

On LP Solvable Models for Portfolio Selection

Renata Mansini*, Włodzimierz Ogryczak†, M. Grazia Speranza‡

October 31, 2002

Abstract

The Markowitz model for single period portfolio optimization quantifies the problem by means of only two criteria: the mean, representing the expected outcome, and the risk, a scalar measure of the variability of outcomes. The classical Markowitz model uses the variance as the risk measure, thus resulting in a quadratic optimization problem. Following Sharpe's work on linear approximation to the mean–variance model, many attempts have been made to linearize the portfolio optimization problem. There were introduced several alternative risk measures which are computationally attractive as (for discrete random variables) they result in solving Linear Programming (LP) problems. The LP solvability is very important for applications to real-life financial decisions where the constructed portfolios have to meet numerous side constraints and take into account transaction costs. This paper provides a systematic overview of the LP solvable models with a wide discussion of their properties.

Keywords: Portfolio optimization, mean–risk model, linear programming

*University of Brescia, Department of Electronics for Automation, via Branze 38, 25123 Brescia, Italy

†Warsaw University of Technology, Institute of Control and Computation Engineering, Nowowiejska 15/19, 00-665 Warsaw, Poland

‡University of Brescia, Department of Quantitative Methods, C. da S.Chiera 48/B, 25122 Brescia, Italy

1 Introduction

The portfolio optimization problem considered in this paper follows the original Markowitz' formulation and is based on a single period model of investment. At the beginning of a period, an investor allocates the capital among various securities, thus assigning a nonnegative weight (share of the capital) to each security. During the investment period, a security generates a random rate of return. This results in a change of the capital invested (observed at the end of the period) which is measured by the weighted average of the individual rates of return.

Let $J = \{1, 2, \dots, n\}$ denote a set of securities considered for an investment. For each security $j \in J$, its rate of return is represented by a random variable R_j with a given mean $\mu_j = \mathbb{E}\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,2,\dots,n}$ denote a vector of decision variables x_j expressing the weights defining a portfolio. To represent a portfolio, the weights must satisfy a set of constraints that form a feasible set $\mathcal{P}$. The simplest way of defining a feasible set is by a requirement that the weights must sum to one and short sales are not allowed, i.e.

$$\mathcal{P} = \{\mathbf{x} : \sum_{j=1}^n x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \dots, n\}. \quad (1)$$

Hereafter, it is assumed that $\mathcal{P}$ is a general LP feasible set given in a canonical form as a system of linear equations with nonnegative variables: This allows one to include upper bounds on single shares as well as several more complex portfolio structure restrictions which may be faced by a real-life investor.

Each portfolio $\mathbf{x}$ defines a corresponding random variable $R_{\mathbf{x}} = \sum_{j=1}^n R_j x_j$ that represents the portfolio rate of return. The mean rate of return for portfolio $\mathbf{x}$ is given as:

$$\mu(\mathbf{x}) = \mathbb{E}\{R_{\mathbf{x}}\} = \sum_{j=1}^n \mu_j x_j.$$

Following the seminal work by Markowitz [17], the portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem where $\mu(\mathbf{x})$ is maximized and some risk measure $\varrho(\mathbf{x})$ is minimized. In the original Markowitz model [17] the risk is measured by the standard deviation or variance: $\sigma^2(\mathbf{x}) = \mathbb{E}\{(\mu(\mathbf{x}) - R_{\mathbf{x}})^2\}$. Several other risk measures have been later considered thus creating the entire family of mean-risk (Markowitz-type) models. While the original Markowitz model forms a quadratic programming problem, following Sharpe [32], many attempts have been made to linearize the portfolio optimization procedure (c.f., [37] and references therein). The LP solvability is very important for applications to real-life financial decisions where the constructed portfolios have to meet numerous side constraints (including the minimum transaction lots [16] and the transaction costs [10, 13]). All these lead to the mixed integer LP structure of the portfolio feasible set $\mathcal{P}$. Some papers also appeared in the literature which consider restrictions on the number of securities in the portfolio and other side constraints in the Markowitz model (see for instance [5] and [9]). Certainly, in order to guarantee that the portfolio takes advantage of diversification, no risk measure can be a linear function of $\mathbf{x}$. Nevertheless, a risk measure can be LP computable in the case of discrete random variables, i.e., in the case of returns defined by their realizations under the specified scenarios. We will consider T scenarios with probabilities p_t (where $t = 1, \dots, T$). We will assume that for each random variable R_j its realization r_{jt} under the scenario t is known. Typically, the realizations are derived from historical data treating T historical periods as equally probable scenarios ($p_t = 1/T$). The realizations of the portfolio return $R_{\mathbf{x}}$ are given as $y_t = \sum_{j=1}^n r_{jt} x_j$ and the expected value $\mu(\mathbf{x})$ can be computed as:

$$\mu(\mathbf{x}) = \sum_{t=1}^T y_t p_t = \sum_{t=1}^T \left[\sum_{j=1}^n r_{jt} x_j \right] p_t.$$

Similarly, several risk measures can be LP computable with respect to the realizations y_t .

The assumption on equally probable scenarios is the most typically applied when using historical data, but other ways to compute scenario probabilities have been proposed. In [37], for instance, historical realizations are weighted in a semi-absolute deviation model for portfolio selection by means of the exponential smoothing technique.

When, in particular, securities have to be priced at some future time period the pricing models may be based on Monte Carlo simulation of the term structure (see [31]). In this case possible states of the economy at a given time period are usually obtained by means of a binomial lattice. More precisely, each price is obtained as the expected discounted value of its cash flow with discounting done at the risk free rate. Notice that any suitable term structure can be used for the purpose of Monte Carlo simulation (see the binomial lattice model described in Black et al. [4]).

In [6] while analyzing the technical aspects of the Russell-Yasuda Kasai financial planning model, the authors also consider different models for discrete distribution scenario generation. In particular the proposed software allows the user to select among different discrete scenarios generation: the scenarios can be independent period by period, dependent through some factor model or general. In all these cases the number and the form of the scenarios are crucial factors in the size and complexity of the models to be solved.

Finally, when considering historical data the impact of parameter estimation on optimal portfolio selection has been recognized by a number of authors who show that practical application of portfolio analysis can be seriously hampered by estimation error, especially in expected return. As pointed out by Simaan [35] in order to reduce estimation error the number of historical periods taken into account should be sufficiently large and be strictly dependent on the used measure of risk.

The mean absolute deviation was very early considered in the portfolio analysis ([33] and references therein) while quite recently Konno and Yamazaki [12] presented and analyzed the complete portfolio LP solvable optimization model based on this risk measure — the so-called MAD model. Yitzhaki [40] introduced the mean-risk model using Gini's mean (absolute) difference as the risk measure. For a discrete random variable represented by its realizations y_t , the Gini's mean difference is LP computable (when minimized). Recently, Young [41] analyzed the LP solvable portfolio optimization model based on risk defined by the worst case scenario (minimax approach), while Ogryczak [22] introduced the multiple criteria LP model covering all the above as special aggregation techniques.

The Markowitz model is frequently criticized as not consistent with axiomatic models of preferences for choice under risk [29]. Models consistent with the preference axioms are based on the relations of stochastic dominance or on expected utility theory [38, 15]. The relations of stochastic dominance allow a pairwise comparison of given portfolios but do not offer any computational recipe to analyze the portfolio selection problem. The expected utility theory, when used for the portfolio selection problem, is restrictive in modeling preferences of investors. If the rates of return are normally distributed, then most of the LP computable risk measures become proportional to the standard deviation $\sigma(\mathbf{x})$ [14, pp. 1216–1217]. Hence the corresponding LP solvable mean-risk models are then equivalent to the Markowitz mean-variance model. However, the LP solvable mean-risk models do not require any specific type of return distributions. Moreover, opposite to the mean-variance approach, for general random variables some consistency with the stochastic dominance relations was shown for the Gini's mean difference [40], for the MAD model [24] and for many other LP solvable models as well [22].

It is often argued that the variability of the rate of return above the mean should not be penalized since the investors are concerned of an underperformance rather than the overperformance of a portfolio. This led Markowitz [18] to propose downside risk measures such as (downside) semivariance to replace variance as the risk measure. Consequently, one observes

growing popularity of downside risk models for portfolio selection [36]. Actually, most of the LP solvable models may be viewed as based on some downside risk measures. Moreover, the models may be extended with some piece-wise linear penalty (risk) functions to provide opportunities for more specific modeling of the downside risk [6, 11, 19, 20].

The variety of LP solvable portfolio optimization models presented in the literature generates a need for their classification and comparison. This is the major goal of this paper. We provide a systematic overview of the models with a wide discussion of their theoretical properties.

The paper is organized as follows. In the next section we show how various LP computable performance measures can be derived from shortfall criteria related to the stochastic dominance. Section 3 gives a detailed revue of the LP solvable portfolio optimization models we examine. Finally, in Section 4 some concluding remarks are stated.

2 Shortfall criteria and performance measures

In this section we first recall the concepts of shortfall criteria and stochastic dominance. Then, we show how various possible portfolio performance measures can be derived from shortfall criteria and that some are consistent with the stochastic dominance relations. Some of the performance measures are risk measures (to be minimized) and some are safety measures (to be maximized). We also show how these measures become LP computable in the case of returns defined by discrete random variables.

2.1 Shortfall criteria and stochastic dominance

The notion of risk is related to a possible failure of achieving some targets. It was formalized as the so-called safety-first strategies [30] and later led to the concept of below-target risk measures [8] or shortfall criteria. The simplest shortfall criterion for the specific target value τ is the mean below-target deviation

$$\bar{\delta}_\tau(\mathbf{x}) = \mathbb{E}\{\max\{\tau - R_{\mathbf{x}}, 0\}\} \quad (2)$$

which is LP computable for returns represented by their realizations y_t as:

$$\bar{\delta}_\tau(\mathbf{x}) = \min \sum_{t=1}^T d_t^- p_t \quad \text{subject to} \quad d_t^- \geq \tau - y_t, \quad d_t^- \geq 0 \quad \text{for } t = 1, \dots, T. \quad (3)$$

We show that the concept of mean below-target deviation is related to the second degree stochastic dominance relation [38] which is based on an axiomatic model of risk-averse preferences [29]. In stochastic dominance, uncertain returns (random variables) are compared by pointwise comparison of functions constructed from their distribution functions. The first function $F_{\mathbf{x}}^{(1)}$ is given as the right-continuous cumulative distribution function of the rate of return $F_{\mathbf{x}}^{(1)}(\eta) = F_{\mathbf{x}}(\eta) = \mathbb{P}\{R_{\mathbf{x}} \leq \eta\}$ and it defines the weak relation of the *first degree stochastic dominance* (FSD) as follows :

$$R_{\mathbf{x}'} \succeq_{FSD} R_{\mathbf{x}''} \quad \Leftrightarrow \quad F_{\mathbf{x}'}(\eta) \leq F_{\mathbf{x}''}(\eta) \quad \text{for all } \eta.$$

The second function is derived from the first as:

$$F_{\mathbf{x}}^{(2)}(\eta) = \int_{-\infty}^{\eta} F_{\mathbf{x}}(\xi) d\xi \quad \text{for real numbers } \eta,$$

and defines the (weak) relation of *second degree stochastic dominance* (SSD):

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Leftrightarrow \quad F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta) \quad \text{for all } \eta.$$

We say that portfolio $\mathbf{x}'$ *dominates* $\mathbf{x}''$ under the SSD ($R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$), if $F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta)$ for all η , with at least one strict inequality. A feasible portfolio $\mathbf{x}^0 \in \mathcal{P}$ is called *SSD efficient* if there is no $\mathbf{x} \in \mathcal{P}$ such that $R_{\mathbf{x}} \succ_{SSD} R_{\mathbf{x}^0}$. If $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$, then $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ within all risk-averse preference models where larger outcomes are preferred.

Note that the SSD relation covers increasing and concave utility functions, while the first stochastic dominance is less specific as it covers all increasing utility functions [15], thus neglecting a risk averse attitude. It is therefore a matter of primary importance that a model for portfolio optimization be consistent with the SSD relation, in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies that the performance measure in $\mathbf{x}'$ takes a value not worse than (lower than or equal to, in the case of a risk measure) in $\mathbf{x}''$. The consistency with the SSD relation implies that an optimal portfolio is SSD efficient.

Function $F_{\mathbf{x}}^{(2)}$, used to define the SSD relation, can also be presented as follows [24]:

$$F_{\mathbf{x}}^{(2)}(\eta) = \mathbb{P}\{R_{\mathbf{x}} \leq \eta\} \mathbb{E}\{\eta - R_{\mathbf{x}} | R_{\mathbf{x}} \leq \eta\} = \mathbb{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\} = \bar{\delta}_{\eta}(\mathbf{x}).$$

Hence, the SSD relation can be seen as a dominance for mean below-target deviations from all possible targets. The mean below-target deviation from a specific target (2) represents only a single criterion. One may consider several, say m , targets $\tau_1 > \tau_2 > \dots > \tau_m$ and use the weighted sum of the shortfall criteria as a risk measure

$$\sum_{k=1}^m w_k \bar{\delta}_{\tau_k}(\mathbf{x}) = \sum_{k=1}^m w_k \mathbb{E}\{\max\{\tau_k - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\left\{\sum_{k=1}^m w_k \max\{\tau_k - R_{\mathbf{x}}, 0\}\right\} \quad (4)$$

where w_k (for $k = 1, \dots, m$) are positive weights which maintains LP computability of the measure (when minimized). Actually, the measure (4) can be interpreted as a single mean below-target deviation applied with a penalty function: $\mathbb{E}\{u(\max\{\tau_1 - R_{\mathbf{x}}, 0\})\}$ where u is increasing and convex piece-wise linear penalty function with breakpoints $b_k = \tau_1 - \tau_k$ and slopes $s_k = w_1 + \dots + w_k$, $k = 1, \dots, m$. Such a piece-wise linear penalty function is used in the Russel-Yasuda-Kasai financial planning model [6] to define the corresponding risk measure.

2.2 MAD and downside versions

When an investment situation involves minimal acceptable returns, then the below-target deviation and its extensions are considered to be good risk measures [8]. However, when the mean portfolio return is used as a target, then in (2) the mean $\mu(\mathbf{x})$ can be used instead of the fixed target τ .

This results in the risk measure known as the *downside mean semideviation* from the mean

$$\bar{\delta}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = F_{\mathbf{x}}^{(2)}(\mu(\mathbf{x})). \quad (5)$$

For a discrete random variable represented by its realizations y_t , the mean semideviation (5), when minimized, is LP computable by formula(3) with $\tau = \mu(\mathbf{x})$.

The downside mean semideviation is always equal to the upside one and therefore we refer to it hereafter as to the mean semideviation. The mean semideviation is a half of the mean absolute deviation from the mean (see [37]), i.e.

$$\delta(\mathbf{x}) = \mathbb{E}\{|R_{\mathbf{x}} - \mu(\mathbf{x})|\} = 2\bar{\delta}(\mathbf{x}).$$

Hence the corresponding mean-risk model is equivalent to the MAD model which is LP computable as:

$$\delta(\mathbf{x}) = \min \sum_{t=1}^T (d_t^- + d_t^+) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \mu(\mathbf{x}) - y_t, \quad d_t^-, d_t^+ \geq 0 \quad \forall t = 1, \dots, T. \quad (6)$$

Due to the use of distribution dependent target value $\mu(\mathbf{x})$, the mean semideviation cannot be directly considered an SSD shortfall criterion. However, as shown by Ogryczak and Ruszczyński [24], the mean semideviation is closely related to the graph of $F_{\mathbf{x}}^{(2)}$. The function $F_{\mathbf{x}}^{(2)}$ is continuous, convex, nonnegative and nondecreasing. The graph $F_{\mathbf{x}}^{(2)}(\eta)$, referred to as the Outcome–Risk (O–R) diagram, has two asymptotes which intersect at the point $(\mu(\mathbf{x}), 0)$ (Figure 1). Specifically, the η -axis is the left asymptote and the line $\eta - \mu(\mathbf{x})$ is the right asymptote. In the case of a deterministic (risk-free) return ($R_{\mathbf{x}} = \mu(\mathbf{x})$), the graph of $F_{\mathbf{x}}^{(2)}(\eta)$ coincides with the asymptotes, whereas any uncertain return with the same expected value $\mu(\mathbf{x})$ yields a graph above (precisely, not below) the asymptotes. The space between the curve $(\eta, F_{\mathbf{x}}^{(2)}(\eta))$, and its asymptotes represents the dispersion (and thereby the riskiness) of $R_{\mathbf{x}}$ in comparison to the deterministic return $\mu(\mathbf{x})$. Therefore, it is called the dispersion space. The mean semideviation turns out to be the largest vertical diameter of the dispersion space while the variance represents its doubled area [24].

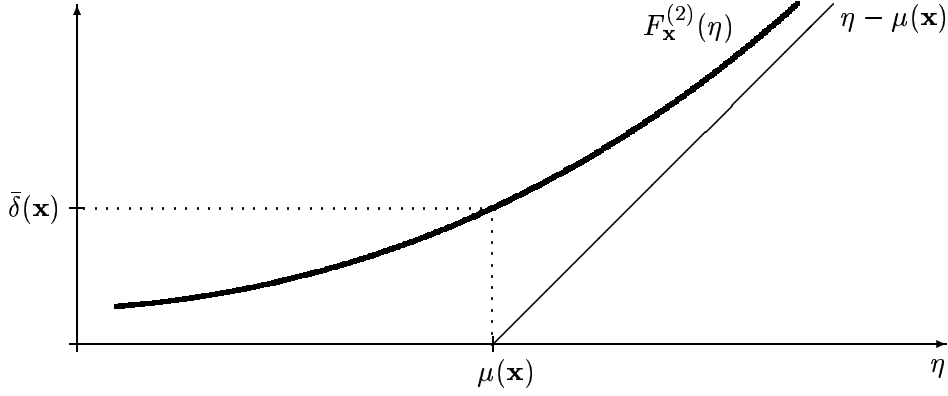


Figure 1: The O–R diagram and the mean semideviation

Every shortfall risk measure or, more precisely, every pair of a target value τ and the corresponding downside deviation defines also the quantity of mean below-target underachievement

$$\tau - \bar{\delta}_{\tau}(\mathbf{x}) = \mathbb{E}\{\tau - \max\{\tau - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\min\{R_{\mathbf{x}}, \tau\}\}.$$

The latter portfolio performance measure can be considered a safety measure as the larger values are preferred. In the case of a fixed target τ one gets $\tau - \bar{\delta}_{\tau}(\mathbf{x}') \geq \tau - \bar{\delta}_{\tau}(\mathbf{x}'')$ iff $\bar{\delta}_{\tau}(\mathbf{x}') \leq \bar{\delta}_{\tau}(\mathbf{x}'')$. Hence, the minimization of the mean below-target deviation (risk measure) and the maximization of the corresponding mean below-target underachievement (safety measure) are equivalent. The latest property is no longer valid when $\mu(\mathbf{x})$ is used as the target. One may introduce the safety measure of *mean downside underachievement*

$$\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) = \mathbb{E}\{\mu(\mathbf{x}) - \max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}\} \quad (7)$$

but the minimization of the mean semideviation is, in general, not equivalent to the maximization of the mean downside underachievement. Note that, as shown in [24], $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies the inequality $\mu(\mathbf{x}') - \bar{\delta}(\mathbf{x}') \geq \mu(\mathbf{x}'') - \bar{\delta}(\mathbf{x}'')$ while the corresponding inequality on the mean semideviations $\bar{\delta}(\mathbf{x}') \leq \bar{\delta}(\mathbf{x}'')$ may not be valid. Thus, the mean downside underachievement is consistent with the SSD relation, while the consistency is not guaranteed for the mean semideviation.

For better modeling of the downside risk, one may consider a risk measure defined by the mean semideviation applied with a piece-wise linear penalty function [11] to penalize larger downside deviations. It turns out, however, that for maintaining both the LP computability and SSD consistency [20], the breakpoints (or additional target values) must be located at

the corresponding mean downside underachievements (7). Namely, when using m distribution dependent targets $\mu_1(\mathbf{x}) = \mu(\mathbf{x})$, $\mu_2(\mathbf{x}), \dots, \mu_m(\mathbf{x})$ and the corresponding mean semideviations $\bar{\delta}_1(\mathbf{x}) = \bar{\delta}(\mathbf{x})$, $\bar{\delta}_2(\mathbf{x}), \dots, \bar{\delta}_m(\mathbf{x})$ defined recursively according to the formulas:

$$\begin{aligned}\bar{\delta}_k(\mathbf{x}) &= \mathbb{E}\{\max\{\mu_k(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\max\{\mu(\mathbf{x}) - \sum_{i=1}^{k-1} \bar{\delta}_i(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \\ \mu_{k+1}(\mathbf{x}) &= \mu_k(\mathbf{x}) - \bar{\delta}_k(\mathbf{x}) = \mu(\mathbf{x}) - \sum_{i=1}^k \bar{\delta}_i(\mathbf{x}) = \mathbb{E}\{\min\{R_{\mathbf{x}}, \mu_k(\mathbf{x})\}\},\end{aligned}$$

one may combine the semideviations by the weighted sum to the measure

$$\bar{\delta}^{(m)}(\mathbf{x}) = \sum_{k=1}^m w_k \bar{\delta}_k(\mathbf{x}) \quad 1 = w_1 \geq w_2 \geq \dots \geq w_m \geq 0 \quad (8)$$

as in the m -MAD model [20]. Actually, the measure can be interpreted as a single mean semideviation (from the mean) applied with a penalty function: $\bar{\delta}^{(m)}(\mathbf{x}) = \mathbb{E}\{u(\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\})\}$ where u is increasing and convex piece-wise linear penalty function with breakpoints $b_k = \mu(\mathbf{x}) - \mu_k(\mathbf{x})$ and slopes $s_k = w_1 + \dots + w_k$, $k = 1, \dots, m$. Therefore, we will refer to the measure $\bar{\delta}^{(m)}(\mathbf{x})$ as to the *mean penalized semideviation*. The mean penalized semideviation (8) defines the corresponding safety measure $\mu(\mathbf{x}) - \bar{\delta}^{(m)}(\mathbf{x})$ which may be expressed directly as the *weighted sum of the mean downside underachievements* $\mu_k(\mathbf{x})$:

$$\mu(\mathbf{x}) - \bar{\delta}^{(m)}(\mathbf{x}) = (w_1 - w_2)\mu_2(\mathbf{x}) + (w_2 - w_3)\mu_3(\mathbf{x}) + \dots + (w_{m-1} - w_m)\mu_m(\mathbf{x}) + w_m\mu_{m+1}(\mathbf{x}) \quad (9)$$

where the weights coefficients are nonnegative and total to 1. This safety measure was shown [20] to be SSD consistent in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \bar{\delta}^{(m)}(\mathbf{x}') \geq \mu(\mathbf{x}'') - \bar{\delta}^{(m)}(\mathbf{x}'')$.

2.3 Minimax and the CVaR measures

For a discrete random variable represented by its realizations y_t , the *worst realization*

$$M(\mathbf{x}) = \min_{t=1, \dots, T} y_t \quad (10)$$

is a well appealing safety measure, while the *maximum (downside) semideviation*

$$\Delta(\mathbf{x}) = \mu(\mathbf{x}) - M(\mathbf{x}) = \max_{t=1, \dots, T} (\mu(\mathbf{x}) - y_t) \quad (11)$$

represents the corresponding risk measure. The latter is well defined in the O-R diagram (Fig. 1) as it represents the maximum horizontal diameter of the dispersion space. The measure $M(\mathbf{x})$ is known to be SSD consistent and it was applied to portfolio optimization by Young [41].

A natural generalization of the measure $M(\mathbf{x})$ is the worst conditional expectation or the conditional value-at-risk (CVaR) [28] defined as the mean of the specified size (quantile) of worst realizations. For the simplest case of equally probable scenarios ($p_t = 1/T$), one may define the CVaR measure $M_{\frac{k}{T}}(\mathbf{x})$ as the mean return under the k worst scenarios. In general, the *conditional value-at-risk (CVaR)* and the *worst conditional semideviation* (conditional drawdown) for any real tolerance level $0 < \beta \leq 1$ (replacing the quotient k/T) are defined as

$$M_{\beta}(\mathbf{x}) = \frac{1}{\beta} \int_0^{\beta} F_{\mathbf{x}}^{(-1)}(\alpha) d\alpha \quad \text{for } 0 < \beta \leq 1 \quad (12)$$

and

$$\Delta_{\beta}(\mathbf{x}) = \mu(\mathbf{x}) - M_{\beta}(\mathbf{x}) \quad \text{for } 0 < \beta \leq 1 \quad (13)$$

where $F_{\mathbf{x}}^{(-1)}(p) = \inf \{ \eta : F_{\mathbf{x}}(\eta) \geq p \}$ is the left-continuous inverse of the cumulative distribution function $F_{\mathbf{x}}$. For any tolerance level $0 < \beta \leq 1$ the corresponding CVaR measure $M_{\beta}(\mathbf{x})$ is an SSD consistent measure. Actually, the CVaR measures provide an alternative characterization of the SSD relation [23, 25] in the sense of the following equivalence:

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \Leftrightarrow M_{\beta}(\mathbf{x}') \geq M_{\beta}(\mathbf{x}'') \quad \text{for all } 0 < \beta \leq 1. \quad (14)$$

Note that $M_1(\mathbf{x}) = \mu(\mathbf{x})$ and $M_{\beta}(\mathbf{x})$ tends to $M(\mathbf{x})$ for β approaching 0. By the theory of convex conjugent (dual) functions [27], the worst conditional expectation may be defined by optimization [25]:

$$M_{\beta}(\mathbf{x}) = \max_{\eta \in R} \left[\eta - \frac{1}{\beta} F_{\mathbf{x}}^{(2)}(\eta) \right] = \max_{\eta \in R} \left[\eta - \frac{1}{\beta} \bar{\delta}_{\eta}(\mathbf{x}) \right] \quad (15)$$

where η is a real variable taking the value of β -quantile $Q_{\beta}(\mathbf{x})$ at the optimum [25]. Hence, the value of CVaR and the corresponding worst conditional semideviation express the results of the O-R diagram analysis according to a slant direction defined the slope β (Fig. 2).

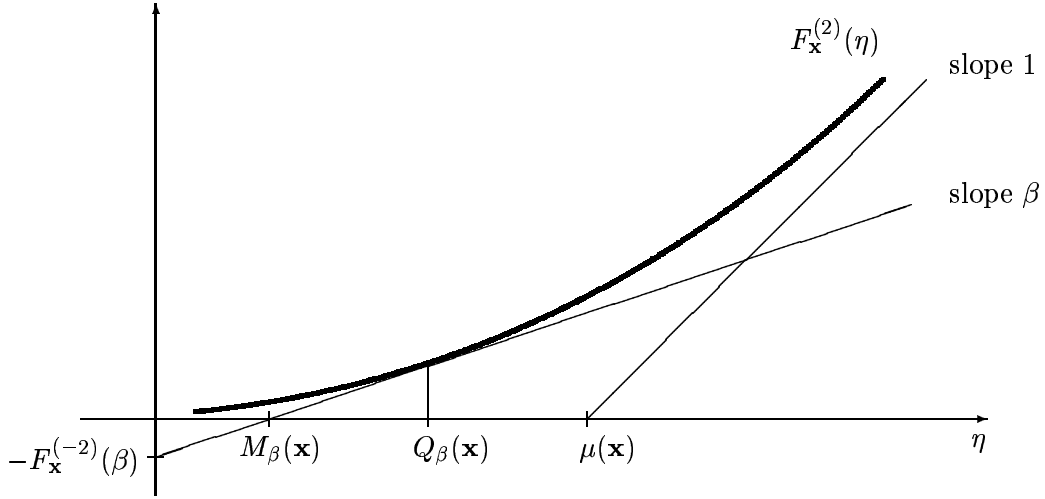


Figure 2: Quantile safety measures in the O-R diagram

For a discrete random variable represented by its realizations y_t , due to (3), problem (15) becomes an LP. Thus

$$M_{\beta}(\mathbf{x}) = \max \left[\eta - \frac{1}{\beta} \sum_{t=1}^T d_t^- p_t \right] \quad \text{s.t.} \quad d_t^- \geq \eta - y_t, \quad d_t^- \geq 0 \quad \text{for } t = 1, \dots, T \quad (16)$$

whereas the worst conditional semideviations may be computed as the corresponding differences from the mean ($\Delta_{\beta}(\mathbf{x}) = \mu(\mathbf{x}) - M_{\beta}(\mathbf{x})$) or directly as:

$$\Delta_{\beta}(\mathbf{x}) = \min \sum_{t=1}^T (d_t^+ + \frac{1-\beta}{\beta} d_t^-) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \eta - y_t, \quad d_t^+, d_t^- \geq 0 \quad \text{for } t = 1, \dots, T$$

where η is an auxiliary (unbounded) variable. Note that for $\beta = 0.5$ one has $1 - \beta = \beta$. Hence, $\Delta_{0.5}(\mathbf{x})$ represents the mean absolute deviation from the median, the risk measure suggested by Sharpe [33]. The LP problem for computing this measure takes the form:

$$\Delta_{0.5}(\mathbf{x}) = \min \sum_{t=1}^T (d_t^- + d_t^+) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \eta - y_t, \quad d_t^+, d_t^- \geq 0 \quad \text{for } t = 1, \dots, T.$$

One may notice that the above models differs from the classical MAD formulation (6) only due to replacement of $\mu(\mathbf{x})$ with (unrestricted) variable η .

2.4 Gini's mean difference

Yitzhaki [40] introduced the mean-risk model using Gini's mean (absolute) difference as the risk measure. For a discrete random variable represented by its realizations y_t , the *Gini's mean difference*

$$\Gamma(\mathbf{x}) = \frac{1}{2} \sum_{t'=1}^T \sum_{t''=1}^T |y_{t'} - y_{t''}| p_{t'} p_{t''} \quad (17)$$

is obviously LP computable (when minimized).

In the case of equally probable T scenarios with $p_t = 1/T$ the Gini's mean difference may be expressed as the weighted average of the worst conditional semideviations $\Delta_{\frac{k}{T}}(\mathbf{x})$ for $k = 1, \dots, T$ [22]. Exactly, using weights $w_k = (2k)/T^2$ for $k = 1, 2, \dots, T-1$ and $w_T = 1/T = 1 - \sum_{k=1}^{T-1} w_k$, one gets $\Gamma(\mathbf{x}) = \sum_{k=1}^T w_k \Delta_{\frac{k}{T}}(\mathbf{x})$. On the other hand, for general discrete distributions, directly from the definition (17):

$$\Gamma(\mathbf{x}) = \sum_{t'=1}^T \left[\sum_{t'' : y_{t''} < y_{t'}} (y_{t'} - y_{t''}) p_{t''} \right] p_{t'} = \sum_{t=1}^T F_{\mathbf{x}}^{(2)}(y_t) p_t = \sum_{t=1}^T \bar{\delta}_{y_t}(\mathbf{x}) p_t.$$

Hence, $\Gamma(\mathbf{x})$ can be interpreted as the weighted sum of multiple mean below-target deviations (4) but both the targets and the weights are distribution dependent. This corresponds to an interpretation of $\Gamma(\mathbf{x})$ as the integral of $F_{\mathbf{x}}^{(2)}$ with respect to the probability measure induced by $R_{\mathbf{x}}$ [25]. Thus although not representing directly any shortfall criterion, the Gini's mean difference is a combination of the basic shortfall criteria.

Note that the Gini's mean difference defines the corresponding safety measure

$$\mu(\mathbf{x}) - \Gamma(\mathbf{x}) = \mathbb{E}\{R_{\mathbf{x}} \wedge R_{\mathbf{x}}\} \quad (18)$$

where the cumulative distribution function of $R_{\mathbf{x}} \wedge R_{\mathbf{x}}$ for any $\eta \in \mathbb{R}$ is given as $F_{\mathbf{x}}(\eta)(2 - F_{\mathbf{x}}(\eta))$. Hence, (18) is the expectation of the minimum of two independent identically distributed random variables (i.i.d.r.v.) $R_{\mathbf{x}}$ [40] thus representing the *mean worse return*. This safety measure is SSD consistent [40, 25] in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \Gamma(\mathbf{x}') \geq \mu(\mathbf{x}'') - \Gamma(\mathbf{x}'')$. In the case of equally probable T scenarios with $p_t = 1/T$ the safety measure $\mu(\mathbf{x}) - \Gamma(\mathbf{x})$ may be expressed as the weighted average of the CVaR values $M_{\frac{k}{T}}(\mathbf{x})$, for $k = 1, \dots, T$, with weights $w_k = (2k)/T^2$ for $k = 1, 2, \dots, T-1$ and $w_T = 1/T$.

3 LP solvable models

In this section we present the complete set of LP solvable models we consider and their LP formulation.

3.1 Risk and safety measures

Following Markowitz [17], the portfolio optimization problem is modeled as a mean-risk bicriteria optimization problem

$$\max\{[\mu(\mathbf{x}), -\varrho(\mathbf{x})] : \mathbf{x} \in \mathcal{P}\} \quad (19)$$

where the mean $\mu(\mathbf{x})$ is maximized and the risk measure $\varrho(\mathbf{x})$ is minimized. A feasible portfolio $\mathbf{x}^0 \in \mathcal{P}$ is called efficient solution of problem (19) or μ/ϱ -efficient portfolio if there is no $\mathbf{x} \in \mathcal{P}$ such that $\mu(\mathbf{x}) \geq \mu(\mathbf{x}^0)$ and $\varrho(\mathbf{x}) \leq \varrho(\mathbf{x}^0)$ with at least one inequality strict.

The original Markowitz model [17] uses the standard deviation $\sigma(\mathbf{x})$ as the risk measure. As shown in the previous section, several other risk measures may be used instead of the standard deviation thus generating the corresponding LP solvable mean-risk models. In this paper we restrict our analysis to the risk measures which, similar to the standard deviation, are shift independent dispersion parameters. Thus, they are equal to 0 in the case of a risk free portfolio and take positive values for any risky portfolio. This excludes the direct use of the mean below-target deviation (2) and its extensions with penalty functions (4). Nevertheless, as shown in Section 2, there is a gamut of LP computable risk measures fitting the requirements.

In Section 2 we have seen that in the literature some of the LP computable measures are dispersion type risk measures and some are safety measures, which, when embedded in an optimization model, are maximized instead of being minimized. Moreover, we have shown that each risk measure $\varrho(\mathbf{x})$ has a well defined corresponding safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$ and viceversa. Although the risk measures are more "natural", due to the consolidated familiarity with Markowitz model, we have seen that the safety measures, contrary to the dispersion type risk measures, satisfy axioms of the so-called coherent risk measurement [1]. Moreover, they are SSD consistent in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \varrho(\mathbf{x}') \geq \mu(\mathbf{x}'') - \varrho(\mathbf{x}'')$ [20, 24, 23, 40, 41].

The practical consequence of the lack of SSD consistency can be illustrated by two portfolios $\mathbf{x}'$ and $\mathbf{x}''$ (with rate of return given in percents):

$$\mathbb{P}\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 1, & \xi = 1.0 \\ 0, & \text{otherwise} \end{cases} \quad \mathbb{P}\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 1/2, & \xi = 3.0 \\ 1/2, & \xi = 5.0 \\ 0, & \text{otherwise} \end{cases}$$

Note that the risk free portfolio $\mathbf{x}'$ with the guaranteed result 1.0 is obviously worse than the risky portfolio $\mathbf{x}''$ giving 3.0 or 5.0. Certainly, in all models consistent with the preference axioms [1, 15, 38], portfolio $\mathbf{x}'$ is dominated by $\mathbf{x}''$, in particular $R_{\mathbf{x}''} \succeq_{SSD} R_{\mathbf{x}'}$. When a dispersion type risk measure $\varrho(\mathbf{x})$ is used, then both the portfolios may be efficient in the corresponding mean-risk model since for each such measure $\varrho(\mathbf{x}'') > 0$ while $\varrho(\mathbf{x}') = 0$. This is a common flaw of all Markowitz-type mean-risk models where risk is measured with some dispersion measure.

It is interesting to note that, in order to overcome this weakness of the Markowitz model already in 1964 Baumol [2] suggested to consider a safety measure, which he called the expected gain-confidence limit criterion, $\mu(\mathbf{x}) - \lambda\sigma(\mathbf{x})$ to be maximized instead of the $\sigma(\mathbf{x})$ minimization. Thus, on the basis of the above remarks, for each risk measure, it is reasonable to consider also an alternative mean-safety bicriteria model:

$$\max\{\mu(\mathbf{x}), \mu(\mathbf{x}) - \varrho(\mathbf{x})\} : \quad \mathbf{x} \in \mathcal{P}. \quad (20)$$

Summarizing, in a computational analysis one should consider the full set of risk and safety measures presented in Table 1.

Note that the MAD model was first introduced [12] with the risk measure of mean absolute deviation $\delta(\mathbf{x})$ whereas the mean semideviation $\bar{\delta}(\mathbf{x})$ we consider is half of it. This is due to the fact that the resulting optimization models are equivalent and that the model with the semideviation is more efficient (see [37]). For the MAD model, the safety measure represents the mean downside underachievement. For the m -MAD model the two measures represent the mean penalized semideviation and the weighted sum of the mean downside underachievements, respectively.

The Minimax model was considered and tested [41] with the safety measure of the worst realization $M(\mathbf{x})$ but it was also analyzed with the maximum semideviation $\Delta(\mathbf{x})$ [22]. The

Table 1: Risk and safety measures

	Risk measure $\varrho(\mathbf{x})$	Safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$
MAD model [12]	$\bar{\delta}(\mathbf{x})$ (5)	$\mathbb{E}\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}\}$ (7)
m -MAD model [20]	$\bar{\delta}^{(m)}(\mathbf{x})$ (8)	$\mu(\mathbf{x}) - \bar{\delta}^{(m)}(\mathbf{x})$ (9)
Minimax model [41]	$\Delta(\mathbf{x})$ (11)	$M(\mathbf{x})$ (10)
CVaR model [28]	$\Delta_{\beta}(\mathbf{x})$ (12)	$M_{\beta}(\mathbf{x})$ (13)
GMD model [40]	$\Gamma(\mathbf{x})$ (17)	$\mathbb{E}\{R_{\mathbf{x}} \wedge R_{\mathbf{x}}\}$ (18)

CVaR model was considered with the safety measure [28]. The corresponding risk measure represents the worst conditional semideviation.

Yitzhaki [40] introduced the GMD model with the Gini's mean difference $\Gamma(\mathbf{x})$ but he also analyzed the model implementation with the corresponding safety measure of the mean worse return $\mathbb{E}\{R_{\mathbf{x}} \wedge R_{\mathbf{x}}\}$.

As shown in the previous section, all the risk measures we consider may be derived from the shortfall criteria of SSD. However, they are quite different in modeling of the downside risk aversion. Definitely, the strongest in this respect is the maximum semideviation $\Delta(\mathbf{x})$ used in the Minimax model. It is a strict worst case measure where only the worst scenario is taken into account. The CVaR model allows to extend the analysis to a specified β quantile of the worst returns. The measure of worst conditional semideviation $\Delta_{\beta}(\mathbf{x})$ offers a continuum of models evolving from the strongest downside risk aversion (β close to 0) to the complete risk neutrality ($\beta = 1$). The MAD model with risk measured by the mean (downside) semideviation from the mean is formally a downside risk model. However, due to the symmetry of mean semideviations from the mean [24], it is equally appropriate to interpret the MAD model as an upside risk model. Actually, the m -MAD model has been introduced to incorporate downside risk modeling capabilities into the MAD model. The Gini's mean difference although related to all the worst conditional semideviations is, similar to the mean absolute deviation, a symmetric risk measure (in the sense that for $R_{\mathbf{x}}$ and $-R_{\mathbf{x}}$ it has exactly the same value).

Note that having $\mu(\mathbf{x}') \geq \mu(\mathbf{x}'')$ and $\varrho(\mathbf{x}') \leq \varrho(\mathbf{x}'')$ with at least one inequality strict, one gets $\mu(\mathbf{x}') - \varrho(\mathbf{x}') > \mu(\mathbf{x}'') - \varrho(\mathbf{x}'')$. Hence, a portfolio dominated in the mean-risk model (19) is also dominated in the corresponding mean-safety model (20). In other words, the efficient portfolios of problem (20) form a subset of the entire μ/ϱ -efficient set. Due to the SSD consistency of the safety measures, except for portfolios with identical mean and risk measure, every portfolio belonging to this subset is SSD efficient. Although very important, the SSD efficiency is only a theoretical property. For specific types of distributions or feasible sets the subset of portfolios with guaranteed SSD efficiency may be larger [23, 24]. Hence, the mean-safety model (20) may be too restrictive in some practical investment decisions. In conclusion, we believe the computational testing of the models resulting from all the risk and safety measures summarized in Table 1 is an important part of the models comparison.

3.2 How to solve bicriteria problems

In order to compare on real-life data the performance of various mean-risk models, one needs to deal with specific investor preferences expressed in the models. There are two ways of modeling risk averse preferences and therefore two major approaches to handle bicriteria mean-risk problems (19). First, having assumed a trade-off coefficient λ between the risk and the mean, the so-called *risk aversion coefficient*, one may directly compare real values $\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x})$ and find the best portfolio by solving the optimization problem:

$$\max \{\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x}) : \mathbf{x} \in \mathcal{P}\}. \quad (21)$$

Various positive values of parameter λ allow to generate various efficient portfolios. By solving the parametric problem (21) with changing $\lambda > 0$ one gets the so-called *critical line approach* [18]. Due to convexity of risk measures $\varrho(\mathbf{x})$ with respect to $\mathbf{x}$, $\lambda > 0$ provides the parameterization of the entire set of the μ/ϱ -efficient portfolios (except of its two ends which are the limiting cases). Note that $(1-\lambda)\mu(\mathbf{x}) + \lambda(\mu(\mathbf{x}) - \varrho(\mathbf{x})) = \mu(\mathbf{x}) - \lambda\varrho(\mathbf{x})$. Hence, bounded trade-off $0 < \lambda \leq 1$ in the Markowitz-type mean-risk model (19) corresponds to the complete weighting parameterization of the model (20). The critical line approach allows to select an appropriate value of the risk aversion coefficient λ and the corresponding optimal portfolio through a graphical analysis in the mean-risk image space.

Unfortunately, in practical investment situations, the risk aversion coefficient does not provide a clear understanding of the investor preferences. The commonly accepted approach to implementation of the Markowitz-type mean-risk model is then based on the use of a specified lower bound μ_0 on expected returns which results in the following problem:

$$\min\{\varrho(\mathbf{x}) : \mu(\mathbf{x}) \geq \mu_0, \mathbf{x} \in \mathcal{P}\}. \quad (22)$$

This bounding approach provides a clear understanding of investor preferences and a clear definition of solution portfolios to be used in the models comparison. Therefore, we use the bounding approach (22) in our analysis.

Due to convexity of risk measures $\varrho(\mathbf{x})$ with respect to $\mathbf{x}$, by solving the parametric problem (22) with changing $\mu_0 \in [\min_{j=1,\dots,n} \mu_j, \max_{j=1,\dots,n} \mu_j]$ one gets various efficient portfolios. Actually, for μ_0 smaller than the expected return of the MRP (the minimum risk portfolio defined as solution of $\min_{\mathbf{x} \in \mathcal{P}} \varrho(\mathbf{x})$) problem (22) generates always the MRP while larger values of μ_0 provide the parameterization of the μ/ϱ -efficient set. As a complete parameterization of the entire μ/ϱ -efficient set, the approach (22) generates also those portfolios belonging to the subset of efficient solutions of the corresponding mean-safety problem (20). The latter correspond to larger values of bound μ_0 exceeding the expected return of the portfolio defined as solution of $\max_{\mathbf{x} \in \mathcal{P}} [\mu(\mathbf{x}) - \varrho(\mathbf{x})]$. Thus, opposite to the critical line approach, having a specified value of parameter μ_0 one does not know if the optimal solution of (22) is also an efficient portfolio with respect to the corresponding mean-safety model (20) or not. Therefore, when using the bounding approach to the mean-risk models (19), we need to consider explicitly a separate problem

$$\max\{\mu(\mathbf{x}) - \varrho(\mathbf{x}) : \mu(\mathbf{x}) \geq \mu_0, \mathbf{x} \in \mathcal{P}\} \quad (23)$$

for the corresponding mean-safety model (20).

An alternative approach looks for a risky portfolio offering the maximum increase of the mean return while comparing to the risk-free investment opportunities. Namely having given the risk-free rate of return r_0 one seeks a risky portfolio $\mathbf{x}$ that maximizes the ratio $(\mu(\mathbf{x}) - r_0)/\varrho(\mathbf{x})$. This leads us to the following ratio optimization problem:

$$\max \left\{ \frac{\mu(\mathbf{x}) - r_0}{\varrho(\mathbf{x})} : \mathbf{x} \in \mathcal{P} \right\} \quad (24)$$

The optimal solution of problem (24) is usually called the *tangent portfolio* or the *market portfolio* [7]. Note that clear identification of dispersion type risk measures $\varrho(\mathbf{x})$ for all the LP computable performance measures allows us to define tangent portfolio optimization for all the models. Rather surprisingly the ratio model (24) can be converted into an LP form [39]. Thus the LP computable portfolio optimization models, we consider, remain within LP environment even in the case of ratio criterion used to define tangent portfolio.

Finally, in practice another frequently used approach for models comparison is the analysis of the efficient risk-return frontiers. Notice that this approach strictly depends on the space risk/return used to compare the models.

3.3 The LP models

We provide here the detailed LP formulations for all the models we have analyzed. For each type of model, the pair of problems (22) and (23), we have analyzed, can be stated as:

$$\max\{\alpha\mu(\mathbf{x}) - \varrho(\mathbf{x}) : \mu(\mathbf{x}) \geq \mu_0, \mathbf{x} \in \mathcal{P}\} \quad (25)$$

covering the minimization of risk measure $\varrho(\mathbf{x})$ (22) for $\alpha = 0$ while for $\alpha = 1$ it represents the maximization of the corresponding safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$ (23). Both optimizations are considered with a given lower bound on the expected return $\mu(\mathbf{x}) \geq \mu_0$. All the models for the case of the simplest feasible set $\mathcal{P}$ (1) are summarized in Table 2.

All the models (25) contain the following core LP constraints:

$$\mathbf{x} \in \mathcal{P} \quad \text{and} \quad z \geq \mu_0 \quad (26)$$

$$\sum_{j=1}^n \mu_j x_j = z \quad (27)$$

$$\sum_{j=1}^n r_{jt} x_j = y_t \quad \text{for } t = 1, \dots, T \quad (28)$$

where z is an unbounded variable representing the mean return of the portfolio $\mathbf{x}$ and y_t ($t = 1, \dots, T$) are unbounded variables to represent the realizations of the portfolio return under the scenario t . In addition to these common variables and constraints, each model contains its specific linear constraints to define the risk or safety measure.

MAD models. The standard MAD model [12], when implemented with the mean semideviation as the risk measure ($\varrho(\mathbf{x}) = \bar{\delta}(\mathbf{x})$), leads to the following LP problem:

$$\begin{aligned} & \text{maximize} && \alpha z - z_1 \\ & \text{subject to} && (26)-(28) \text{ and} \\ & && \sum_{t=1}^T p_t d_{1t} = z_1 \end{aligned} \quad (29)$$

$$d_{1t} + y_t \geq z, \quad d_{1t} \geq 0 \quad \text{for } t = 1, \dots, T \quad (30)$$

where nonnegative variables d_{1t} represent downside deviations from the mean under several scenarios t and z_1 is a variable to represent the mean semideviation itself. The latter can be omitted by using the direct formula for mean semideviation in the objective function instead of equation (29). The above LP formulation uses $T + 1$ variables and $T + 1$ constraints to model the mean semideviation.

In order to incorporate downside risk aversion by techniques of the m -MAD model [20], one needs to repeat constraints of type (29)–(30) for each penalty level $k = 2, \dots, m$.

$$\begin{aligned} & \text{maximize} && \alpha z - z_1 - \sum_{k=2}^m w_k z_k \\ & \text{subject to} && (26)-(28), (29)-(30) \text{ and for } k = 2, \dots, m : \\ & && \sum_{t=1}^T p_t d_{kt} = z_k \\ & && d_{kt} + \sum_{i=1}^{k-1} z_i + y_t \geq z, \quad d_{kt} \geq 0 \quad \text{for } t = 1, \dots, T \end{aligned}$$

This results in the LP formulation that uses $m(T + 1)$ variables and $m(T + 1)$ constraints to model the m -level penalized mean semideviation.

CVaR models. For any $0 < \beta < 1$ the CVaR model [28] with $\varrho(\mathbf{x}) = \Delta_\beta(\mathbf{x})$, due to (15), may be implemented as the following LP problem:

$$\begin{aligned} & \text{maximize} && y - (1 - \alpha)z - \frac{1}{\beta} \sum_{t=1}^T p_t d_t \\ & \text{subject to} && (26)-(28) \text{ and } d_t + y_t \geq y, d_t \geq 0 \quad \text{for } t = 1, \dots, T \end{aligned}$$

It is very similar to that of MAD but while defining the downside deviations an independent free variable y is used instead of z representing the mean. Recall that the optimal value of y represents the value of β -quantile. $T + 1$ variables and T constraints are used here to model the worst conditional semideviation.

As the limiting case when β tends to 0 one gets the standard Minimax model [41]. The latter may be additionally simplified by dropping the explicit use of the deviational variables:

$$\begin{aligned} & \text{maximize} && y - (1 - \alpha)z \\ & \text{subject to} && (26)-(28) \text{ and } y_t \geq y \quad \text{for } t = 1, \dots, T \end{aligned}$$

thus resulting in T constraints and a single variable used to model the maximum semideviation.

GMD model. The model with risk measured by the Gini's mean difference ($\varrho(\mathbf{x}) = \Gamma(\mathbf{x})$) [40], according to (17) takes the form:

$$\begin{aligned} & \text{maximize} && \alpha z - \sum_{t'=1}^T \sum_{t'' \neq t'} p_{t'} p_{t''} d_{t' t''} \\ & \text{subject to} && (26)-(28) \text{ and } d_{t' t''} \geq y_{t'} - y_{t''}, d_{t' t''} \geq 0 \quad \text{for } t', t'' = 1, \dots, T; t'' \neq t' \end{aligned}$$

which contains $T(T - 1)$ nonnegative variables $d_{t' t''}$ and $T(T - 1)$ inequalities to define them. However, variables $d_{t' t''}$ are associated with the singleton coefficient columns. Hence, while solving the dual instead of the original primal, the corresponding dual constraints take the form of simple upper bounds (SUB) [21] which are handled implicitly outside the LP matrix. In other words, the dual contains $T(T - 1)$ variables but the number of constraints (excluding the SUB structure) is then proportional to T . Such a dual approach may dramatically improve the LP model efficiency in the case of large number of scenarios. Certainly, one cannot take the advantages of solving dual in the case of some discrete constraints defining the portfolios set $\mathcal{P}$.

As mentioned, an alternative approach to bicriteria mean-risk problem of portfolio selection depends on search for the tangent portfolio which maximizes the ratio $\mu(\mathbf{x}) - r_0 / \varrho(\mathbf{x})$. The corresponding ratio optimization problem (24) can be converted into an LP form by the following transformation: introduce variables $v = \mu(\mathbf{x}) / \varrho(\mathbf{x})$ and $v_0 = 1 / \varrho(\mathbf{x})$, then replace the original decision variables x_j with $\tilde{x}_j = x_j / \varrho(\mathbf{x})$, getting the linear criterion $\max v - r_0 v_0$ and an LP feasible set. Below we demonstrate a sample transformation for the MAD and CVaR models while more formulations is summarized in Table 3. All they are based on the simplest form of feasible set (1) but the transformation itself can easily be applied to any form of the LP feasible set. Once the transformed problem is solved the values of the portfolio variables x_j can be found by dividing $\tilde{x}_j$ by v_0 while $\varrho(\mathbf{x}) = 1 / v_0$ and $\mu(\mathbf{x}) = v / v_0$ (as stated in the last row of Table 3).

MAD ratio model. In the MAD model, risk measure $\varrho(\mathbf{x}) = \bar{\delta}(\mathbf{x})$ is directly represented by variable z_1 defined in equation (29). Hence, the entire MAD ratio model can be written as

$$\max \frac{z - r_0 z_1}{z_1} \quad \text{subject to } \mathbf{x} \in \mathcal{P}, \quad (27)-(28) \text{ and } (29)-(30)$$

Introducing variables $v = z/z_1$ and $v_0 = 1/z_1$ we get linear criterion $v - r_0 v_0$. Further, we divide all the constraints by z_1 and make the substitutions: $\tilde{d}_{1t} = d_{1t}/z_1$, $\tilde{y}_t = y_t/z_1$ for $t = 1, \dots, T$, as well as $\tilde{x}_j = x_j/z_1$, for $j = 1, \dots, n$. Finally, we get the following LP formulation:

$$\begin{aligned}
& \text{maximize} && v - r_0 v_0 \\
& \text{subject to} && \\
& && \sum_{t=1}^T p_t \tilde{d}_{1t} = 1 \\
& && \tilde{d}_{1t} + \tilde{y}_t \geq v, \quad \tilde{d}_{1t} \geq 0 \quad \text{for } t = 1, \dots, T \\
& && \sum_{j=1}^n \mu_j \tilde{x}_j = v \\
& && \sum_{j=1}^n r_{jt} \tilde{x}_j = \tilde{y}_t \quad \text{for } t = 1, \dots, T \\
& && \sum_{j=1}^n \tilde{x}_j = v_0, \quad \tilde{x}_j \geq 0 \quad \text{for } j = 1, \dots, n
\end{aligned}$$

where the last constraints correspond to the set $\mathcal{P}$ definition (1).

CVaR ratio model. In the CVaR model, risk measure $\varrho(\mathbf{x}) = \Delta_\beta(\mathbf{x})$ is not directly represented. We can introduce, however, the equation:

$$z - y + \frac{1}{\beta} \sum_{t=1}^T p_t d_t = z_0$$

allowing us to represent $\Delta_\beta(\mathbf{x})$ with variable z_0 . Introducing variables $v = z/z_0$ and $v_0 = 1/z_0$ we get linear criterion $v - r_0 v_0$ of the corresponding ratio model. Further, we divide all the constraints by z_0 and make the substitutions: $\tilde{d}_t = d_t/z_0$, $\tilde{y}_t = y_t/z_0$ for $t = 1, \dots, T$, as well as $\tilde{x}_j = x_j/z_0$, for $j = 1, \dots, n$ and $\tilde{y} = y/z_0$. Finally, we get the following LP formulation:

$$\begin{aligned}
& \text{maximize} && v - r_0 v_0 \\
& \text{subject to} && \\
& && v - \tilde{y} + \frac{1}{\beta} \sum_{t=1}^T p_t \tilde{d}_t = 1 \\
& && \tilde{d}_t + \tilde{y}_t \geq \tilde{y}, \quad \tilde{d}_t \geq 0 \quad \text{for } t = 1, \dots, T \\
& && \sum_{j=1}^n \mu_j \tilde{x}_j = v \\
& && \sum_{j=1}^n r_{jt} \tilde{x}_j = \tilde{y}_t \quad \text{for } t = 1, \dots, T \\
& && \sum_{j=1}^n \tilde{x}_j = v_0, \quad \tilde{x}_j \geq 0 \quad \text{for } j = 1, \dots, n
\end{aligned}$$

4 Concluding remarks

The classical Markowitz model uses the variance as the risk measure, thus resulting in a quadratic optimization problem. Several alternative risk measures were later introduced which are computationally attractive as (for discrete random variables) they result in solving linear programming

(LP) problems. The LP solvability is very important for applications to real-life financial decisions where the constructed portfolios have to meet numerous side constraints and take into account transaction costs. A gamut of LP solvable portfolio optimization models has been presented in the literature thus generating a need for their classification and comparison. In this paper we have provided a systematic overview of the models with a wide discussion of their theoretical properties. We have shown that all the risk measures used in the LP solvable models can be derived from the SSD shortfall criteria.

The presented formulations show that all the mean-risk models, we consider, can easily be implemented within the LP methodology. In order to implement the corresponding risk measures, the simplest models require only T (the number of scenarios) auxiliary variables and inequalities. The m -MAD model, providing more detailed downside risk modeling capabilities, requires also more complex LP formulations. The number of auxiliary variables and constraints is there multiplied by m (the combination size) but still remain proportional to the number of scenarios T . The Gini's mean difference requires essentially T^2 auxiliary variables and constraints, but taking advantage of the dual formulation allows to reduce the auxiliary structure size. Moreover, we have shown that the LP computable portfolio optimization models, we consider, remain within LP environment even in the case of ratio criterion taking into account the risk-free return.

The portfolio optimization problem considered in this paper follows the original Markowitz' formulation and is based on a single period model of investment. Certainly, the LP computable risk measures can be applied to multi-period problems of portfolio management [6, 26] and to many other financial problems [42]. Similar, the LP portfolio optimization models can be applied together with more complex models for the rates of return. In particular, one may consider they applied to the Sharpe's type models (c.f. [7]) with distinguished nondiversifiable part of risky returns. However, all these extensions exceed the scope of our analysis and they can be considered as potential directions of further research.

Theoretical properties, although crucial for understanding the modeling concepts, provide only a very limited background for comparison of the final optimization models. Computational results are known only for individual models and not all the models were tested in a real-life decision environment. In a future work we will present a comprehensive experimental study comparing practical performances of various LP solvable portfolio optimization models on real-life stock market data.

Acknowledgments

Research conducted by W. Ogryczak was supported by the grant PBZ-KBN-016/P03/99 from The State Committee for Scientific Research (Poland).

References

- [1] P. Artzner, F. Delbaen, J.-M. Eber and D. Heath, Coherent measures of risk, *Mathematical Finance* 9 (1999) 203–228.
- [2] W.J. Baumol, An expected gain-confidence limit criterion for portfolio selection, *Management Science* 10 (1964) 174–182.
- [3] V.S. Bawa, Stochastic Dominance: A Research Bibliography, *Management Science* 28 (1982) 698–712.
- [4] F. Black, E. Derman and W. Toy,(1990), A one-factor model of interest rates and its application to treasury bond options, *Fin. Anal. Journ.*, 33–39.

- [5] T.-J. Chang, N. Meade, J.E. Beasley and Y.M. Sharaiha, Heuristics for cardinality constrained portfolio optimisation, *Computers and Operations Research* 27 (2000) 1271–1302.
- [6] D.R. Carino, D.H. Myers and W.T. Ziemba, Concepts, technical issues and uses of the Russel-Yasuda Kasai financial planning model, *Operations Research* 46 (1998) 450–463.
- [7] E.J. Elton and M.J. Gruber, *Modern Portfolio Theory and Investment Analysis* (Third Edition), Wiley, New York, 1987.
- [8] P.C. Fishburn, Mean–Risk Analysis with Risk Associated with Below Target Returns, *The American Economic Review* 67 (1977) 116–126.
- [9] N.J. Jobst, M.D. Horniman, C.A. Lucas and G. Mitra, Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints, *Quantitative Finance* 1 (2001) 1–13.
- [10] H. Kellerer, R. Mansini and M.G. Speranza, Selecting portfolios with fixed costs and minimum transaction lots, *Annals of Operations Research* 99 (2000) 287–304.
- [11] H. Konno, Piecewise Linear Risk Function and Portfolio Optimization, *Journal of the Operations Research Society of Japan* 33 (1990) 139–156.
- [12] H. Konno and H. Yamazaki, Mean–absolute deviation portfolio optimization model and its application to Tokyo stock market, *Management Science* 37 (1991) 519–531.
- [13] H. Konno and A. Wijayanayake, Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints, *Mathematical Programming* 89 (2001) 233–250.
- [14] W.H. Kruskal and J.M. Tanur (Eds.), *International Encyclopedia of Statistics*, Macmillan, London, 1978.
- [15] H. Levy, Stochastic dominance and expected utility: survey and analysis, *Management Science* 38 (1992) 555–593.
- [16] R. Mansini and M.G. Speranza, Heuristic algorithms for the portfolio selection problem with minimum transaction lots, *European Journal of Operational Research* 114 (1999) 219–233.
- [17] H.M. Markowitz, Portfolio selection, *Journal of Finance* 7 (1952) 77–91.
- [18] H.M. Markowitz, *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, 1959.
- [19] W. Michalowski and W. Ogryczak, A Recursive Procedure for Selecting Optimal Portfolio According to the MAD Model, *Control and Cybernetics* 28 (1999) 725–738.
- [20] W. Michalowski and W. Ogryczak, Extending the MAD portfolio optimization model to incorporate downside risk aversion, *Naval Research Logistics* 48 (2001) 185–200.
- [21] J.L. Nazareth, *Computer Solutions of Linear Programs*, Oxford Univ. Press, New York, 1987.
- [22] W. Ogryczak, Multiple Criteria Linear Programming Model for Portfolio Selection, *Annals of Operations Research* 97 (2000) 143–162.

- [23] W. Ogryczak, Risk Measurement: Mean Absolute Deviation versus Gini's Mean Difference. In: Decision Theory and Optimization in Theory and Practice, G. Wanka (ed.), pp. 33–51, Shaker Verlag, Aachen, 2000.
- [24] W. Ogryczak and A. Ruszczyński, From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures, *European Journal of Operational Research* 116 (1999) 33–50.
- [25] W. Ogryczak and A. Ruszczyński, Dual Stochastic Dominance and Related Mean–Risk Models, *SIAM Journal on Optimization* 13 (2002) 60–78.
- [26] G.Ch. Pflug and A. Świetanowski, Dynamic Asset Allocation under Uncertainty for Pension Fund Management, *Control and Cybernetics* 28 (1999) 755–777.
- [27] R.T. Rockafellar, *Convex Analysis*, Princeton Univ. Press, Princeton, NJ, 1970.
- [28] R.T. Rockafellar and S. Uryasev, Optimization of Conditional Value-at-Risk, *Journal of Risk* 2 (2000) 21–41.
- [29] M. Rothschild and J.E. Stiglitz, Increasing Risk: I. A Definition, *Journal of Economic Theory* 2 (1969) 225–243.
- [30] A.D. Roy, Safety-first and the holding of assets, *Econometrica* 20 (1952) 431–449.
- [31] L.L. Sakalauskas, Nonlinear stochastic programming by Monte-Carlo estimators, *European Journal of Operational Research* 137 (2002) 558–573.
- [32] W.F. Sharpe, A linear programming approximation for the general portfolio analysis problem, *Journal of Financial and Quantitative Analysis* 6 (1971) 1263–1275.
- [33] W.F. Sharpe, Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios, *Management Science* 18 (1971) B1–B13.
- [34] A.F. Shorrocks, Ranking Income Distributions, *Economica* 50 (1983) 3–17.
- [35] Y. Simaan, Estimation risk in portfolio selection: the mean variance model and the mean-absolute deviation model, *Management Science* 43 (1997) 1437–1446.
- [36] F.A. Sortino and H.J. Forsey, On the Use and Misuse of Downside Risk, *Journal of Portfolio Management*, Winter 1996 35–42.
- [37] M.G. Speranza, Linear programming models for portfolio optimization, *Finance* 14 (1993) 107–123.
- [38] G.A. Whitmore and M.C. Findlay (eds.), *Stochastic Dominance: An Approach to Decision-Making Under Risk*, D.C.Heath, Lexington MA, 1978.
- [39] H.P. Williams, *Model Building in Mathematical Programming* (Third Edition), Wiley, Chichester, 1993.
- [40] S. Yitzhaki, Stochastic dominance, mean variance, and Gini's mean difference, *American Economic Review* 72 (1982) 178–185.
- [41] M.R. Young, A minimax portfolio selection rule with linear programming solution, *Management Science* 44 (1998) 673–683.
- [42] S.A. Zenios and P. Kang, Mean–Absolute Deviation Portfolio Optimization for Mortgage-Backed Securities, *Annals of Operations Research* 45 (1993) 433–450.

Table 2: Risk and safety LP models

Model	$\max\{\alpha\mu(\mathbf{x}) - \varrho(\mathbf{x}) : \mu(\mathbf{x}) \geq \mu_0, \mathbf{x} \in \mathcal{P}\}$
Core constraints	$\sum_{j=1}^n x_j = 1$ $x_j \geq 0 \quad \forall j = 1, \dots, n$ $\sum_{j=1}^n \mu_j x_j = z \text{ and } z \geq \mu_0$ $\sum_{j=1}^n r_{jt} x_j = y_t \quad \forall t = 1, \dots, T$
MAD model	$\max \alpha z - z_1$ $\text{s.t. core constraints and}$ $\sum_{t=1}^T p_t d_{1t} = z_1$ $d_{1t} + y_t \geq z, \quad d_{1t} \geq 0 \quad \forall t = 1, \dots, T$
m -MAD model	$\max \alpha z - z_1 - \sum_{k=2}^m w_k z_k$ $\text{s.t. MAD constraints and for } k = 2, \dots, m:$ $\sum_{t=1}^T p_t d_{kt} = z_k$ $d_{kt} + \sum_{i=1}^{k-1} z_i + y_t \geq z, \quad d_{kt} \geq 0 \quad \forall t = 1, \dots, T$
Minimax model	$\max y - (1 - \alpha)z$ $\text{s.t. core constraints and}$ $y_t \geq y \quad \forall t = 1, \dots, T$
CVaR model	$\max y - (1 - \alpha)z - \frac{1}{\beta} \sum_{t=1}^T p_t d_t$ $\text{s.t. core constraints and}$ $d_t + y_t \geq y, \quad d_t \geq 0 \quad \forall t = 1, \dots, T$
GMD model	$\max \alpha z - \sum_{t'=1}^T \sum_{t'' \neq t'} p_{t'} p_{t''} d_{t' t''}$ $\text{s.t. core constraints and}$ $d_{t' t''} \geq y_{t'} - y_{t''}, \quad d_{t' t''} \geq 0 \quad \forall t', t'' = 1, \dots, T; \quad t'' \neq t'$

Table 3: LP formulations of the ratio optimization models

Model	$\max \left\{ \frac{\mu(\mathbf{x}) - r_0}{\varrho(\mathbf{x})} : \mathbf{x} \in \mathcal{P} \right\}$
Core constraints	$\sum_{j=1}^n \tilde{x}_j = v_0$ $\tilde{x}_j \geq 0 \quad \forall j = 1, \dots, n$ $\sum_{j=1}^n \mu_j \tilde{x}_j = v$ $\sum_{j=1}^n r_{jt} \tilde{x}_j = \tilde{y}_t \quad \forall t = 1, \dots, T$
MAD model	$\max v - r_0 v_0$ $\text{s.t. core constraints and}$ $\sum_{t=1}^T p_t \tilde{d}_{1t} = 1$ $\tilde{d}_{1t} + \tilde{y}_t \geq v, \tilde{d}_{1t} \geq 0 \quad \forall t = 1, \dots, T$
m -MAD model	$\max v - r_0 v_0$ $\text{s.t. core constraints and}$ $v_1 + \sum_{k=2}^m w_k v_k = 1$ $\sum_{t=1}^T p_t \tilde{d}_{kt} = v_k \quad \forall k = 1, \dots, m$ $\tilde{d}_{kt} + \sum_{i=1}^{k-1} v_i + \tilde{y}_t \geq v, \tilde{d}_{kt} \geq 0 \quad \forall k = 1, \dots, m; t = 1, \dots, T$
Minimax model	$\max v - r_0 v_0$ $\text{s.t. core constraints and}$ $\tilde{y}_t \geq v - 1 \quad \forall t = 1, \dots, T$
CVaR model	$\max v - r_0 v_0$ $\text{s.t. core constraints and}$ $v - \tilde{y} + \frac{1}{\beta} \sum_{t=1}^T p_t \tilde{d}_t = 1$ $\tilde{d}_t + \tilde{y}_t \geq \tilde{y}, \tilde{d}_t \geq 0 \quad \forall t = 1, \dots, T$
GMD model	$\max v - r_0 v_0$ $\text{s.t. core constraints and}$ $\sum_{t'=1}^T \sum_{t'' \neq t'} p_{t'} p_{t''} \tilde{d}_{t' t''} = 1$ $\tilde{d}_{t' t''} \geq \tilde{y}_{t'} - \tilde{y}_{t''}, \tilde{d}_{t' t''} \geq 0 \quad \forall t', t'' = 1, \dots, T; t'' \neq t'$
Final solution	$x_j = \frac{\tilde{x}_j}{v_0} \quad \forall j = 1, \dots, n, \quad \mu(\mathbf{x}) = \frac{v}{v_0}, \quad \varrho(\mathbf{x}) = \frac{1}{v_0}$