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, CA 93106, USA

Frank J. Fabozzi
Frederick Frank Adjunct Professor of Finance
Yale University, School of Management

February 21, 2005

Abstract

This article considers 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.

*Prof 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

There are two basic approaches to the problem of portfolio selection under uncertainty. 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 the 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 latter approach is that it cannot capture the richness of the former. As a matter of fact, the relationship between the two approaches is still a research topic (see Ogryczak, 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), which is based on the mean-variance analysis, some new performance measures like the STARR ratio, the Minimax measure, Sortino-Satchell ratio, Farinelli-Tibletti ratio and most recently the Rachev ratio and the Generalized Rachev ratio have been proposed (for an empirical comparison, see Biglova et. al. (2004), Rachev et. al. (2005) and the references therein). 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.

In this paper, we focus on general performance measure optimization. We continue 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, positive reward and risk measures respectively. 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 approach the recently proposed Rachev and Generalized Rachev ratio.

2 The mean-variance analysis and the Sharpe ratio

The classical mean-variance framework introduced by Markowitz in the 1950’s (Markowitz (1952)) is the first proposed model of the second 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 $r = (r_1, r_2, \ldots, r_n)^T$ is stochastic with expected value $E_r = (E_{r_1}, E_{r_2}, \ldots, E_{r_n})^T$. The result of the investment decision is a portfolio with composition $w = (w_1, w_2, \ldots, w_n)^T$ where $w_i$ is the portfolio weight corresponding to the $i$-th item, i.e. the share of the initial endowment invested in the $i$-th asset. We require that the weights of all portfolio items sum up to 1, $w^T e = \sum_{i=1}^n w_i = 1$ where $e = (1, 1, \ldots, 1) \in \mathbb{R}^n$. The expected portfolio return, expressed in terms of the individual items, equals
A key point in Markowitz's approach is that the standard deviation of portfolio return \( \sigma_p \) is assumed to be the measure of risk. If we denote \( \Sigma = \{ \text{cov}(r_i, r_j) \}_{i,j=1}^{n} \) to be the covariance matrix of the portfolio items, 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 can be summarized briefly in two ways:

1. From all feasible portfolios with a given upper bound on \( \sigma_p \), find the ones that have maximum expected return \( \mu_p \);
2. From all feasible portfolios with a given lower bound on \( \mu_p \), find the ones that have minimum risk \( \sigma_p \);

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

\[
\begin{align*}
\max_w & \quad w^T E r \\
\text{subject to} & \quad w^T e = 1 \\
& \quad w^T \Sigma w \leq R^* \\
& \quad Lb \leq Aw \leq Ub \\
\end{align*}
\]

and

\[
\begin{align*}
\min_w & \quad w^T \Sigma w \\
\text{subject to} & \quad w^T e = 1 \\
& \quad w^T E r \geq R^* \\
& \quad Lb \leq Aw \leq Ub \\
\end{align*}
\]

where \( R^* \) is the upper bound on portfolio risk, \( R^* \) is the lower bound on portfolio return, \( A \in \mathbb{R}^{n \times k} \) 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. As a matter of fact, the originally proposed problem by Markowitz in his seminal work is Problem (2) and it is known as the Markowitz problem\(^1\). The optimal

\(^1\)For the general duality theorems of the type of Problems (1) and (2), see Rachev, Rüschendorf (1998)
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 \( (w^{oT}E_r, w^{oT}\Sigma w^o) \) where \( w^o \in \mathcal{E}_o \) is called the Efficient frontier.

We need to remark that there is a third way to arrive at the mean-variance efficient set. It is by considering the optimization problem:

\[
\max_{w} \quad w^{T}E_r - \lambda w^{T}\Sigma w
\]
\[
\text{subject to} \quad w^{T}e = 1
\]
\[
Lb \leq Aw \leq Ub
\]

where \( \lambda > 0 \) is a parameter. In this representation, the objective function \( w^{T}E_r - \lambda w^{T}\Sigma w \) is interpreted as a utility function and \( \lambda \) is called the risk-aversion parameter. Since it is possible to show that the three problems are equivalent (see, for example, Rockafellar, Uryasev (2002) and Palmquist, Uryasev (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

\[
SR(w^o) = \frac{w^{oT}E_r}{\sqrt{w^{oT}\Sigma w^o}}
\]

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

\[
\max_{w \in \mathcal{E}_o} \quad \frac{w^{T}E_r}{\sqrt{w^{T}\Sigma w}}
\]

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 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).

\[\text{The optimal portfolio is known as the Markowitz market portfolio, or can also be called tangent portfolio with zero risk-free rate.}\]
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 correct 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. 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 potential loss as well as potential profit.

An example of a class of risk measures recently 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: $X, Y \in V; Y \geq X \implies \rho(Y) \leq \rho(X)$
2. sub-additive: $X, Y, X + Y \in V \implies \rho(X + Y) \leq \rho(Y) + \rho(X)$
3. positively homogeneous: $X \in V, h > 0, hX \in V \implies \rho(hX) = h\rho(X)$
4. translation invariant: $X \in V, a \in \mathbb{R} \implies \rho(X + a) = \rho(X) - a$

These axioms are desirable for the purpose of risk management.

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:

![Figure 1: The Efficient frontier and the tangent portfolio](image)
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

\[
\max_{w \in \mathcal{E}_o} \frac{w^T Er}{\rho(w^T r)} \tag{7}
\]

The solution of Problem (7) will be called the $\rho$-market portfolio or $\rho$-tangent portfolio with zero risk-free rate. Note that this portfolio is not necessarily unique. Actually, it is easy to show that the following simple result holds.

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

\[
\max_{w} \frac{w^T Er}{\rho(w^T r)} \text{ subject to } w^T e = 1, Lb \leq Aw \leq Ub \tag{8}
\]

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

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

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

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

\[
\frac{\bar{w}^o T Er}{\rho(\bar{w}^o T Er)} \leq \frac{\bar{w}^T Er}{\rho(\bar{w}^T Er)} \tag{10}
\]

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

\[
\max_{w \in X} \frac{w^T Er}{\rho(w^T Er)} = \frac{\bar{w}^o T Er}{\rho(\bar{w}^o T Er)} \leq \frac{\bar{w}^T Er}{\rho(\bar{w}^T Er)} \leq \max_{w \in \mathcal{E}_o} \frac{w^T Er}{\rho(w^T r)}
\]

The equality holds because $\bar{w}$ 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 $\bar{w}^o$. 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.

It appears that we can find the \( \rho \)-tangent portfolio by considering directly Problem (8) and maximize the corresponding performance measure without the need to find first the set \( \mathcal{E}_o \).

### 3.1 General reward-risk ratio optimization

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

\[
\begin{align*}
\max_w & \quad \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \\
\text{subject to} & \quad w^T r = 1 \\
& \quad Lb \leq Aw \leq Ub
\end{align*}
\]

where \( \mu(\cdot) \) is a general non-linear reward measure, \( \rho(\cdot) \) is a risk measure and \( r_b \) 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. Our analysis will be based on (quasi-) convex functions, their optimal properties and extensions of some techniques for programming with fractional objectives (see Charnes, Cooper (1962)). We briefly recall that a function \( f : D \subset \mathbb{R}^n \rightarrow \mathbb{R} \) 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 \( f : D \subset \mathbb{R}^n \rightarrow \mathbb{R} \) 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.

First we establish the next simple

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

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

b) the ratio \( \rho/\mu : D_1 \cap D_2 \rightarrow \mathbb{R}^{++} \) 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. a) Consider the set \( L_t = \{ x : \mu(x)/\rho(x) \geq t \} = \{ x : \mu(x) - t\rho(x) \geq 0 \} = \{ x : -\mu(x) + t\rho(x) \leq 0 \} \) where the first equality holds because the numerator and the denominator are strictly positive. If \( t < t_{\text{min}} = \min_{x \in D_1 \cap D_2} \mu(x)/\rho(x) \), then \( L_t = D_1 \cap D_2 \) because all admissible \( x \) satisfy the condition in \( L_t \). Therefore \( L_t \) in this case is convex because both \( D_1 \) and \( D_2 \) are convex. If \( t_{\text{min}} \leq t \leq t_{\text{max}} = \max_{x \in D_1 \cap D_2} \mu(x)/\rho(x) \), \( L_t \) is convex because it is a level set of the convex function \(-\mu(x) + t\rho(x)\). If \( t > t_{\text{max}} \), then \( L_t = \emptyset \). It follows that \( \mu(x)/\rho(x) \) is quasi-concave.

b) Consider the set \( L_t = \{ x : \rho(x)/\mu(x) \leq t \} = \{ x : \rho(x) - t\mu(x) \leq 0 \} \). If \( t < t_{\text{min}} = \min_{x \in D_1 \cap D_2} \rho(x)/\mu(x) \), then \( L_t = \emptyset \). If \( t_{\text{min}} \leq t \), \( L_t \) is convex following the same reasoning as in a). Therefore \( \rho(x)/\mu(x) \) is quasi-convex.

c) Let \( r(x) = \rho(x)/\mu(x) \). Suppose that \( x^o = \arg \min_x r(x), x \in D_1 \cap D_2 \). 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 \( x^o = \arg \max_x r^{-1}(x), x \in D_1 \cap D_2 \). The converse is also true.

It should be remarked that 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, 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.

4 Reward-risk ratio of the form \( \frac{\mu(w^Tr)}{\rho(w^Tr)} \)

In this section we consider performance measures that have the general form \( \mu(w^Tr)/\rho(w^Tr) \) where \( \mu(\cdot) \) is a reward measure and \( \rho(\cdot) \) is a risk measure. The reward measure is assumed to be a positive 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 positive 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 $(w_1, w_2, \ldots, w_n)$ 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.

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

**Proof.** We shall show that $\mu$ is concave, the same approach works for the risk function. Let $f(w) = \mu(w^T r - r_b)$ and $w_1, w_2 \in X$. Then

$$f(\alpha w_1 + (1 - \alpha)w_2) = \mu(\alpha w_1^T r + (1 - \alpha)w_2^T r - r_b) = \mu(\alpha (w_1^T r - r_b) + (1 - \alpha)(w_2^T r - r_b))$$
$$\geq \alpha \mu(w_1^T r - r_b) + (1 - \alpha)\mu(w_2^T r - r_b) = \alpha f(w_1) + (1 - \alpha)f(w_2)$$

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

\[
\begin{align*}
\max_{\{x,t\}} & \quad \mu(x^T r - tr_b) \\
\text{subject to} & \quad \rho(x^T r - tr_b) \leq 1 \\
& \quad x^T e = t \\
& \quad t Lb \leq Ax \leq t Ub \\
& \quad t \geq 0
\end{align*}
\]  
\begin{align*}
\min_{\{x,t\}} & \quad \rho(x^T r - tr_b) \\
\text{subject to} & \quad \mu(x^T r - tr_b) \geq 1 \\
& \quad x^T e = t \\
& \quad t Lb \leq Ax \leq t Ub \\
& \quad t \geq 0
\end{align*}
\]
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, \(\mu(x_A^o T r - t_A^o r_b) = \mu(w^o T r - r_b) / \rho(w^o T r - r_b) = \rho^{-1}(x_B^o T r - t_B^o r_b)\). Conversely, if \(w^o\) is a solution to Problem (11) and \(t^o = \rho^{-1}(w^o T r - r_b)\), then the pair \((t^o w^o, t^o)\) is a solution to Problem (A). If \(t^o = \mu^{-1}(w^o T r - r_b)\), 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. 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 \(X\) 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 \(X \cap \{w : \mu(w^T r - r_b) \geq \epsilon > 0\}\) which is a convex set and does not change the nature of the problem. The parameter \(\epsilon\) 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 \(w \in X\) such that \(\mu(w^T r - r_b) > 0\).

Another important assumption is that the risk measure is a positive functional. If \(\rho(\cdot)\) is not bounded away from zero for all feasible portfolios, i.e. there exists \(w \in X\) 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 \(t = \rho^{-1}(w^T r - r_b)\) (for more details see the proof in the Appendix), which makes the problem unbounded. The solution of Problem B will be a portfolio \(w \in X\) such that \(\rho(w^T r - r_b) = 0\) because for the other portfolios the objective is strictly positive. As a result, we can conclude that strict positivity for the risk measure is crucial for the stated equivalence between the three problems.

5 Reward-risk ratio of the form \(\frac{w^T E r}{\rho(w^T r)}\)

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

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 (8) is equivalent to the following two problems

---

3Actually 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 De Giorgi (2004).
in the sense that if the pair \((x^a, t^a)\) is a solution to Problem (A1) and \((x^b, t^b)\) is a solution to Problem (B1), then \(w^\alpha = x^a / t^a = x^b / t^b\) solves Problem (II). Moreover \(x^a T r - t^a E r b = (w^{\alpha T} r - E r b) / \rho(w^{\alpha T} r - r_b) = \rho^{-1}(x^a T r - t^a E r b)\). Conversely, if \(w^\alpha\) is a solution to Problem (II) and \(t^\alpha = \rho^{-1}(w^{\alpha T} r - r_b)\), then the pair \((t^\alpha w^\alpha, t^\alpha)\) is a solution to Problem (A1). If \(t^\alpha = (w^{\alpha T} r - E r b)^{-1}\), then \((t^\alpha w^\alpha, t^\alpha)\) is a solution to Problem (B1).

\[ \begin{align*}
\max_{(x,t)} & \quad x^T E r - t E r b \\
\text{subject to} & \quad \rho(x^T r - t r_b) \leq 1 \\
& \quad x^T e = t \\
& \quad t L b \leq Ax \leq t U b \\
& \quad t \geq 0 \\
\end{align*} \]

(\text{A1})

\[ \begin{align*}
\min_{(x,t)} & \quad \rho(x^T r - t r_b) \\
\text{subject to} & \quad x^T E r - t E r b = 1 \\
& \quad x^T e = t \\
& \quad t L b \leq Ax \leq t U b \\
& \quad t \geq 0 \\
\end{align*} \]

(\text{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 the 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 risk measure is the standard deviation of portfolio returns. If \( r_b \) is non-zero, then the variance of the excess return is

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

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

\[
D(w^T r - r_b) = (-1, w)^T \Sigma_1 (-1, w), \quad \Sigma_1 = \begin{pmatrix} \sigma_b^2 & \sigma_{br} \\ \sigma_{br}^T & \Sigma \end{pmatrix},
\]

\( \sigma_b^2 \) is the variance of the benchmark asset \( r_b \), \( \sigma_{br} = (\text{cov}(r_b, r_1), \text{cov}(r_b, r_2), \ldots, \text{cov}(r_b, r_n)) \) is the vector of covariances of the returns of the benchmark asset with the returns of the other assets \( r = (r_1, r_2, \ldots, r_n) \) and \( \Sigma = \{\text{cov}(r_i, r_j)\}_{i,j=1}^{n} \) is the covariance matrix of the assets’ returns. The notation \((-1, w)^T\) stands for the vector row \((-1, w_1, \ldots, w_n)^T\) and \((-1, w)\) is the corresponding vector column.

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

\[
\begin{align*}
\max_{w} \quad & \frac{w^T E r - E r_b}{\sqrt{(-1, w)^T \Sigma_1 (-1, w)}} \\
\text{subject to} \quad & w^T e = 1 \\
& Lb \leq Aw \leq Ub
\end{align*}
\]
The result is contained in the

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

\[
\begin{align*}
\max_{(x,t)} & \quad x^T Er - tEr_b \\
\text{subject to} & \quad (-t, x)^T \Sigma_1 (-t, x) \leq 1 \\
& \quad x^T e = t \\
& \quad tLb \leq Ax \leq tUb \\
& \quad t \geq 0
\end{align*}
\]

(SR A)

and

\[
\begin{align*}
\min_{(x,t)} & \quad (-t, x)^T \Sigma_1 (-t, x) \\
\text{subject to} & \quad x^T Er - tEr_b = 1 \\
& \quad x^T e = t \\
& \quad tLb \leq Ax \leq tUb \\
& \quad t \geq 0
\end{align*}
\]

(SR B)

in the sense that if the pair \((x^o_A, t^o_A)\) is a solution to Problem (SR A) and \((x^o_B, t^o_B)\) is a solution to Problem (SR B), then \(w^o = x^o_A/t^o_A = x^o_B/t^o_B\) solves Problem (12). Moreover \(x^o_A^T Er - t^o_A Er_b = (w^o)^T Er - Er_b)/\rho(w^o^T r - r_b) = \rho^{-1}(x^o_B^T r - t^o_B r_b)\) where \(\rho(z) = \sqrt{z^T \Sigma z}\). Conversely, if \(w^o\) is a solution to Problem (12) and \(t^o = \rho^{-1}(w^o^T r - r_b)\), then the pair \((t^o w^o, t^o)\) is a solution to Problem (SR A). If \(t^o = (w^o^T Er - Er_b)^{-1}\), then \((t^o w^o, t^o)\) is a solution to Problem (SR B).

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

In the problem of type B in Proposition 5, we have as objective

\[
\sqrt{z^T \Sigma z} \to \min_{(x,t)}
\]

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

Problem (SR B) is quadratic and can be solved as a quadratic programming problem. Note that the reciprocal of the objective function value obtained at the 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 at the solution equals
In the case of the Markowitz problem, the risk measure is strictly convex if the covariance matrix is non-singular and therefore Figure 1 presents the general case for the view of the graph of the Efficient frontier.

\[((-t_B, x_B)^T \Sigma_1 (-t_B, x_B))^{-1} = ((-1, w_B)^T \Sigma_1 (-1, w_B))^{-1} = \left(\frac{(w_B^T \mu - E r_b)^2}{(w_B^T \mu - E r_b)^2}\right)^2\]

In the case of the Markowitz problem, the risk measure is strictly convex if the covariance matrix is non-singular and therefore 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 conditional value-at-risk (CVaR), is proposed in the literature as a coherent measure of risk, a superior alternative to the industry standard Value-at-Risk (VaR). For a discussion, see Yamai, Yoshiba (2002). The ETL of portfolio returns at the 100(1 - \alpha) percent confidence level is defined as

\[ETL_\alpha(w^T r) = E(-w^T r - w^T r > VaR_\alpha(w^T r))\]  \hspace{1cm} (13)

where \(VaR_\alpha(w^T r)\) is implicitly defined as \(\mathbb{P}(w^T r < -VaR_\alpha(w^T r)) = \alpha\) and is the Value-at-Risk measure. 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

\[STARR(w) = \frac{w^T \mu - E r_b}{ETL_\alpha(w^T r - r_b)}\]  \hspace{1cm} (14)

and is proposed in Martin, Rachev, Siboulet (2003). Since the ETL 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 \(\rho(w^T r - r_b) = ETL_\alpha(w^T r - r_b)\). For both problems of type A and B, we can apply the linearization technique developed in Rockafellar, Uryasev (2002) when there are scenarios for portfolio items returns.

Suppose that we have \(N\) vectors \(r^k \in \mathbb{R}^n\), \(k = 1, 2, \ldots, N\) 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 \(s_k = w^T r^k\), \(k = 1, 2, \ldots, N\) where \(w\) is a concrete vector of weights. Using the scenarios, we can compute an estimate \(\bar{ETL}\) of portfolio ETL:

\[\bar{ETL}_\alpha(w^T r) = \frac{1}{N_\alpha}\sum_{k=1}^{[N_\alpha]} (-1)^{s(k)}\]  \hspace{1cm} (15)
where \( s(k), k = 1, 2, \ldots, N \) denote the sorted portfolio scenarios in increasing order and \( [a] \) is the largest integer smaller than \( a \). It is shown in Rockafellar, Uryasev (2002) that the same estimate can be obtained via minimization

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

(16)

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

\[
\max_{(x, t, d, \theta)} \quad x^T Er - t Er_b \\
\text{subject to} \quad \theta + \frac{1}{N\alpha} \sum_{k=1}^{N} d_k \leq 1 \\
\text{(STARR A)} \\
-\frac{x^T r^k + tr^b \leq \theta \leq d_k, \quad k = 1, 2, \ldots, N}{x^T e = t} \\
tLb \leq Ax \leq tUb \\
t \geq 0, \quad d_k \geq 0, \quad k = 1, 2 \ldots N
\]

and

\[
\min_{(x, t, d, \theta)} \quad \theta + \frac{1}{N\alpha} \sum_{k=1}^{N} d_k \\
\text{subject to} \quad x^T Er - t Er_b = 1 \\
\text{(STARR B)} \\
-\frac{x^T r^k + tr^b \leq \theta \leq d_k, \quad k = 1, 2, \ldots, N}{x^T e = t} \\
tLb \leq Ax \leq tUb \\
t \geq 0, \quad d_k \geq 0, \quad k = 1, 2 \ldots N
\]

In both problems \( r^k_b, k = 1, 2, \ldots, N \) denote the scenarios for the benchmark portfolio, \( d = (d_1, d_2, \ldots, d_N) \) 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. Due to the translation invariance property, the ETL measure can violate this assumption. The ETL of a portfolio can be represented as

\[
ETL_{\alpha}(w^T r) = ETL_{\alpha}(w^T (r - \mu_p) + \mu_p) = ETL_{\alpha}(w^T (r - \mu_p)) - \mu_p
\]

In this expression, \( ETL_{\alpha}(w^T (r - \mu_p)) \) is the ETL of the centralized portfolio returns. The defining equation (13) implies the inequality

\[
ETL_{\alpha}(w^T r) \leq ETL_{\beta}(w^T r), \quad \beta < \alpha
\]

meaning that we can always choose \( \alpha \) to be sufficiently high, such that \( ETL_{\alpha}(w^T r) = 0 \) because \( \mu_p \) does not depend on \( \alpha \). Geometrically increasing \( \alpha \) continuously will shift the
expected Portfolio Return

Portfolio Risk, ETL

The tangent portfolios

The Efficient frontier changes as the parameter \( \alpha \) of the ETL increases, \( \alpha_1 < \alpha_2 < \alpha_3 < \alpha_4; r_b = 0 \)

The entire Efficient frontier to the left until the "tangent" line becomes vertical, see Figure 3.

We need to remark that 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 ETL 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) and Acerbi, Simonetti (2002).

7 Reward-risk ratio of the form \( \frac{\mu(w^T r)}{\rho(w^T r)} \) 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 positive and convex functional and \( \mu(\cdot) \) — a positive and concave functional. 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 problems from 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, Vandenberghe (2004). According to Proposition 2, the ratio optimization Problem (11) can be equivalently re-stated.
The objective function in the above problem is quasi-convex. Now consider the following feasibility problem

\[
\min_{w} \frac{\rho(w^T r - r_b)}{\mu(w^T r - r_b)} \quad \text{subject to} \quad \begin{align*}
\rho(w^T r - r_b) - t\mu(w^T r - r_b) &\leq 0 \\
w^T e &= 1 \\
Lb &\leq Aw \leq Ub
\end{align*}
\]  

(18)

where \(t\) is a fixed positive. For every fixed \(t \geq t_{\text{min}} = \min_{w \in X} \rho(w^T r - r_b)/\mu(w^T r - r_b)\), the set \(\{w : \rho(w^T r - r_b) - t\mu(w^T r - r_b) \leq 0\}\) is convex and therefore we have a convex feasibility problem. Also for every \(t \geq t_{\text{min}}\), Problem (18) 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 \(t_{\text{min}} \in [a, b]\) and we have a tolerance level \(\epsilon\). Then

1. Set \(t = (a + b)/2\).
2. Solve the feasibility Problem 18.
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 \(t_{\text{min}}\) since \(t_{\text{min}} \in [a, b]\) 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, Vandenberghe (2004) for a general description the quasi-convex problem.

In the first section, we related the Sharpe ratio to the Markowitz problem. Next we related the ratio of the form \(w^T E r/\rho(w^T r)\) 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

\[
\max_{w} \mu(w^T r - r_b) \quad \text{subject to} \quad \begin{align*}
\rho(w^T r - r_b) &\leq R^* \\
w^T e &= 1 \\
Lb &\leq Aw \leq Ub
\end{align*}
\]  

(19)
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 the next

**Proposition 9.** The Efficient frontier generated by Problem (19) — in which \( \mu(\cdot) \) is a positive, concave reward measure and \( \rho(\cdot) \) is a positive, 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

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

and

\[
\mu\left(\left(w^o\left(R_\lambda^*\right)\right)^T r - r_b\right) \geq \mu\left(\left(w^o\left(R_2^*\right)\right)^T r - r_b\right)
\]

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. On the basis of the above 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

\[
RR(w) = \frac{ETL_\alpha(r_b - w^T r)}{ETL_\beta(w^T r - r_b)}
\]  

(20)

The reward measure is the ETL 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

\[
\max_w \frac{ETL_\alpha(r_b - w^T r)}{ETL_\beta(w^T r - r_b)}
\]

subject to

\[
w^T e = 1
\]

\[
Lb \leq Aw \leq Ub
\]

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

**Proposition 10.** Problem (21) is equivalent to the following two optimization problems

**\( \text{RR A} \)**

\[
\max_{(x,t)} \quad \text{ETL}_\alpha(tr_b - x^T r)
\]

subject to

\[
\begin{align*}
\text{ETL}_\beta(x^T r - tr_b) &\leq 1 \\
x^T e &= t \\
tLb &\leq Ax \leq tUb \\
t &\geq 0
\end{align*}
\]

and

**\( \text{RR B} \)**

\[
\min_{(x,t)} \quad \text{ETL}_\beta(x^T r - tr_b)
\]

subject to

\[
\begin{align*}
\text{ETL}_\alpha(tr_b - x^T r) &\geq 1 \\
x^T e &= t \\
tLb &\leq Ax \leq tUb \\
t &\geq 0
\end{align*}
\]

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 (21). Moreover

\[
\text{ETL}_\alpha(t^o_A r_b - x^o_A^T r) = \text{ETL}_\alpha(r_b - w^o T r)/\text{ETL}_\beta(w^o T r - r_b) = (ETL_\beta(x^o_B^T r - t^o_B r_b))^{-1}
\]

Conversely, if \(w^o\) is a solution to Problem (21) and \(t^o = (ETL_\beta(x^o_B^T r - r_b))^{-1}\), then the pair \((t^o w^o, t^o)\) is a solution to Problem (RR A). If \(t^o = (ETL_\alpha(r_b - w^o T r))^{-1}\), 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 \( \{(x, t) : ETL_\delta(x^T r - tr_b) \leq 1\} \) is convex. In the objective we have maximization of a convex function and for this reason the problem is non-convex. In Problem (RR B), in the objective there is minimization of a convex function but the set \( \{(x, t) : ETL_\alpha(tr_b - x^T r) \geq 1\} \) is non-convex and so is the entire problem.

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

\[
\max_{(x, t, \delta, \theta, d)} \quad \delta + \frac{1}{[N_{\alpha}]} \sum_{k=1}^{N} g_k \\
\text{subject to} \quad x^T r^k - \delta \leq g_k, \quad k = 1, 2, \ldots, N \\
g_k \geq 0, \quad k = 1, 2, \ldots, N \\
\theta + \frac{1}{[N_{\alpha}]} \sum_{k=1}^{N} d_k \leq 1 \\
x^T r^k - \theta \leq d_k, \quad k = 1, 2, \ldots, N \\
d_k \geq 0, \quad k = 1, 2, \ldots, N \\
x^T e = t \\
tLb \leq Ax \leq tUb \\
t \geq 0
\]

(22)

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 \( x^T r^k - \delta \leq g_k \) and therefore all \( g_k \) increase indefinitely. There is no such a restriction for \( \theta \) as well and it also increases indefinitely. Now suppose that we try to “fix” the problem by changing the constraints to be \( x^T r^k - \delta \geq g_k \). The new problem could turn out infeasible if for one \( k \), \( 0 > x^T r^k - \delta \geq g_k \) because we require all \( g_k \) to be positive.

We propose a linearization introducing binary variables in the optimization. Thus the resulting problem is a mixed-integer linear programming (MIP) problem. The linearization is based on equation (15) and on the consideration that if the ETL estimate is computed as the average of the largest \([N_{\alpha}]\) observations in the sample multiplied by (-1), then the average of any other \([N_{\alpha}]\) observations will be smaller than the ETL. The above statement transforms into the optimization problem
\[
\max_{(g,\lambda)} \frac{1}{|N\alpha|} \sum_{k=1}^{N} g_k
\]
subject to
\[
\begin{align*}
g_k &\leq B\lambda_k, \quad k = 1, 2, \ldots, N \\
g_k &\geq -w^T r^k - B(1 - \lambda_k), \quad k = 1, 2, \ldots, N \\
g_k &\leq -w^T r^k + B(1 - \lambda_k), \quad k = 1, 2, \ldots, N \\
\lambda^T e &= \lfloor N\alpha \rfloor \\
\lambda_k &\in \{0, 1\}, \quad g_k \geq 0, \quad k = 1, 2, \ldots, N
\end{align*}
\]  

where \( r^k \in \mathbb{R}^n \), \( k = 1, 2, \ldots, N \) are vectors of assets returns scenarios, \( \alpha \) is the parameter of the ETL function, \( B \) is a very large number, such that \( |w^T r^k| \leq B \) for all scenarios and \( \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_N) \) is a vector of binary variables. Let us verify if Problem (23) is equivalent to the above statement. For this reason, we have to examine two cases:

1. Suppose that \( \lambda_k = 0 \). Then

\[
\begin{align*}
g_k &\leq 0 \\
g_k &\geq -w^T r^k - B \\
g_k &\leq -w^T r^k + B \\
g_k &\geq 0
\end{align*}
\]

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 \( \lambda_k = 1 \).

\[
\begin{align*}
g_k &\leq B \\
g_k &\geq -w^T r^k \\
g_k &\leq -w^T r^k \\
g_k &\geq 0
\end{align*}
\]

From the last three equalities it follows that \( g_k = -w^T r^k \geq 0 \). The first is redundant.

Thus the binary variable \( \lambda_k \) indicates when 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 (-1). 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 \( \lambda_k \), is controlled by the constraint \( \lambda^T e = \lfloor N\alpha \rfloor \) according to which the sum of all \( \lambda_k \) should equal \( \lfloor N\lambda \rfloor \). Certainly \( \mathcal{ETL}_\alpha(w^T r) \) matches the solution of Problem (23).
We have demonstrated that for each $w$ fixed, solving Problem (23) we obtain $\text{ETL}_\alpha(w^T r)$. We can combine Problem (RR A) and Problem (23) as the objective of the former is also maximization. As a result we have

$$\max_{(x,t,g,\lambda,d,\theta)} \frac{1}{|N\alpha|} \sum_{k=1}^{N} g_k$$

subject to

$$g_k \leq B\lambda_k, \quad k = 1, 2, \ldots, N$$
$$g_k \geq x^T r_k - B(1 - \lambda_k), \quad k = 1, 2, \ldots, N$$
$$g_k \leq x^T r_k + B(1 - \lambda_k), \quad k = 1, 2, \ldots, N$$
$$\lambda^T e = |N\alpha|$$
$$\lambda_k \in \{0, 1\}, \quad g_k \geq 0, \quad k = 1, 2, \ldots, N$$
$$\theta + \frac{1}{|N\alpha|} \sum_{k=1}^{N} d_k \leq 1$$
$$-x^T r_k - \theta \leq d_k, \quad k = 1, 2, \ldots, N$$
$$d_k \geq 0, \quad k = 1, 2 \ldots N$$
$$x^T e = t$$
$$t Lb \leq Ax \leq t Ub$$
$$t \geq 0$$

Solving Problem (24) we obtain the global maximum of the Rachev ratio. The number of binary variables $\lambda_k$ 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 ETL linearization is applicable in the STARR ratio ratio optimization problem and in general, in the optimal portfolio problem with the ETL as a risk function, but the method of Rockafellar, Uryasev (2002) is more efficient as it involves continuous variables only.

### 8.2 The Generalized Rachev ratio

The Generalized Rachev ratio is defined as

$$GRR(w) = \frac{\text{ETL}_{[\gamma, \alpha]}(r_b - w^T r)}{\text{ETL}_{[\delta, \beta]}(w^T r - r_b)}$$

where $\text{ETL}_{[\gamma, \alpha]}(X) = E((\max(-X, 0))^\gamma) - X > VaR_\alpha(X)$ and $VaR_\alpha(X)$ is the Value-at-Risk measure. 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 $\text{ETL}_{[\gamma, \alpha]}(r_b - w^T r)$ : $X \subset \mathbb{R}^n \rightarrow \mathbb{R}^{+\cdot}$ is convex for $\gamma \geq 1$, provided that $X$ 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 $ETL_{(\gamma, \alpha)}(X)$ as defined in equation (25) is positive homogeneous of degree $\gamma$, that is

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

where $t > 0$.

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

$$\max_w \frac{ETL_{(\gamma, \alpha)}(r_b - w^T r)}{ETL_{(\delta, \beta)}(w^T r - r_b)}$$

subject to

$$w^T e = 1$$

$$Lb \leq Aw \leq Ub$$

(26)

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 (26) with $\gamma = \delta$ is equivalent to the following two optimization problems

$$\max_{(x, t)} \frac{ETL_{(\delta, \alpha)}(tr_b - x^T r)}{ETL_{(\delta, \beta)}(x^T r - tr_b)}$$

subject to

$$x^T e = t$$

$$tLb \leq Ax \leq tUb$$

$$t \geq 0$$

(GRR A)

and

$$\min_{(x, t)} \frac{ETL_{(\delta, \alpha)}(tr_b - x^T r)}{ETL_{(\delta, \beta)}(x^T r - tr_b)}$$

subject to

$$x^T e = t$$

$$tLb \leq Ax \leq tUb$$

$$t \geq 0$$

(GRR B)

in the sense that if the pair $(x^*_A, t^*_A)$ is a solution to Problem (GRR A) and $(x^*_B, t^*_B)$ is a solution to Problem (GRR B), then $w^* = x^*_A/t^*_A = x^*_B/t^*_B$ solves Problem (21). Moreover

$$ETL_{(\delta, \alpha)}(t^*_A r_b - x^*_A^T r) = ETL_{(\delta, \alpha)}(r_b - w^* e^T r)/ETL_{(\delta, \beta)}(w^* e^T r - r_b) = (ETL_{(\delta, \beta)}(x^* e^T r - t^*_B r_b))^{-1}$$. Conversely, if $w^*$ is a solution to Problem (21) and $t^* = (ETL_{(\delta, \beta)}(w^* e^T r - r_b))^{-1}$, then the pair $((t^*)^{1/\delta} w^*, (t^*)^{1/\delta})$ is a solution to Problem (GRR A). If $t^* = (ETL_{(\delta, \alpha)}(r_b - w^* e^T r))^{-1}$, then $((t^*)^{1/\delta} w^*, (t^*)^{1/\delta})$ 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 (25) as

$$MGRR(w) = \left( \frac{ETL_{(\gamma, \alpha)}(r_b - w^T r)}{ETL_{(\delta, \beta)}(w^T r - r_b)} \right)^{1/\delta}$$ (27)

In this modification, it is clear from Proposition 12 that 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 (26) with equation (27) as objective is equivalent to the following two optimization problems

\[
\begin{align*}
\text{max}_{(x,t)} & \quad ETL_{(\gamma, \alpha)}(tr_b - x^T r) \\
\text{subject to} & \quad ETL_{(\delta, \beta)}(x^T r - tr_b) \leq 1 \\
& \quad x^T e = t \\
& \quad tl_b \leq Ax \leq tU_b \\
& \quad t \geq 0
\end{align*}
\]  
\( (MGRR A) \)

and

\[
\begin{align*}
\text{min}_{(x,t)} & \quad ETL_{(\delta, \beta)}(x^T r - tr_b) \\
\text{subject to} & \quad ETL_{(\gamma, \alpha)}(tr_b - x^T r) \geq 1 \\
& \quad x^T e = t \\
& \quad tl_b \leq Ax \leq tU_b \\
& \quad t \geq 0
\end{align*}
\]  
\( (MGRR B) \)

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

$$ETL_{(\gamma, \alpha)}(tr_b - x^T r) \rightarrow \text{max}_{(x,t)}$$

\( (26 \text{a}) \)
and risk constraint $(ETL_{(\delta,\theta)}(x^T r - tr_b))^{1/\delta} \leq 1$. Raising the objective to the power $\gamma$ does not change the solution point as the transformation is strictly increasing. Raising the inequality to power $\delta$ does not change the feasibility either. Thus we receive Problem (MGRR A).

It should be remarked that 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, Tibletti (2003) and Biglova et. al. (2004)):

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

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

9 Conclusion

In this article 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 linearizable, 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 some of the Propositions.

• Proof of Proposition 4.

Proof: First we show that Problem (11) is equivalent to Problem (A). We make the proof in two steps. As a first step, we show that Problem (11) is equivalent to

\[
\begin{align*}
\max_{(x,t)} & \quad \mu(x^T r - tr_b) \\
\text{subject to} & \quad \rho(x^T r - tr_b) = 1 \\
& \quad x^Te = t \\
& \quad tLb \leq Ax \leq tUb \\
& \quad t \geq 0
\end{align*}
\]

(28)

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 (28). Thus

\[
\max_{w \in X} \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \leq \max_{(x,t) \in X_1} \mu(x^T r - tr_b)
\]

where \( X \) denotes the set of feasible portfolios in Problem (11) and \( X_1 \) denotes the set of feasible \((x,t)\) pairs in Problem (28). Conversely, if \((x,t)\) is feasible in Problem (28), then \( w = x/t \) is feasible in Problem (11) and

\[
\max_{w \in X} \frac{\mu(w^T r - r_b)}{\rho(w^T r - r_b)} \geq \max_{(x,t) \in X_1} \mu(x^T r - tr_b)
\]

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 (28) and if \((x,t)\) is a solution to Problem (28) then \( w = x/t \) is a solution to Problem (11).

As a second step, let us consider Problem (28) and Problem (A) where the risk function constraint has been relaxed. We claim that if \((x^o, t^o)\) is a solution to Problem (28), then \( \rho(x^{oT} r - t^o r_b) = 1 \) and therefore the pair is a solution to Problem (28). 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

\[
\begin{align*}
\max_{(w,t)} & \quad t\mu(w^T r - r_b) \\
\text{subject to} & \quad t\rho(w^T r - r_b) \leq 1 \\
& \quad tw^Te = t \\
& \quad tLb \leq Aw \leq tUb \\
& \quad t \geq 0
\end{align*}
\]

Obviously multiplying \( t \) by any constant \( \alpha \) such that \( 1 < \alpha \leq \rho^{-1}(w^{oT} r - r_b) \) does not change the feasibility of \((\alpha t^o w^o, \alpha t^o)\) and
since according to the assumed properties of \( \mu(\cdot) \), it is strictly positive. Therefore we arrive at a contradiction that \( (x^0, t^0) \) 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:

\[
\begin{align*}
\min_{w} & \quad \frac{\rho(w^T r - r_b)}{\mu(w^T r - r_b)} \\
\text{subject to} & \quad w^T e = a \\
& \quad Lb \leq Aw \leq Ub
\end{align*}
\]

and

\[
\begin{align*}
\min_{(x,t)} & \quad \rho(x^T r - t r_b) \\
\text{subject to} & \quad \mu(x^T r - t r_b) = 1 \\
& \quad x^T e = t \\
& \quad t Lb \leq Ax \leq t Ub \\
& \quad t \geq 0
\end{align*}
\]

We arrive at Problem (B) by changing the reward constraint to \( \mu(x^T r - t r_b) \geq 1 \). 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 \( (x^\lambda, t^\lambda) \) satisfies all constraints in the problem. We show how to deal with all of them:

- the first one: \( x^T \lambda E r - t \lambda E r_b = 1 \).

\[
x^T \lambda E r - t \lambda E r_b = \lambda(x^T \lambda E r - t^1 \lambda E r_b) + (1 - \lambda)(x^T \lambda E r - t^2 \lambda E r_b) = 1
\]

The first equality follows since \( \lambda \) and \( t^\lambda \) 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: \( x^T e = t \).

Since both points \( (x^1, t^1) \) and \( (x^2, t^2) \) are feasible, then \( x^1 T e = t^1 \) and \( x^2 T e = t^2 \). Therefore

\[
\lambda x^1 T e + (1 - \lambda)x^2 T e = \lambda t^1 + (1 - \lambda)t^2
\]

and it follows that \( x^T e = t^\lambda \).

- the third one: \( t Lb \leq Ax \leq t Ub \).

Since both points \( (x^1, t^1) \) and \( (x^2, t^2) \) are feasible, then \( t^1 Lb \leq Ax^1 \leq t^1 Ub \) and \( t^2 Lb \leq Ax^2 \leq t^2 Ub \). Multiplying by \( \lambda \) and \( (1 - \lambda) \) respectively and summing parts by parts we obtain
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

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

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

\[ \rho(x^{\lambda T}r - t^\lambda r_b) = \rho(\lambda(x^{1T}r - t^1r_b) + (1-\lambda)(x^{2T}r - t^2r_b)) \]
\[ \leq \lambda \rho(x^{1T}r - t^1r_b) + (1-\lambda)\rho(x^{2T}r - t^2r_b) = \rho_{\min} \]

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. \( \Box \)

**Proof of Proposition 9.**

*Proof.* The monotonic property is easiest to establish. Since \( R_i^* < R_2^* \) and \( \rho(w^T r - r_b) \) is a convex function of \( w \) (see Proposition 3), then \( X_1 = \{ w : \rho(w^T r - r_b) \leq R_i^* , w^Te = 1, Lb \leq Aw \leq Ub \} \subset X_2 = \{ w : \rho(w^T r - r_b) \leq R_2^* , w^Te = 1, Lb \leq Aw \leq Ub \} \). Thus

\[ \max_{x \in X_1} \mu(w^T r - r_b) \leq \max_{x \in X_2} \mu(w^T r - r_b) \]

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 \( w^o(R_i^*) \) 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 \( w^\lambda \) satisfies all linear constraints. In addition,

\[ \rho(w^{\lambda T}r - r_b) \leq \lambda \rho\left((w^o(R_1^*))^T r - r_b\right) + (1-\lambda)\rho\left((w^o(R_2^*))^T r - r_b\right) \]
\[ = \lambda R_1^* + (1-\lambda)R_2^* \]

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 \( R^* = R_1^* = \lambda R_1^* + (1-\lambda)R_2^* \), we obtain the optimal portfolio \( w^o(R_1^*) \) with reward

\[ \mu \left((w^o(R_1^*))^T r - r_b\right) \geq \mu(w^{\lambda T}r - r_b) \]
which, combined with the above inequality of the reward proves the claim. Moreover
the risk constraint is binding at \( \rho((w^o(R^*_\lambda))^T r - r_b) < R^*_\lambda \)
and let \( X_\lambda = \{ w : \rho(w^T r - r_b) \leq R^*_\lambda, w^T e = 1, Lb \leq Aw \leq Ub \} \). Since \( R^*_1 \leq R^*_\lambda \leq R^*_2 \) and \( \rho(w^T r - r_b) \) is a convex function of \( w \), \( X_\lambda \subset X_2 \). 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 \( \rho((w^o(R^*_2))^T r - r_b) = R^*_2 \).

\[\square\]

- **Proof of Proposition 11.**

**Proof.** The function \( z(y) = (\max(y, 0))^\gamma \) is convex for \( \gamma \geq 1 \). This is straightforward to verify

\[
z(\lambda y_1 + (1 - \lambda)y_2) = (\max(\lambda y_1 + (1 - \lambda)y_2, 0))^\gamma \leq (\lambda \max(y_1, 0) + (1 - \lambda) \max(y_2, 0))^\gamma \leq \lambda (\max(y_1, 0))^\gamma + (1 - \lambda)(\max(y_2, 0))^\gamma = \lambda z(y_1) + (1 - \lambda) z(y_2)
\]
The first inequality is because \( \max(\cdot, 0) \) is a convex function and the second inequality follows because \( q(y) = y^\gamma \) is convex for \( \gamma \geq 1 \) and \( y \geq 0 \); \( q''(y) = \gamma(\gamma - 1)y^{\gamma - 2} \geq 0 \). The function \( z(w^T x) = (\max(-w^T x, 0))^\gamma \) is convex in \( w \in X \subset \mathbb{R}^d \) for each fixed \( x \in D_1 \subset \mathbb{R}^d \) because it is a composition of a convex and a linear function which preserves convexity (see, for example, Boyd, Vandenberghe (2004)). Therefore

\[
z(w^{\lambda T} x) \leq \lambda z(w^{1 T} x) + (1 - \lambda) z(w^{2 T} x)
\]
where \( w^{\lambda} = \lambda w^1 + (1 - \lambda) w^2 \) and \( \lambda \in [0, 1] \). Multiplying both sides by \( f(x) > 0 \), \( x \in D_1 \) does not change the inequality

\[
z(w^{\lambda T} x) f(x) \leq \lambda z(w^{1 T} x) f(x) + (1 - \lambda) z(w^{2 T} x) f(x)
\]
The inequality is correct for every \( x \in D_1 \) and \( \lambda \in [0, 1] \). Let us consider the function

\[
g(w) = \int_A z(w^T x) f(x) dx = \int_A (\max(-w^T x, 0))^\gamma f(x) dx, \quad A \subset D_1 \quad (29)
\]
For any convex combination \( w^\lambda \), we have
\[ g(w^\lambda) = \int_A z(w^{\lambda^T} x) f(x) dx \]
\[ \leq \lambda \int_A z(w^{1^T} x) f(x) dx + (1 - \lambda) \int_A z(w^{2^T} x) f(x) dx \]
\[ = \lambda g(w^1) + (1 - \lambda) g(w^2) \]

and therefore \( g(w) \) is convex. According to the definition,

\[ ETL_{(\gamma, \alpha)}(w^T r) = \frac{1}{\alpha} \int_A (\max(-w^T x, 0))^\gamma f_r(x) dx, \quad A = \{ x : VaR_\alpha(w^T r) \leq -w^T x \} \]

where \( f_r(x) \) is the probability density function of the random vector \( r \). The expression for \( ETL_{(\gamma, \alpha)}(w^T r) \) is basically the same as equation (29) and it follows directly that \( ETL_{(\gamma, \alpha)}(w^T r) \) is convex for \( \gamma \geq 1 \). Adding a non-zero benchmark does not change the conclusion. In this case, we can consider the linear function \( w^T x - y \) in the max function and \( f_{(y, r)}(x) \) as a probability density function of the random vector \( (r_b, r) = (r_b, r_1, \ldots, r_n) \).}

- **Proof of Proposition 12.**

  **Proof.** According to the definition, for any \( t > 0 \)

\[ ETL_{(\gamma, \alpha)}(t X) = E\left( (\max(-t X, 0))^\gamma \mid -t X > VaR_\alpha(t X) \right) \]
\[ = \frac{1}{\alpha} \int_{VaR_\alpha(t X)}^\infty (\max(-x, 0))^\gamma f_{-X}(x) dx \]
\[ = \frac{1}{\alpha} \int_{VaR_\alpha(x)}^\infty (\max(-x, 0))^\gamma f_{-X}(x) \frac{dx}{x} \]
\[ = \frac{1}{\alpha} \int_{VaR_\alpha(x)}^\infty (\max(-y, 0))^\gamma f_{-X}(y) dy \]
\[ = \frac{t^\gamma}{\alpha} \int_{VaR_\alpha(x)}^\infty (\max(-y, 0))^\gamma f_{-X}(y) dy \]

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

\[ ETL_{(\gamma, \alpha)}(t X) = t^\gamma E\left( (\max(-X, 0))^\gamma \mid -X > VaR_\alpha(X) \right) \]
\[ = t^\gamma ETL_{(\gamma, \alpha)}(X) \]

and we have proved the claim. \( \square \)

- **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 \( t = (ETL_{(\delta, \beta)}(w^T r - r_b))^{-\delta} \) we obtain
\[
\max_{\{w,t\}} \ ET L_{(\delta,\alpha)} \left( tr_b - tw^T r \right) \\
\text{subject to} \ ET L_{(\delta,\beta)} \left( tw^T r - tr_b \right) = 1 \\
w^T e = 1 \\
Lb \leq Aw \leq Ub \\
t \geq 0
\]

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

\[
\max_{\{x,t\}} \ ET L_{(\delta,\alpha)} \left( tr_b - x^T r \right) \\
\text{subject to} \ ET L_{(\delta,\beta)} \left( x^T r - tr_b \right) = 1 \\
x^T e = t \\
tLb \leq Ax \leq tUb \\
t \geq 0
\]

The same reasoning as in the proof of Proposition 4 establishes the equivalence of Problem (26) with Problem (30). That is, if \( (x^o, t^o) \) is a solution of Problem (30), then \( w^o = x^o/t^o \) is a solution to Problem (26) and conversely, if \( w^o \) is a solution to Problem (26) then the pair \( (w^o, t^o) \) solves Problem (30) with \( t^o = (ET L_{(\delta,\beta)}(w^{oT} r - r_b))^{-\delta} \).

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 \( ET L_{(\delta,\beta)}(x^T r - tr_b) \leq 1 \) since it is necessarily binding. Indeed due to the homogeneity property stated in Proposition 12, we can rewrite the entire problem as

\[
\max_{\{w,t\}} \ t^\delta ET L_{(\delta,\alpha)} \left( r_b - w^T r \right) \\
\text{subject to} \ t^\delta ET L_{(\delta,\beta)} \left( w^T r - r_b \right) \leq 1 \\
tw^T e = t \\
tLb \leq Aw \leq tUb \\
t \geq 0
\]

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

\[
1 < a \leq \left( (t^o)^\delta ET L_{(\delta,\beta)}(w^o^{T} r - r_b))^{-1/\delta}
\]

The point \( (at^o w^o, at^o) \) is feasible and

\[
(t^o)^\delta ET L_{(\delta,\alpha)} \left( r_b - w^{oT} r \right) < a^\delta (t^o)^\delta ET L_{(\delta,\alpha)} \left( r_b - w^{oT} r \right)
\]

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

The equivalence with Problem (GRR B) is proved with the same arguments applied to the problems.
\begin{equation}
\begin{align*}
\min_{w} & \quad \frac{ETL(\delta, \beta)(w^T r - r_b)}{ETL(\delta, \alpha)(r_b - w^T r)} \\
\text{subject to} & \quad w^T e = 1 \\
& \quad Lb \leq Aw \leq Ub
\end{align*}
\tag{31}
\end{equation}

and

\begin{equation}
\begin{align*}
\min_{(x,t)} & \quad ETL(\delta, \beta) \left( x^T r - tr_b \right) \\
\text{subject to} & \quad ETL(\delta, \alpha) \left( tr_b - x^T r \right) = 1 \\
& \quad x^T e = t \\
& \quad tLb \leq Ax \leq tUb \\
& \quad t \geq 0
\end{align*}
\end{equation}

where \( t = (ETL(\delta, \alpha)(r_b - w^T r))^{-\delta} \) and then considering the relaxation \( ETL(\delta, \alpha)(tr_b - x^T r) \geq 1 \) of the risk constraint. The fact that Problem (31) is equivalent to the generic ration optimization Problem (26) follows from the same arguments as in the proof of Proposition 2, part c). Actually the result there is not dependent of any assumption of convexity.
References


