore, since $H(\max(X(w),0))$ is non-negative and non-decreasing, we obtain the bound

$$
\begin{aligned}
A &\leq H(\max(q_\epsilon(X),0)) \int_{S_1} f(w) dP + H(\max(q_\epsilon(X),0)) \int_{\overline{S}_1} f(w) dP \\
&= H(\max(q_\epsilon(X),0)) \int_{\Omega} f(w) dP \\
&= H(\max(q_\epsilon(X),0)) E\left(\mathbf{1}_{\{Z(w)>q_\epsilon(Z)\}} - \mathbf{1}_{\{X(w)>q_\epsilon(X)\}}\right) \\
&= H(\max(q_\epsilon(X),0))(1-\epsilon-1+\epsilon) = 0
\end{aligned}
$$

Similarly we receive $B \leq 0$ and thus finally we arrive at $L \leq 0$. As a result, for any $a \in [0,1]$,

$$
G_{H,\epsilon}(aX+(1-a)Y) \leq aG_{H,\epsilon}(X) + (1-a)G_{H,\epsilon}(Y)
$$

□

9.0.1 Proof of the main result

Proof. It is easy to check that $z(y) = y^\gamma$ is convex and non-decreasing for $\gamma \geq 1$ and $y \geq 0$; $(z'(y) = \gamma y^{\gamma-1} \geq 0,\ z''(y) = \gamma(\gamma-1)y^{\gamma-2} \geq 0)$ and therefore it satisfies the conditions in Lemma 1.

Then, noticing that

$$
\begin{aligned}
CVaR_{\gamma,\alpha}(X) &= E((\max(-X,0))^\gamma | -X > VaR_\alpha(X)) \\
&= E((\max(-X,0))^\gamma | -X > -q_\alpha(X)) \\
&= E((\max(-X,0))^\gamma | -X > q_{1-\alpha}(-X)) \\
&= G_{z,(1-\alpha)}(-X)
\end{aligned}
$$

we are ready with the proof making use of the result in Lemma 1. □

- **Proof of Proposition 12.**

  Proof. According to the definition, for any $t > 0$

$$\begin{aligned}
CVaR_{(\gamma,\alpha)}(tX) &= E\left((\max(-tX,0))^{\gamma}| - tX > VaR_{\alpha}(tX)\right) \\
&= \frac{1}{\alpha}\int_{VaR_{\alpha}(tX)}^{\infty}(\max(-x,0))^{\gamma} f_{-tX}(x)dx \\
&= \frac{1}{\alpha}\int_{tVaR_{\alpha}(X)}^{\infty}(\max(-x,0))^{\gamma} f_{-X}\left(\frac{x}{t}\right)\frac{dx}{t} \\
&= \frac{1}{\alpha}\int_{VaR_{\alpha}(X)}^{\infty}(\max(-ty,0))^{\gamma} f_{-X}\left(y\right)dy \\
&= \frac{t^{\gamma}}{\alpha}\int_{VaR_{\alpha}(X)}^{\infty}(\max(-y,0))^{\gamma} f_{-X}\left(y\right)dy
\end{aligned}$$

  where $f_X(x)$ is the probability density function of the random variable X. Hence

$$\begin{aligned}
CVaR_{(\gamma,\alpha)}(tX) &= t^{\gamma}E\left((\max(-X,0))^{\gamma}| - X > VaR_{\alpha}(X)\right) \\
&= t^{\gamma}CVaR_{(\gamma,\alpha)}(X)
\end{aligned}$$

  and we have proved the claim. □

- **Next we give a proof of Proposition 13.**

  Proof. We sketch the proof of equivalence with Problem (GRR A). Using the homogeneity property in Proposition 12 and substituting $t = (CVaR_{(\delta,\beta)}(w^T r - r_b))^{-\delta}$ we obtain

$$\begin{array}{cl}
\max\limits_{(w,t)} & CVaR_{(\delta,\alpha)}\left(tr_b - tw^T r\right) \\
\text{subject to} & CVaR_{(\delta,\beta)}\left(tw^T r - tr_b\right) = 1 \\
& w^T e = 1 \\
& Lb \leq Aw \leq Ub \\
& t \geq 0
\end{array}$$

  We multiply by t the second and the third constraint and then change the variables $x = tw$. In effect we have

$$\begin{array}{cl}
\max\limits_{(x,t)} & CVaR_{(\delta,\alpha)}\left(tr_b - x^T r\right) \\
\text{subject to} & CVaR_{(\delta,\beta)}\left(x^T r - tr_b\right) = 1 \\
& x^T e = t \\
& tLb \leq Ax \leq tUb \\
& t \geq 0
\end{array} \tag{31}$$

  The same reasoning as in the proof of Proposition 4 establishes the equivalence of Problem (27) with Problem (31). That is, if (x^o, t^o) is a solution of Problem (31), then $w^o = x^o/t^o$ is a solution to Problem (27) and conversely, if w^o is a solution to Problem (27) then the pair $(w^o t^o, t^o)$ solves Problem (31) with $t^o = (CVaR_{(\delta,\beta)}(w^{oT} r - r_b))^{-\delta}$.

Also in the same way as in the proof of Proposition 4, we can show that it is possible to relax the risk constraint to $CVaR_{(\delta,\beta)}\left(x^Tr - tr_b\right) \leq 1$ since it is necessarily binding. Indeed, due to the homogeneity property stated in Proposition 12, we can rewrite the entire problem as

$$\begin{array}{ll} \max\limits_{(w,t)} & t^\delta CVaR_{(\delta,\alpha)}\left(r_b - w^Tr\right) \\ \text{subject to} & t^\delta CVaR_{(\delta,\beta)}\left(w^Tr - r_b\right) \leq 1 \\ & tw^Te = t \\ & tLb \leq Atw \leq tUb \\ & t \geq 0 \end{array}$$

We assume that the solution is attained at (x^o, t^o) which is such that $(t^o)^\delta CVaR_{(\delta,\beta)}(w^{oT}r - r_b) < 1$. Now let us consider the pair (at^ow^o, at^o) where

$$1 < a \leq \left((t^o)^\delta CVaR_{(\delta,\beta)}\left(w^{oT}r - r_b\right)\right)^{-1/\delta}$$

The point (at^ow^o, at^o) is feasible and

$$(t^o)^\delta CVaR_{(\delta,\alpha)}\left(r_b - w^{oT}r\right) < a^\delta (t^o)^\delta CVaR_{(\delta,\alpha)}\left(r_b - w^{oT}r\right)$$

because $a^\delta > 1$. This inequality shows that the objective function value increases at (at^ow^o, at^o) and therefore $(t^ow^o, t^o) = (x^o, t^o)$ is not the solution according to the assumption. Because of the established contradiction, the risk constraint is satisfied as equality at the solution point.

The equivalence with Problem (GRR B) is proved with the same arguments applied to the problems

$$\begin{array}{ll} \min\limits_{w} & \dfrac{CVaR_{(\delta,\beta)}(w^Tr - r_b)}{CVaR_{(\delta,\alpha)}(r_b - w^Tr)} \\ \text{subject to} & w^Te = 1 \\ & Lb \leq Aw \leq Ub \end{array} \tag{32}$$

and

$$\begin{array}{ll} \min\limits_{(x,t)} & CVaR_{(\delta,\beta)}\left(x^Tr - tr_b\right) \\ \text{subject to} & CVaR_{(\delta,\alpha)}\left(tr_b - x^Tr\right) = 1 \\ & x^Te = t \\ & tLb \leq Ax \leq tUb \\ & t \geq 0 \end{array}$$

where $t = (CVaR_{(\delta,\alpha)}(r_b - w^Tr))^{-\delta}$ and then considering the relaxation $CVaR_{(\delta,\alpha)}(tr_b - x^Tr) \geq 1$ of the risk constraint. The fact that Problem (32) is equivalent to the generic ratio optimization, Problem (27) follows from the same arguments as in the proof of Proposition 2, part (c). Actually the result there is not dependent on any assumption of convexity.

□

References

Acerbi, C. (2002), 'Spectral measures of risk: a coherent representation of subjective risk aversion', *working paper, Abaxbank, Milano* .

Acerbi, C. (2004), 'Coherent representation of subjective risk aversion', *Chapter 10 in G. Szego (ed.), Risk measures for the 21st century, Wiley, Chichester* pp. 147–208.

Acerbi, C. and D. Tasche (2002), 'On the coherence of expected shortfall', *Journal of Banking and Finance* **26, (7)**, 1487–1503.

Acerbi, C. and P. Simonetti (2002), 'Portfolio optimization with spectral measures of risk', *working paper, Abaxbank, Milano* .

Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1998), 'Coherent measures of risk', *Math. Fin.* **6**, 203–228.

Biglova, A., S. Ortobelli, S. Rachev and S. Stoyanov (2004), 'Different approaches to risk estimation in portfolio theory', *The Journal of Portfolio Management* **Fall 2004**, 103–112.

Black, F. (1972), 'Capital market equilibrium with restricted borrowing', *The Journal of Business* **45, (3)**, 444–458.

Boyd, S. and L. Vandenberghe (2004), *Convex optimization*, Cambridge University Press.

Charnes, A. and W. W. Cooper (1962), 'Programming with linear fractional functionals', *Naval Research Logistics Quarterly* **9**, 181–186.

DeGiorgi, E. (2004), 'Reward-risk portfolio selection and stochastic dominance', *Journal of Banking and Finance, forthcoming* .

Föllmer, H. and A. Schied (2002*a*), 'Robust preferences and convex risk measures', *Advances in Finance and Stochastics, Essays in Honor of Dieter Sondermann, Eds. K. Sandmann, Ph. Schonbucher. Berlin, Heidelberg, New York: Springer.* pp. 39–56.

Föllmer, H. and A. Schied (2002*b*), *Stochastic finance: An introduction in discrete time, second revised and extended revision*, Walter de Gruyter, Berlin.

Hanoch, G. and H. Levy (1969), 'The efficiency analysis of choices involving risk', *Review of economic studies* **36**, 335–346.

Mangasarian, O. (1970), 'Convexity, pseudo-convexity and quasi-convexity of composite functions', *Cahiers du Centre dEtudes de Recherche Operationnelle* **12**, 114–122.

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

Markowitz, H. M. (1959), *Portfolio selection: Efficient diversification of investments*, John Wiley & Sons, New York.

Markowitz, H. M. (1987), *Mean-variance analysis in portfolio choice and capital markets*, Basil Blackwell, New York.

Martin, D., S. T. Rachev and F. Siboulet (2003), 'Phi-alpha optimal portfolios and extreme risk management', *Wilmott* **November**, 70–83.

Ogryczak, W. and A. Ruszczynski (2001), 'On consistency of stochastic dominance and mean-semideviation models', *Mathematical Programming* **89**, 217–232.

Palmquist, J., S. Uryasev and P. Krokhmal (2002), 'Portfolio optimization with conditional value-at-risk objective and constraints', *The Journal of Risk* **4, (2)**.

Rachev, S. and L. Rüschendorf (1998), *Mass transportation problems, V1, V2*, Springer-Verlag, NY.

Rachev, S., T. Jasik, A. Biglova and F. Fabozzi (2005), 'Risk and return in momentum strategies: Profitability from portfolios based on risk-adjusted stock ranking criteria', *Technical report, Department of probability and applied statistics, University of California, Santa Barbara.* .

Rachev, S.T. and S. Mittnik (2000), *Stable Paretian Models in Finance*, John Wiley & Sons, Series in Financial Economics.

Rockafellar, R. T. and S. Uryasev (2002), 'Conditional value-at-risk for general loss distributions', *Journal of Banking and Finance* **26, (7)**, 1443–1471.

Rockafellar, R. T., S. Uryasev and M. Zabarankin (2005), 'Master funds in portfolio analysis with general deviation measures', *The Journal of Banking and Finance* **29**.

Rockafellar, R. T., S. Uryasev and M. Zabarankin (2006), 'Generalized deviations in risk analysis', *Finance and Stochastics* **10**, 51–74.

Rockafellar, R. T. and W. Ziemba (2000), 'Modified risk measures and acceptance sets'.

Sharpe, W. F. (1966), 'Mutual funds performance', *Journal of Business* **January**, 119–138.

Sharpe, W. F. (1994), 'The sharpe ratio', *Journal of Protfolio Management* **Fall**, 45–58.

Yamai, Y. and T. Yoshiba (2002*a*), 'Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition and optimization', *Monetary and Economic Studies, Institute for Monetary and Economic Studies, Bank of Japan* **20, (1)**, 87–122.

Yamai, Y. and T. Yoshiba (2002*b*), 'On the validity of value-at-risk: comparative analyses with expected shortfall', *Monetary and Economic Studies, Institute for Monetary and Economic Studies, Bank of Japan* **20, (1)**, 57–86.

Ziemba, W. (1974), 'Choosing investment portfolios when the returns have stable distributions', *P. L. Hammer and G. Zoutendijk, Eds., Mathematical Programing in Theory and Practice, North-Holland* pp. 443–482.

Ziemba, W. (2005), 'The symmetric downside-risk sharpe ratio', *Journal of Portfolio Management,* **Fall**, 108–122.

Ziemba, W., C. Parkan and R. Brooks-Hill (1974), 'Calculation of investment portfolios with risk free borrowing and lending', *Management Science* **21, (2)**, 209–222.

Ziemba, W. and J. Board (1996), 'Mean-variance portfolio theory', *Handbook of Operations Research., S.T. Gass and C. Harris, Eds. Kluwer Publishing* pp. 498–503.

Ziemba, W. and R. Vickson (2006), *Stochastic optimization models in finance*, second edition, World Scientific, Singapore.