IMA Journal of Management Mathematics (2003) **14**, 187–220

LP solvable models for portfolio optimization: a classification and computational comparison

RENATA MANSINI†
Department of Electronics for Automation, University of Brescia, via Branze 38, 25123 Brescia, Italy

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

AND

M. GRAZIA SPERANZA§
Department of Quantitative Methods, University of Brescia, C. da S.Chiara 48/B, 25122 Brescia, Italy

The Markowitz model of portfolio optimization quantifies the problem in a lucid form 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. The variety of LP solvable portfolio optimization models presented in the literature generates a need for their classification and comparison. It is the main goal of our work. The paper introduces a systematic overview of the LP solvable models with a wide discussion of their theoretical properties. This allows us to classify the models with respect to the types of risk or safety measures they use. The paper also provides the first complete computational comparison of the discussed models on real-life data.

Keywords: portfolio optimization; mean–risk and mean–safety model; linear programming; experimental analysis.

1. Introduction

The portfolio optimization problem considered in this paper follows the original Markowitz formulation which is based on a single period model of investment. At the beginning of a period, an investor allocates his capital among various securities, thus assigning a non-

†Email: rmansini@ing.unibs.it
‡Email: ogryczak@ia.pw.edu.pl
§Email: speranza@eco.unibs.it

negative 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 $J = \{1, 2, \ldots, n\}$ 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 $\mu_j = E\{R_j\}$. Further, let $\mathbf{x} = (x_j)_{j=1,2,\ldots,n}$ 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. $\sum_{j=1}^{n} x_j = 1$ and $x_j \geqslant 0$ for $j = 1, \ldots, n$. 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 non-negative variables.

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 $\mu(\mathbf{x}) = \mathbb{E}\{R_{\mathbf{x}}\} = \sum_{j=1}^{n} \mu_j x_j$. Hence, the mean rate of return is a linear function of portfolio $\mathbf{x}$.

Following the seminal work by Markowitz (1952), the portfolio optimization problem is modelled 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 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 models (Mitra *et al.*, 2003, and references therein). While the original Markowitz model forms a quadratic programming problem, following Sharpe (1971a), many attempts have been made to linearize the portfolio optimization procedure (cf. Speranza, 1993 and references therein). The LP solvability is very important for applications dealing with real-life financial decisions where the constructed portfolios have to meet numerous side constraints, such as minimum transaction lots (Mansini & Speranza, 1999), cardinality constraints (Jobst *et al.*, 2001), and to take into account transaction costs (Kellerer *et al.*, 2000; Konno & Wijayanayake, 2001; Chiodi *et al.*, 2003; Bonaglia *et al.*, 2002).

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 S_t (where $t = 1, \ldots, T$) with corresponding probabilities p_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

$$y_t = \sum_{j=1}^{n} r_{jt} x_j \tag{1}$$

and thus they are linear functions of portfolio $\mathbf{x}$. The expected value $\mu(\mathbf{x})$ can be then expressed as a linear function of the realizations y_t as $\mu(\mathbf{x}) = \sum_{t=1}^{T} y_t p_t$. Similarly, several risk measures can be LP computable with respect to the realizations y_t. The mean absolute deviation was very early considered in portfolio analysis (Sharpe, 1971b) and references therein) while more recently Konno & Yamazaki (1991) presented and analysed the complete portfolio LP solvable optimization model based on this risk measure—the so-called MAD model. Yitzhaki (1982) introduced the mean–risk model using Gini's mean (absolute) difference as the risk measure (the GMD model). Recently, Young (1998) analysed the LP solvable portfolio optimization model based on risk defined by the worst case scenario (the minimax approach), while Ogryczak (2000) introduced the multiple criteria LP model covering all the above as special aggregation techniques. During the achievement of this study, some risk measures for portfolio management have been proposed (Chekhlov *et al.*, 2000), which result in models reducible to LP solvable problems. This is a further evidence of the high interest shown for the subject dealt with in our paper and of the constant evolution of this research domain.

The Markowitz model is frequently criticized as not consistent with axiomatic models of preferences for choice under risk (Rothschild & Stiglitz, 1969). Models consistent with the preference axioms are based on the relations of stochastic dominance or on expected utility theory (Whitmore & Findlay, 1978; Bawa, 1982; Levy, 1992). If the rates of return are normally distributed, then the mean absolute deviation and the Gini mean difference become proportional to the standard deviation $\sigma(\mathbf{x})$ (Kruskal & Tanur, 1978, 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 mean difference (Yitzhaki, 1982), for the MAD model (Ogryczak & Ruszczyński, 1999) and for many other LP solvable models as well (Ogryczak, 2000). Recently, in Artzner *et al.* (1999), a class of coherent risk measures has been defined by means of several axioms. Again, the coherence has been shown for the MAD model (Ogryczak & Ruszczyński, 2002) and for some other LP computable measures (Acerbi & Tasche, 2002).

It is often argued that the variability of the rate of return above the mean should not be penalized since the investors are concerned with an underperformance rather than the overperformance of a portfolio. This led Markowitz (1959) 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 (Bawa, 1978; Fishburn, 1977; Zagst, 2002). Some authors pointed out that the MAD model opens up opportunities for more specific modelling of the downside risk (Feinstein & Thapa, 1993; Speranza, 1993). In fact, most of the LP solvable models may be viewed as based on some downside risk measures. Moreover, the models may be extended with some piecewise linear penalty (risk) functions to provide opportunities for more specific modelling of the downside risk (Carino *et al.*, 1998; Konno, 1990; Michalowski & Ogryczak, 2001).

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 such as SSD consistency (Ogryczak & Ruszczyński, 2001) and the coherence in the sense of Artzner *et al.* (1999). In particular, we classify the performance measures of the models in risk measures (to be minimized) and safety measures (to be maximized). We show that for each risk measure there exists a corresponding well-defined safety measure and vice versa.

Since theoretical results provide only a limited background for models comparison, we also present extensive computational results. The literature provides computational results only for individual models and not all the models were tested in a real-life decision environment. While the MAD model was quite extensively tested (Konno & Yamazaki, 1991) including its application to portfolios of mortgage-backed securities (Zenios & Kang, 1993) where distribution of rate of return is known to be not symmetric, the other LP solvable models seem to get much less recognition from applied studies.

The paper is organized as follows. In the next section we consider the stochastic dominance and the related shortfall criteria. We show how various LP computable performance measures can be derived from the shortfall criteria. Section 3 gives a detailed review and classification of the LP solvable portfolio optimization models we examine. Section 4 is devoted to the experimental analysis on real-life data from the Milan Stock Exchange. Extensive in-sample and out-of-sample computational results are provided and commented on. Finally, some concluding remarks are given.

2. Shortfall criteria and performance measures

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 (Roy, 1952; Bawa, 1978) and later led to the concept of below-target risk measures (Fishburn, 1977; Zagst, 2002) or shortfall criteria. The simplest shortfall criterion for the specific target value τ is the *mean below-target deviation*

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

In the case of returns represented by their realizations, the mean below-target deviation is a convex piecewise linear function of realizations y_t given as $\sum_{t=1}^{T} \max\{\tau - y_t, 0\} p_t$. Hence, due to (1), the mean below-target deviation is also a convex piecewise linear function of the portfolio $\mathbf{x}$ itself and it is LP computable as

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

The concept of mean below-target deviation is related to the second-degree stochastic dominance relation (Whitmore & Findlay, 1978) which is based on an axiomatic model of risk-averse preferences (Rothschild & Stiglitz, 1969; Levy, 1992). 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 $F_\mathbf{x}^{(1)}(\eta) = F_\mathbf{x}(\eta) = \mathbb{P}\{R_\mathbf{x} \leqslant \eta\}$ and it defines the weak relation of the *first-degree stochastic dominance* (FSD) as follows:

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

The second function is derived from the first as

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

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

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

We say that portfolio $\mathbf{x}'$ *dominates* $\mathbf{x}''$ *under the SSD* ($R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$), if $F_{\mathbf{x}'}^{(2)}(\eta) \leqslant F_{\mathbf{x}''}^{(2)}(\eta)$ 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 (Levy, 1992), 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 (Ogryczak & Ruszczyński, 1999):

$$F_{\mathbf{x}}^{(2)}(\eta) = \mathbb{P}\{R_{\mathbf{x}} \leqslant \eta\}\mathbb{E}\{\eta - R_{\mathbf{x}} | R_{\mathbf{x}} \leqslant \eta\} = \mathbb{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\} = \bar{\delta}_{\eta}(\mathbf{x}). \tag{3}$$

Hence, the SSD relation can be seen as a dominance for mean below-target deviations from all possible targets. We call them hereafter the basic SSD shortfall criteria.

The mean below-target deviation from a specific target (2) represents only a single basic SSD shortfall criterion. One may consider several, say m, targets $\tau_1 > \tau_2 > \cdots > \tau_m$ and use the weighted sum of the shortfall criteria as a risk measure

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

where w_k (for $k = 1, \ldots, m$) are positive weights which maintain 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 an increasing and convex piecewise linear penalty function with breakpoints $b_k = \tau_1 - \tau_k$ and slopes $s_k = w_1 + \cdots + w_k$, $k = 1, \ldots, m$. Such a piecewise linear penalty function is used in the Russel–Yasuda–Kasai financial planning model (Carino *et al.*, 1998) to define the corresponding risk measure.

When an investment situation involves minimal acceptable returns, then the below-target deviation and its extensions, as presented in the previous section, are considered to be good risk measures (Fishburn, 1977). However, they are in general not risk relevant as for some targets they may not prevent concentration of risks from remaining undetected. When the mean portfolio return is used to define target achievements, then the corresponding risk measure should relate to shortfalls with respect to the mean $\mu(\mathbf{x})$ rather than to any fixed target τ. We will call such below-mean characteristics downside measures or semideviations (if applicable). In the following sections we show how various possible downside performance measures can be derived from the basic SSD shortfall criteria and that some are consistent with the stochastic dominance relations and are coherent in the sense of Artzner *et al.* (1999). Some of the performance measures are risk measures (to be minimized) and some are safety measures (to be maximized). We show that there are complementary pairs of risk and safety measures. That means, for each risk measure there exists a corresponding safety measure and vice versa. We also show how these measures become LP computable in the case of returns defined by discrete random variables.

2.2 *MAD and downside versions*

Let us simply use the mean portfolio return $\mu(\mathbf{x})$ in the shortfall criterion (2) instead of a fixed target τ. This results in the risk measure known as the *downside mean semideviation* from the mean:

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

The downside mean semideviation is always equal to the upside one $\bar{\delta}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\max\{R_{\mathbf{x}} - \mu(\mathbf{x}), 0\}\}$, therefore we refer to it hereafter as the mean semideviation. Note that the mean semideviation represents both downside as well as upside mean deviations (Kenyon *et al.*, 1999; Ogryczak & Ruszczyński, 1999). Actually, the mean semideviation is a half of the mean absolute deviation from the mean, 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 (Speranza, 1993). For a discrete random variable represented by its realizations, the mean semideviation (5) is a convex piecewise linear function of realizations y_t, given as $\sum_{t=1}^{T} \max\{\mu(\mathbf{x}) - y_t, 0\} p_t$. Hence, due to (1), the mean semideviation is also a convex piecewise linear function of the portfolio $\mathbf{x}$ itself and it is LP computable as

$$\bar{\delta}(\mathbf{x}) = \min \sum_{t=1}^{T} d_t^- p_t \quad \text{subject to} \quad d_t^- \geqslant \mu(\mathbf{x}) - y_t,\ d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T.$$

Due to the use of distribution-dependent target value $\mu(\mathbf{x})$, the mean semideviation cannot be directly considered as a basic SSD shortfall criterion. However, as shown by Ogryczak & Ruszczyński (1999), the mean semideviation is closely related to the graph of $F_{\mathbf{x}}^{(2)}$. The function $F_{\mathbf{x}}^{(2)}$ is continuous, convex, non-negative 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 $(\mu(\mathbf{x}), 0)$ (Fig. 1). Exactly, the η-axis is the left asymptote and the ascent line $\eta - \mu(\mathbf{x})$ is the right asymptote. In the case of a 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. Thus, 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

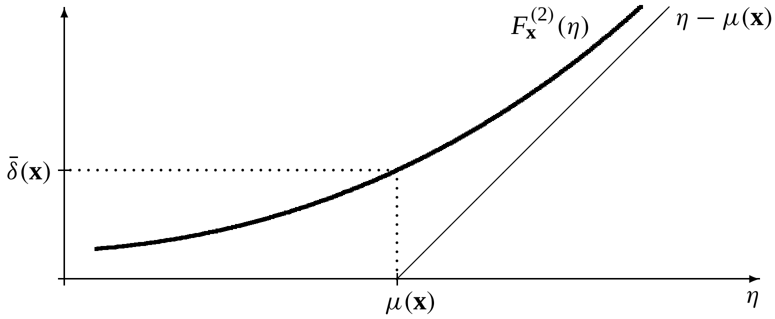


FIG. 1. The O–R diagram and the mean semideviation.

diameter of the dispersion space while the variance represents its doubled area (Ogryczak & Ruszczyński, 1999).

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 $\tau - \bar{\delta}_\tau(\mathbf{x}') \geqslant \tau - \bar{\delta}_\tau(\mathbf{x}'')$ iff $\bar{\delta}_\tau(\mathbf{x}') \leqslant \bar{\delta}_\tau(\mathbf{x}'')$. 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*

$$\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) = \mathbb{E}\{\mu(\mathbf{x}) - \max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}\} \tag{6}$$

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 Ogryczak & Ruszczyński (1999), $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies the inequality $\mu(\mathbf{x}') - \bar{\delta}(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \bar{\delta}(\mathbf{x}'')$ while the corresponding inequality on the mean semideviations $\bar{\delta}(\mathbf{x}') \leqslant \bar{\delta}(\mathbf{x}'')$ 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. In Artzner *et al.* (1999), a class of coherent risk measures has been defined by means of several axioms. In our terms, these measures correspond to composite objectives of form $f(\mathbf{x}) = -\mu(\mathbf{x}) + \varrho(\mathbf{x})$ (note the opposite scalarization via the sign change). The axioms are: translation invariance, positive homogeneity, subadditivity, monotonicity ($R_{\mathbf{x}'} \geqslant R_{\mathbf{x}''} \Rightarrow f(\mathbf{x}') \leqslant f(\mathbf{x}'')$), and relevance ($R_{\mathbf{x}} \leqslant 0, R_{\mathbf{x}} \neq 0 \Rightarrow f(\mathbf{x}) < 0$). As pointed out in Ogryczak & Ruszczyński (2002, Remark 1), $\bar{\delta}(\mathbf{x})$ is seminorm in $\mathcal{L}_1$, is convex and positively homogeneous. Therefore, the composite objective $-\mu(\mathbf{x}) + \bar{\delta}(\mathbf{x})$ does satisfy the first three axioms. Moreover, owing to the consistency with stochastic dominance, it also satisfies monotonicity and relevance, because $R_{\mathbf{x}'} \geqslant R_{\mathbf{x}''} \Rightarrow R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$. Theorems 3 and 4 (see Appendix) generalize this assertion making it applicable to the various LP computable measures we consider.

For better modelling of the downside risk, one may consider a risk measure defined by the mean semideviation applied with a piecewise linear penalty function (Konno, 1990) to penalize larger downside deviations. It turns out, however, that for maintaining both the LP computability and SSD consistency (Michalowski & Ogryczak, 2001), the breakpoints (or additional target values) must be located at the corresponding mean downside underachievements (6). Namely, when using m distribution-dependent targets $\mu_1(\mathbf{x}) = \mu(\mathbf{x}), \mu_2(\mathbf{x}), \ldots, \mu_m(\mathbf{x})$ and the corresponding mean semideviations $\bar{\delta}_1(\mathbf{x}) = \bar{\delta}(\mathbf{x}), \bar{\delta}_2(\mathbf{x}), \ldots, \bar{\delta}_m(\mathbf{x})$ defined recursively according to the formulae:

$$\bar{\delta}_k(\mathbf{x}) = \mathbb{E}\{\max\{\mu_k(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = \mathbb{E}\{\max\{\mu(\mathbf{x}) - \sum_{i=1}^{k-1} \bar{\delta}_i(\mathbf{x}) - R_{\mathbf{x}}, 0\}\}, \qquad (7)$$

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

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

$$\bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}) = \sum_{k=1}^{m} w_k \bar{\delta}_k(\mathbf{x}), \qquad 1 = w_1 \geqslant w_2 \geqslant \cdots \geqslant w_m \geqslant 0, \qquad (8)$$

as in the m-MAD model (Michalowski & Ogryczak, 2001). Actually, the measure can be interpreted as a single mean semideviation (from the mean) applied with a penalty function: $\bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}) = \mathbb{E}\{u(\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\})\}$ where u is an increasing and convex piecewise linear penalty function with breakpoints $b_k = \mu(\mathbf{x}) - \mu_k(\mathbf{x})$ and slopes $s_k = w_1 + \cdots + w_k$, $k = 1, \ldots, m$. Therefore, we will refer to the measure $\bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x})$ as to the *mean penalized semideviation*.

Note that the mean semideviations $\bar{\delta}_k(\mathbf{x})$ defined by the recursive formula (7), in general, may be not convex functions of portfolio $\mathbf{x}$. Nevertheless, the mean penalized semideviation (8) is a convex piecewise linear function of portfolio $\mathbf{x}$ with returns represented by its realizations (1). This follows from the properties of the cumulative deviation function $\bar{\delta}^{(k)}(\mathbf{x}) = \sum_{i=1}^{k} \bar{\delta}_i(\mathbf{x})$ and the restriction used in (8). In the case of returns represented by their realizations (1), $\bar{\delta}^{(1)}(\mathbf{x}) = \bar{\delta}_1(\mathbf{x}) = \bar{\delta}(\mathbf{x})$ is a convex piecewise linear function of $\mathbf{x}$. Due to (7), the following recursive formula is valid:

$$\begin{aligned}\bar{\delta}^{(k)}(\mathbf{x}) &= \bar{\delta}_k(\mathbf{x}) + \bar{\delta}^{(k-1)}(\mathbf{x}) = \mathbb{E}\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, \bar{\delta}^{(k-1)}(\mathbf{x})\}\} \\ &= \sum_{t=1}^{T} \max\{\mu(\mathbf{x}) - y_t, \bar{\delta}^{(k-1)}(\mathbf{x})\} p_t\end{aligned}$$

which justifies $\bar{\delta}^{(k)}(\mathbf{x})$ as a convex piecewise linear function of portfolio $\mathbf{x}$, for any $k \geqslant 1$. Further, the mean penalized semideviation (8) can be expressed as the linear combination of the cumulated deviations:

$$\bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}) = w_m \bar{\delta}^{(m)}(\mathbf{x}) + \sum_{k=1}^{m-1} (w_k - w_{k+1}) \bar{\delta}^{(k)}(\mathbf{x}),$$

where all the coefficients are non-negative. Hence, in the case of returns represented by their realizations, the mean penalized semideviation is a convex piecewise linear function of $\mathbf{x}$.

As defined by a convex piecewise linear function, the penalized mean semideviation is LP computable. Exactly, it can be computed from the following LP problem:

$$\bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}) = \min \sum_{k=1}^{m} w_k z_k \text{ s.t.} \quad z_k = \sum_{t=1}^{T} d_{kt}^- p_t \quad \text{for } k = 1, \ldots, m,$$

$$d_{kt}^- \geqslant \mu(\mathbf{x}) - y_t - \sum_{i=1}^{k-1} z_i, \; d_{kt}^- \geqslant 0 \quad \text{for } t = 1, \ldots, T; \; k = 1, \ldots, m.$$

The mean penalized semideviation (8) defines the corresponding safety measure $\mu(\mathbf{x}) - \bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x})$ which may be expressed directly as the *weighted sum of the mean downside underachievements* $\mu_k(\mathbf{x})$:

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

where the coefficients are non-negative and sum to 1. This safety measure was shown by Michalowski & Ogryczak (2001) to be SSD consistent in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \bar{\delta}_{\mathbf{w}}^{(m)}(\mathbf{x}'')$. Moreover, due to Theorem 3, the corresponding safety measure (its negative) is coherent in the sense of Artzner *et al.* (1999).

2.3 *Minimax and the worst conditional expectation*

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

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

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

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

represents the corresponding risk measure. The latter may be interpreted as the maximal drawdown (Chekhlov *et al.*, 2000). It is also a well defined measure in the O–R diagram (Fig. 1) as it represents the maximum horizontal diameter of the dispersion space. According to (11), the maximum semideviation is a convex piecewise linear function of realizations y_t and, due to (1), it is also a convex piecewise linear function of the portfolio $\mathbf{x}$ itself. Similar to the mean semideviation, it is LP computable as

$$\Delta(\mathbf{x}) = \min \; d_t^- \quad \text{subject to} \quad d_t^- \geqslant \mu(\mathbf{x}) - y_t, \; d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T.$$

The measure $M(\mathbf{x})$ is known to be SSD consistent and it was applied to portfolio optimization by Young (1998). By the use of Theorem 4, one easily gets that $-M(\mathbf{x})$ is a coherent risk measure in the sense of Artzner *et al.* (1999). A natural generalization of the measure $M(\mathbf{x})$ is the worst conditional expectation 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 worst conditional expectation $M_{\frac{k}{T}}(\mathbf{x})$ as the mean return under the k worst scenarios. In general, the *worst conditional expectation* and the *worst conditional semideviation* for any (real) tolerance level $0 < \beta \leqslant 1$ are defined as

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

and

$$\varDelta_\beta(\mathbf{x}) = \mu(\mathbf{x}) - M_\beta(\mathbf{x}) \quad \text{for } 0 < \beta \leqslant 1, \tag{13}$$

respectively, where $F_{\mathbf{x}}^{(-1)}(p) = \inf\{\eta : F_{\mathbf{x}}(\eta) \geqslant p\}$ is the left-continuous inverse of the cumulative distribution function $F_{\mathbf{x}}$. For any $0 < \beta \leqslant 1$, the conditional worst realization $M_\beta(\mathbf{x})$ is an SSD consistent measure. Actually, the conditional worst expectations provide an alternative characterization of the SSD relation (Ogryczak & Ruszczyński, 2002) in the sense of the following equivalence:

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

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 (Rockafellar, 1970), the worst conditional expectation may be defined by the optimization (Ogryczak & Ruszczyński, 2002)

$$M_\beta(\mathbf{x}) = \max_{\eta\in R}\left[\eta - \frac{1}{\beta}F_{\mathbf{x}}^{(2)}(\eta)\right] = \max_{\eta\in R}\left[\eta - \frac{1}{\beta}\mathbb{E}\{\max\{\eta - R_{\mathbf{x}}, 0\}\}\right], \tag{15}$$

where η is a real variable taking the value of β-quantile $Q_\beta(\mathbf{x})$ at the optimum. Formula (15) may be also interpreted as $M_\beta(\mathbf{x}) = \max\{\eta - \frac{1}{\beta}\xi : \xi \geqslant F_{\mathbf{x}}^{(2)}(\eta)\}$. Hence, the worst conditional expectations and the corresponding worst conditional semideviations express the results of the O–R diagram analysis according to a slant direction defined by the slope β (Fig. 2).

For a discrete random variable represented by its realizations y_t, problem (15) becomes an LP:

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

where η is an auxiliary (unbounded) variable. The worst conditional semideviations are then available as the corresponding differences from the mean $\varDelta_\beta(\mathbf{x}) = \mu(\mathbf{x}) - M_\beta(\mathbf{x})$. Alternatively, by using (15) one gets

$$\begin{aligned}\varDelta_\beta(\mathbf{x}) = \mu(\mathbf{x}) - M_\beta(\mathbf{x}) &= \min_{\eta\in R}\ \mathbb{E}\left\{R_{\mathbf{x}} - \eta + \frac{1}{\beta}\max\{\eta - R_{\mathbf{x}}, 0\}\right\}\\ &= \min_{\eta\in R}\ \mathbb{E}\left\{\max\{R_{\mathbf{x}} - \eta, 0\} + \frac{1-\beta}{\beta}\max\{\eta - R_{\mathbf{x}}, 0\}\right\}\end{aligned}$$

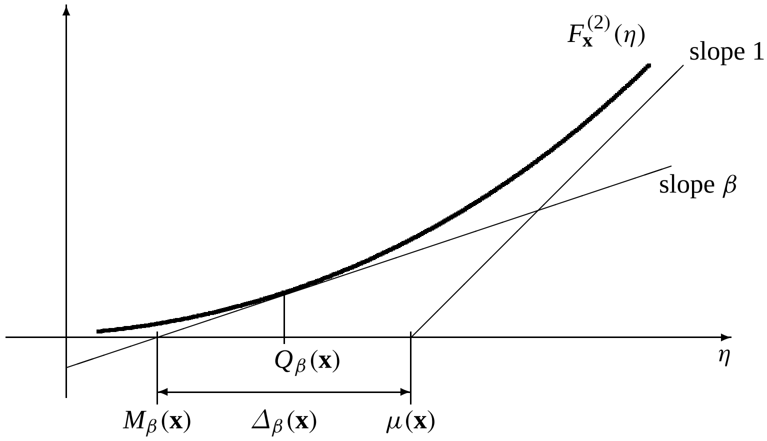


FIG. 2. Quantile safety measures in the O–R diagram.

which allows one to compute the worst conditional semideviation directly from the following LP:

$$\Delta_\beta(\mathbf{x}) = \min \sum_{t=1}^{T} \left(d_t^+ + \frac{1-\beta}{\beta} d_t^- \right) p_t \quad \text{s.t.}$$
$$d_t^- - d_t^+ = \eta - y_t, \quad d_t^+, d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T. \tag{17}$$

Thus, the worst conditional semideviation is a convex piecewise linear function of realizations y_t and, due to (1), it is also a convex piecewise linear function of the portfolio $\mathbf{x}$ itself. It follows from Theorem 3 that $-M_\beta(\mathbf{x})$ is coherent in the sense of Artzner *et al.* (1999).

Note that for $\beta = 0{\cdot}5$ one has $1 - \beta = \beta$. Hence, $\Delta_{0\cdot5}(\mathbf{x})$ represents the mean absolute deviation from the median, the risk measure suggested by Sharpe (1971b). The LP problem for computing this measure takes the form

$$\Delta_{0\cdot5}(\mathbf{x}) = \min \sum_{t=1}^{T} (d_t^+ + d_t^-) p_t \quad \text{s.t.} \quad d_t^- - d_t^+ = \eta - y_t,$$
$$d_t^+, d_t^- \geqslant 0 \quad \text{for } t = 1, \ldots, T.$$

The worst conditional expectation is closely related to the measure called Conditional Concentration (Shalit & Yitzhaki, 1994), Expected Shortfall (Embrechts *et al.*, 1997) or Conditional Value-at-Risk (CVaR) (Rockafellar & Uryasev, 2000) which may be expressed as $\mathrm{CVaR}_\beta(\mathbf{x}) = \mathbb{E}\{R_\mathbf{x} | R_\mathbf{x} \leqslant Q_\beta(\mathbf{x})\}$. Exactly, $M_\beta(\mathbf{x}) = \mathrm{CVaR}_\beta(\mathbf{x})$ in the case of continuous distributions of returns, while they can take different values for discrete distributions (Ogryczak & Ruszczyński, 2002). Nevertheless, recently considered models for portfolio optimization (Rockafellar & Uryasev, 2000) use the LP formula for the worst conditional expectation as a computational approximation to CVaR for continuous distributions. Therefore, the models using the worst conditional expectation or the worst conditional semideviation as a performance measure we will refer to as the CVaR models.

2.4 *Gini mean difference*

Yitzhaki (1982) 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 mean difference*

$$\Gamma(\mathbf{x}) = \frac{1}{2}\sum_{t'=1}^{T}\sum_{t''=1}^{T} |y_{t'} - y_{t''}| p_{t'} p_{t''} \tag{18}$$

is obviously a convex piecewise linear function of realizations y_t and, due to (1), it is also a convex piecewise linear function of the portfolio $\mathbf{x}$. This allows one to compute the Gini mean difference directly from the following LP:

$$\Gamma(\mathbf{x}) = \min \sum_{t'=1}^{T}\sum_{t''\neq t'} d_{t't''} p_{t'} p_{t''} \quad \text{s.t.} \quad d_{t't''} \geqslant y_{t'} - y_{t''},$$
$$d_{t't''} \geqslant 0 \quad \text{for } t', t'' = 1, \dots, T;\ t'' \neq t'.$$

In the case of equally probable T scenarios with $p_t = 1/T$, the Gini mean difference may be expressed as the weighted average of the worst conditional semideviations $\Delta_{\frac{k}{T}}(\mathbf{x})$ for $k = 1, \dots, T$ (Ogryczak, 2000). 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 (18) and from (3),

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

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 $R_{\mathbf{x}}$ (Ogryczak & Ruszczyński, 2002). Thus, although not representing directly any shortfall criterion, the Gini mean difference is a combination of the basic shortfall criteria.

Note that the Gini mean difference defines the corresponding safety measure (Yitzhaki, 1982):

$$\mu(\mathbf{x}) - \Gamma(\mathbf{x}) = \mathbb{E}\{R_{\mathbf{x}} \wedge R_{\mathbf{x}}\} \tag{19}$$

which is the expectation of the minimum of two i.i.d.r.v. $R_{\mathbf{x}}$ thus representing the *mean worse return*. This safety measure is SSD consistent (Yitzhaki, 1982; Ogryczak & Ruszczyński, 2002) in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \Gamma(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \Gamma(\mathbf{x}'')$. Moreover, due to Theorem 3, the safety measure (its negative) is coherent in the sense of Artzner *et al.* (1999).

TABLE 1 *Sample returns*

Scenario	$R_{\mathbf{x}^o}$	$R_{\mathbf{x}'}$	$R_{\mathbf{x}''}$
$\mathbb{P}\{S_1\} = 0{\cdot}5$	1·5	3·5	5·0
$\mathbb{P}\{S_2\} = 0{\cdot}5$	1·5	4·5	4·0

3. Portfolio optimization

3.1 *Risk and safety measures*

Following Markowitz (1952), the portfolio optimization problem is modelled as a mean–risk bicriteria optimization problem:

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

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 an efficient solution of problem (20) or a μ/ϱ*-efficient* portfolio if there is no $\mathbf{x} \in \mathcal{P}$ such that $\mu(\mathbf{x}) \geqslant \mu(\mathbf{x}^0)$ and $\varrho(\mathbf{x}) \leqslant \varrho(\mathbf{x}^0)$ with at least one inequality strict.

The original Markowitz (1952) model 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 vice versa. Although the risk measures are more 'natural', due to the consolidated familiarity with the Markowitz model, we have seen that the safety measures, contrary to the dispersion type risk measures, are SSD consistent in the sense that $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$ implies $\mu(\mathbf{x}') - \varrho(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \varrho(\mathbf{x}'')$ (Michalowski & Ogryczak, 2001; Ogryczak & Ruszczyński, 1999, 2002; Yitzhaki, 1982; Young, 1998). Moreover, one may notice that the safety measures, we consider, satisfy axioms of the so-called coherent risk measurement as in Artzner *et al.* (1999) (with the sign change). We want to emphasize that the convexity of (dispersion type) risk measures is essential for portfolio optimization solvability, while their additional properties of positive homogeneity and appropriate scaling (see Theorem 4) guarantee that the corresponding safety measures are coherent.

The practical consequence of the lack of SSD consistency or the lack of coherence can be illustrated by three portfolios $\mathbf{x}^o$, $\mathbf{x}'$ and $\mathbf{x}''$ with rate of return (given in per cent) under two equally probable scenarios S_1 and S_2 (Table 1). Note that the risk-free portfolio $\mathbf{x}^o$ with the guaranteed result 1·5 is obviously worse than the risky portfolios: $\mathbf{x}'$ giving 3·5 or 4·5 and $\mathbf{x}''$ giving 5·0 or 4·0. Certainly, in all models consistent with the preference axioms

TABLE 2 *Risk and safety measures*

Model	Risk measure $\varrho(\mathbf{x})$		Safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$	
MAD model (Konno & Yamazaki, 1991)	$\bar{\delta}(\mathbf{x})$	(5)	$\mathbb{E}\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}\}$	(6)
m-MAD model (Michalowski & Ogryczak, 2001)	$\bar{\delta}^{(m)}_{\mathbf{w}}(\mathbf{x})$	(8)	$\mu(\mathbf{x}) - \bar{\delta}^{(m)}_{\mathbf{w}}(\mathbf{x})$	(9)
Minimax model (Young, 1998)	$\varDelta(\mathbf{x})$	(11)	$M(\mathbf{x})$	(10)
CVaR model (Rockafellar & Uryasev, 2000)	$\varDelta_{\beta}(\mathbf{x})$	(12)	$M_{\beta}(\mathbf{x})$	(13)
GMD model (Yitzhaki, 1982)	$\varGamma(\mathbf{x})$	(18)	$\mathbb{E}\{R_{\mathbf{x}} \wedge R_{\mathbf{x}}\}$	(19)

of either coherence (Artzner *et al.*, 1999) or SSD (Levy, 1992; Whitmore & Findlay, 1978) portfolio $\mathbf{x}^o$ is dominated by both $\mathbf{x}'$ and $\mathbf{x}''$. When a dispersion type risk measure $\varrho(\mathbf{x})$ is used, then all the portfolios may be efficient in the corresponding mean–risk model. Unfortunately, it also applies to portfolio $\mathbf{x}^o$, since for each such measure $\varrho(\mathbf{x}') > 0$ and $\varrho(\mathbf{x}'') > 0$ while $\varrho(\mathbf{x}^o) = 0$. This is a common flaw of all mean–risk models where risk is measured with some dispersion measure (Markowitz-type models). Further, let us notice that $R_{\mathbf{x}''} \succeq_{SSD} R_{\mathbf{x}'}$ although $R_{\mathbf{x}''} \not\geq R_{\mathbf{x}'}$. Hence, the SSD consistency of a model guarantees that $R_{\mathbf{x}''}$ will be selected while the coherence allows that either $R_{\mathbf{x}''}$ or $R_{\mathbf{x}'}$ may be selected (it only guarantees that $R_{\mathbf{x}^o}$ will not be selected).

It is interesting to note that, in order to overcome this weakness of the Markowitz model, Baumol (1964) suggested considering 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 also consider an alternative mean–safety bicriteria model:

$$\max\{[\mu(\mathbf{x}), \mu(\mathbf{x}) - \varrho(\mathbf{x})] : \quad \mathbf{x} \in \mathcal{P}\}. \tag{21}$$

The full set of risk and safety measures is presented in Table 2. Note that the MAD model was first introduced (Konno & Yamazaki, 1991) 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 (Speranza, 1993). 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 (Young, 1998) with the safety measure of the worst realization $M(\mathbf{x})$ but it was also analysed with the maximum semideviation $\varDelta(\mathbf{x})$ (Ogryczak, 2000). The CVaR model was considered with the safety measure of the worst conditional expectation (Rockafellar & Uryasev, 2000) while the risk measure represents the worst conditional semideviation. Yitzhaki (1982) introduced the GMD model with the Gini mean difference $\varGamma(\mathbf{x})$ but he also analysed 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 basic SSD shortfall criteria. However, they are quite different in their modelling 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 one 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 (Ogryczak & Ruszczyński, 1999), 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 modelling capabilities into the MAD model. The Gini 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 $\mu(\mathbf{x}') \geqslant \mu(\mathbf{x}'')$ and $\varrho(\mathbf{x}') \leqslant \varrho(\mathbf{x}'')$ 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 (20) is also dominated in the corresponding mean–safety model (21). In other words, the efficient portfolios of problem (21) 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 (Ogryczak & Ruszczyński, 1999, 2002) than the corresponding mean–safety efficient set. Hence, the mean–safety model (21) may be too restrictive in some practical investment decisions.

3.2 *Bicriteria portfolio selection*

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 modelling risk-averse preferences and therefore two major approaches to handle bicriteria mean–risk problems (20). 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

$$\max\ \{\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x}) : \mathbf{x} \in \mathcal{P}\}. \tag{22}$$

Various positive values of parameter λ allow one to generate various efficient portfolios. By solving the parametric problem (22) with changing $\lambda > 0$ one gets the so-called *critical line approach* (Markowitz, 1959). Due to the convexity of risk measures $\varrho(\mathbf{x})$ with respect to $\mathbf{x}$, $\lambda > 0$ provides the parametrization 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 $0 < \lambda \leqslant 1$ in the mean–risk model (20) corresponds to the complete weighting parametrization of the mean–safety model (21). The critical line approach allows one 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 mean–risk model is then based on the use of a specified lower bound μ_0 on expected returns which results in the following minimum risk bounded problem:

$$\min\{\varrho(\mathbf{x}): \quad \mu(\mathbf{x}) \geqslant \mu_0, \quad \mathbf{x} \in \mathcal{P}\}. \tag{23}$$

This bounding approach is widely accepted and provides a clear understanding of investor preferences and a clear definition of solution portfolios to be used in the model comparison. Therefore, we use the bounding approach (23) in our analysis.

Due to the convexity of risk measures $\varrho(\mathbf{x})$ with respect to $\mathbf{x}$, by solving the parametric problem (23) 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, the efficient frontier is bounded by the *minimum risk portfolio* (MRP) defined as the solution of $\min_{\mathbf{x}\in\mathcal{P}} \varrho(\mathbf{x})$. For μ_0 smaller than the expected return of the MRP, problem (23) always generates the MRP while larger values of μ_0 provide the parametrization of the μ/ϱ-efficient set as the optimal solution of the fixed return problem

$$\min\ \{\varrho(\mathbf{x}):\ \mu(\mathbf{x}) = \mu_0,\ \mathbf{x} \in \mathcal{P}\} \tag{24}$$

which is then also an optimal solution to (23). This follows from the general properties of a convex bicriteria minimization as shown in Theorem 1 (see Appendix) when applied to $f(\mathbf{x}) = \varrho(\mathbf{x})$.

As a complete parametrization of the entire μ/ϱ-efficient set, the approach (23) also generates portfolios belonging to the subset of efficient solutions of the corresponding mean–safety problem (21). The latter correspond to larger values of bound μ_0 as these portfolios are bounded by the *maximum safety portfolio* (MSP), i.e. the solution to the problem

$$\max\{\mu(\mathbf{x}) - \varrho(\mathbf{x}): \quad \mathbf{x} \in \mathcal{P}\}. \tag{25}$$

Note that, in contrast to the critical line approach, having a specified value of parameter μ_0 does not mean that one knows whether the optimal solution of (23) is also an efficient portfolio with respect to the corresponding mean–safety model (21) or not. Therefore, when using the bounding approach to the mean–risk models (20), essentially, we need to consider explicitly a separate problem of the maximum safety under bounded return

$$\max\{\mu(\mathbf{x}) - \varrho(\mathbf{x}): \quad \mu(\mathbf{x}) \geqslant \mu_0, \quad \mathbf{x} \in \mathcal{P}\} \tag{26}$$

for the corresponding mean–safety model (21). However, the solutions to the bounded maximum safety problem (26) can be found by the analysis of the corresponding minimum risk problem (23), provided that there is already known the MSP. Namely, if $\mu_0 \leqslant \mu(MSP)$, then the MSP is an optimal solution to (26). When $\mu_0 \geqslant \mu(MSP)$, then according to Theorem 2 (see Appendix), the optimal solution of the corresponding problem of risk minimization under fixed return (24) is the optimal solution to both bounded problems: the corresponding minimum risk problem (23) and the maximum safety problem (26).

3.3 *The LP models*

We provide here the detailed LP formulations for all the models we have analysed. For each type of model, the pair of problems (23) and (26) we have analysed can be stated as the problem

$$\max\{\alpha\mu(\mathbf{x}) - \varrho(\mathbf{x}) \ : \ \mu(\mathbf{x}) \geqslant \mu_0, \ \mathbf{x} \in \mathcal{P}\} \tag{27}$$

covering the minimization of risk measure $\varrho(\mathbf{x})$ (23) for $\alpha = 0$ while for $\alpha = 1$ it represents the maximization of the corresponding safety measure $\mu(\mathbf{x}) - \varrho(\mathbf{x})$ (26). Both optimizations are considered with a given lower bound on the expected return $\mu(\mathbf{x}) \geqslant \mu_0$.

By definition, any model (27) contains the following linear constraints:

$$\mathbf{x} \in \mathcal{P} \quad \text{and} \quad z \geqslant \mu_0, \tag{28}$$

where z is an unbounded variable representing the mean return of the portfolio $\mathbf{x}$. Further, all the models contain an equation defining the mean return and explicitly defined realization of the portfolio return, i.e.

$$\sum_{j=1}^{n} \mu_j x_j - z = 0 \quad \text{and} \quad \sum_{j=1}^{n} r_{jt} x_j - y_t = 0 \quad \text{for } t = 1, \ldots, T, \tag{29}$$

where y_t $(t = 1, \ldots, T)$ 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. Note that, in order to use a more standard LP notation and to relate models of the same class, we modify here the notation for some of the variables introduced in Section 2.

MAD and downside versions. The standard MAD model (Konno & Yamazaki, 1991), when implemented with the mean semideviation as the risk measure ($\varrho(\mathbf{x}) = \bar{\delta}(\mathbf{x})$), leads to the following LP problem:

$$\begin{aligned} \text{maximize} \quad & \alpha z - z_1 \\ \text{subject to} \quad & (28)\text{–}(29) \text{ and} \\ & z_1 - \sum_{t=1}^{T} p_t d_{1t} = 0, & (30) \\ & d_{1t} - z + y_t \geqslant 0, \ d_{1t} \geqslant 0 \quad \text{for } t = 1, \ldots, T, & (31) \end{aligned}$$

where non-negative 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 (30). The above LP formulation uses $T + 1$ variables and $T + 1$ constraints to model the mean semideviation.

In the m-MAD model (Michalowski & Ogryczak, 2001) constraints of type (30)–(31) have to be repeated for each penalty level $k = 2, \ldots, m$. This results in the following problem:

$$\begin{aligned}
\text{maximize}\ \ & \alpha z - z_1 - \sum_{k=2}^{m} w_k z_k \\
\text{subject to}\ \ & \text{(28)–(29), (30)–(31) and for } k = 2, \dots, m: \\
& z_k - \sum_{t=1}^{T} p_t d_{kt} = 0, \\
& d_{kt} - z + \sum_{i=1}^{k-1} z_i + y_t \geqslant 0,\ d_{kt} \geqslant 0 \quad \text{for } t = 1, \dots, T.
\end{aligned}$$

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

Minimax and the worst conditional expectation. For any $0 < \beta < 1$ the CVaR model (Rockafellar & Uryasev, 2000) with $\varrho(\mathbf{x}) = \varDelta_\beta(\mathbf{x})$ may be implemented as the following LP problem (the variables d_t^+ which appear in (17) have been substituted in the objective function):

$$\begin{aligned}
\text{maximize}\ \ & y - (1-\alpha)z - \frac{1}{\beta}\sum_{t=1}^{T} p_t d_t \\
\text{subject to}\ \ & \text{(28)–(29) and} \quad d_t - y + y_t \geqslant 0,\ d_t \geqslant 0 \quad \text{for } t = 1, \dots, T.
\end{aligned}$$

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 (Young, 1998). The latter may be additionally simplified by dropping the explicit use of the deviational variables:

$$\begin{aligned}
\text{maximize}\ \ & y - (1-\alpha)z \\
\text{subject to}\ \ & \text{(28)–(29) and} \quad y_t - y \geqslant 0 \quad \text{for } t = 1, \dots, T,
\end{aligned}$$

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

Gini mean difference. The model with risk measured by Gini mean difference ($\varrho(\mathbf{x}) = \varGamma(\mathbf{x})$) (Yitzhaki, 1982), due to the relation $\varGamma(\mathbf{x}) = \mu(\mathbf{x}) - \mathbb{E}\{R_\mathbf{x} \wedge R_\mathbf{x}\}$, may be implemented as follows:

$$\begin{aligned}
\text{maximize}\ \ & (\alpha - 1)z + \sum_{t=1}^{T} p_t^2 y_t + 2\sum_{t'=1}^{T-1}\ \sum_{t''=t'+1}^{T} p_{t'} p_{t''} u_{t't''} \\
\text{subject to}\ \ & \text{(28)–(29) and} \\
& u_{t't''} \leqslant y_{t'},\ u_{t't''} \leqslant y_{t''} \quad \text{for } t' = 1, \dots, T-1;\ t'' = t'+1, \dots, T,
\end{aligned}$$

where $u_{t't''}$ are unbounded variables to represent $\min\{y_{t'}, y_{t''}\}$. The above LP formulation uses $T(T-1)/2$ variables and $T(T-1)$ constraints to model the Gini mean difference.

The direct formulation of the GMD model according to (18) takes the form

$$\text{maximize} \;\; \alpha z - \sum_{t'=1}^{T} \sum_{t'' \neq t'} p_{t'} p_{t''} d_{t't''}$$
$$\text{subject to (28)–(29) and} \quad d_{t't''} \geqslant y_{t'} - y_{t''},\; d_{t't''} \geqslant 0 \quad \text{for } t', t'' = 1, \dots, T;\; t'' \neq t',$$

which contains $T(T-1)$ non-negative variables $d_{t't''}$ and $T(T-1)$ inequalities to define them. The symmetry property $d_{t't''} = d_{t''t'}$ is here ignored and, therefore, the number of variables is doubled in comparison to the previous model. 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) (Nazareth, 1987) 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.

4. Computational results

4.1 *Testing environment*

This section is devoted to a comprehensive analysis and comparison of all the discussed models in a real-life decision environment. In particular, we have compared the practical performance of the models using datasets from Milan Stock Exchange. Tests have been conducted on a PC with the Pentium 200 MHz processor by using the CPLEX 6.5 package (ILOG Inc., 1997). The Barrier solver of CPLEX has been applied to the quadratic programs resulting from the Markowitz model. Below, firstly we describe the analytical framework of the experiments. Secondly, we present and discuss the results of an in-sample analysis, and, thirdly, we provide an extensive out-of-sample comparison of the models. Due to the large number of solved instances, we have summarized and commented only the main results.

Since the measure of conditional maximum 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 ($\beta = 1$), some decisions need to be taken on these parameter values. We have decided to consider two different values, i.e. $\beta = 0{\cdot}1$ and $0{\cdot}5$, in order to compare the corresponding CVaR models versus the standard Minimax and the other models. From now on, we will refer to such models as CVaR(0·1) and CVaR(0·5), respectively. Note that $\beta = 0{\cdot}5$ corresponds to the median and, therefore, the two analysed CVaR models can be considered as downside risk models. The mean penalized semideviation $\bar{\delta}^{(m)}(\mathbf{x})$ of the m-MAD model is defined by the number of penalty levels m and by the weights $1 \geqslant w_2 \geqslant \dots w_m > 0$. The latter may be simplified by the use of a single parameter $0 < a \leqslant 1$ and the power sequence $w_k = a^{k-1}$ for $k = 2, \dots, m$. Larger m (or larger a) implies larger downside risk aversion, while a approaching 0 (or simply $m = 1$) reduces the m-MAD to the standard MAD model. In particular, when $a = 1$ and m tends to infinity the Minimax model (maximum semideviation) is obtained as the limiting case. In order to compare the m-MAD model versus the standard MAD, the selected CVaR models and the

Minimax rule, we have considered the following parameter values: $m = 2$, while $a = 1$ and $a = 0{\cdot}4$. We will identify these models as 2-MAD(1) and 2-MAD(0·4), respectively.

We have prepared four sets of data, consisting of the rates of return of different sets of stocks over periods of about 104 weeks (2 years). Historical realizations have been derived from the stock index prices as listed in the Milan Stock Exchange. In-sample datasets are as follows:

Period A (1994–95): 103 weekly observations and 209 securities available;

Period B (1995–96): 104 weekly observations and 220 securities available;

Period C (1996–97): 105 weekly observations and 235 securities available;

Period D (1997–98): 105 weekly observations and 246 securities available.

The choice of a weekly periodicity for the rates of return is consistent with the requirement of having a large historical sample to reduce estimation errors (see Simaan, 1997). The same number of realizations, but with monthly rates of return, would have implied, for each set, a historical period longer than 8 years. The rates of return have been computed as relative variations of the prices P_{jt}, i.e. $r_{jt} = (P_{j,t+1} - P_{jt})/P_{jt}$; no dividends have been taken into account. Results are reported with annualized returns.

The number of securities available in each in-sample dataset is different, since the number of securities quoted with continuity on the market varies according to the period taken into account. Each security has to meet predefined BSE (Board Stock Exchange) standards to be quoted with continuity on the market. It is worth noticing that, over the four periods, more than 90% of the available securities have been suspended at least once and usually for not less than one week.

Each dataset, corresponding to one of the periods from A to D, has been used to find the mean–risk/safety portfolios through the solution of all the described models including that of Markowitz. The target weekly required return has been set to seven different values (corresponding to the yearly rates 5, 7·5, 10, 12·5, 15, 17·5 and 20%, respectively). Moreover, for each period and model, the MRP and MSP have been computed. The ex-post behaviour of all the selected portfolios has been examined out-of-sample at the end of the 12 month investment periods following the portfolios selection date (the last date of the corresponding in-sample period).

The objective of this section is to provide (when possible) experimental evidence of what has been discussed from a theoretical point of view.

One may expect that some models will show a more aggressive behaviour, typically providing larger returns and lower diversification. This is indeed one of the main results from our computational analysis: while MAD and Markowitz might be classified as the most ‘risk seeking’ models, Minimax and CVaR(0·1) are the most ‘risk averse’ providing portfolios with a stable diversification and lower returns.

4.2 *In-sample analysis*

For each dataset and each level of required rate of return, we have solved all the LP problems defined in Section 3.3 and the Markowitz (mean–variance) model. As an example of information one can get from the analysis of the optimal portfolio selected by a given model, we show Table 3 representing the findings for the Minimax model over Period

TABLE 3 *Minimax model—Period A: optimal portfolios characteristics*

	Min. risk models ($\alpha = 0$)					Max. safety models ($\alpha = 1$)				
μ_0	obj.	z	div.	shares		obj.	z	div.	shares	
%	10^{-2}	%	#	min	max	10^{-3}	%	#	min	max
—	−0·972	1·08	27	2·30 10^{-5}	0·138	−8·804	8·34	25	4·51 10^{-4}	0·129
5·0	−0·989	5·00	28	9·92 10^{-4}	0·122	−8·804	8·34	25	4·51 10^{-4}	0·129
7·5	−1·021	7·50	27	4·53 10^{-4}	0·127	−8·804	8·34	25	4·51 10^{-4}	0·129
10·0	−1·065	10·0	25	2·61 10^{-4}	0·109	−8·817	10·0	25	2·61 10^{-4}	0·109
12·5	−1·111	12·5	26	7·13 10^{-4}	0·109	−8·840	12·5	26	7·13 10^{-4}	0·109
15·0	−1·156	15·0	26	6·14 10^{-4}	0·106	−8·864	15·0	26	6·14 10^{-4}	0·106
17·5	−1·200	17·5	26	1·57 10^{-3}	0·102	−8·891	17·5	26	1·57 10^{-3}	0·102
20·0	−1·243	20·0	27	4·19 10^{-4}	0·102	−8·917	20·0	27	4·19 10^{-4}	0·102

A. Table 3 is divided into two parts: the first corresponds to the problem formulated as minimization of the risk measure ($\alpha = 0$), while the second refers to the maximization of the corresponding safety measure ($\alpha = 1$). Each part consists of five columns showing: the objective function value (obj.), the portfolio per cent average return (z), the portfolio diversification (div.) represented by the number of selected securities, and the minimum and the maximum share within the portfolio, respectively. The average return is reported as converted onto a yearly basis (e.g. $z = 8{\cdot}34\%$ per year is equivalent to a mean return of 0·1542% per week). Each row of the table corresponds to a level of the required return (μ_0). The first row refers to the MRP and to the MSP, respectively (no required return bound). Recall that, for a bound on the expected return larger than the MRP (equal to 1·08% per year), the mean–risk model provides the complete parametrization of the μ/ρ-efficient set. The latter includes the portfolios belonging to the subset of the efficient solutions of the corresponding mean–safety problem, i.e. the portfolios with required return exceeding the mean return of the MSP (8·34% per year), as proven in Section 3.2. The full set of tables for all the models in the four periods, along with extensive details on the in-sample computational results, can be found in Mansini *et al.* (2002).

In Period A (see Table 3) the Minimax model shows a small gap between the return of the MRP and that of the MSP. However, this is not always the case. Table 4 shows the mean returns of the MRP and the MSP, respectively, in the four periods for all the models. The symbol '**' means that the mean return is larger than 10 000% per year. Such large values may be caused by low diversified portfolios, the securities of which have dramatically large weekly mean returns (sometimes close to 100%) over the period. Usually, large weekly rates of return are a direct outcome of a stock quotation suspension for excessive price increase: when newly admitted to quotation, the price of a security may be drastically higher than its last quotation. For instance, the MSP for the MAD model in Period A consists of only three securities (namely, Bintermo, Saiag and Simint) and it returns 17·47% per week, which is greater than 40 000% on a yearly basis. This is due to stock Simint, the average weekly return of which, over Period A, is about 94·25%. This security price during the period 1994–95 moved from a minimum of 0·0054 € to a maximum of 1·54 €. Simint's quotation was suspended several times in Period A, and in one case for more than two weeks. Similarly, the 2063·7% yearly return for the model

TABLE 4 *MRP and MSP mean returns over the four periods (in %)*

	Period A		Period B		Period C		Period D	
	MRP	MSP	MRP	MSP	MRP	MSP	MRP	MSP
Minimax	1·08	8·34	1·60	12·59	10·41	26·17	34·51	51·93
MAD	6·31	**	1·21	**	8·93	301·22	27·39	555·50
2-MAD(0·4)	6·65	2063·7	1·18	**	9·39	257·27	30·34	222·97
2-MAD(1)	5·88	510·63	1·32	343·86	10·86	166·09	28·22	110·09
GMD	0·00	76·44	0·00	427·97	0·00	61·45	8·79	104·25
CVaR(0·1)	3·92	8·29	1·60	12·59	10·41	27·36	35·28	48·35
CVaR(0·5)	6·75	16·06	1·16	101·70	9·10	45·95	27·48	73·65
Markowitz	3·07	344·65	1·39	297·34	8·89	292·05	27·44	409·30

2-MAD(0·4) in period A stems from the high return of the same security. Through the analysis of Table 4 some conclusions on market trend can also be drawn: the MRP return tends to increase over time, being lower than 7% for all the models in Period A and being always larger than 27% (except for GMD) in Period D.

In the following, for the sake of simplicity, we summarize the main figures available for all the selected portfolios.

Table 5 shows, for all the models and all the time periods, the diversification of the optimal portfolios obtained for various required rates of return. For instance, in Period A with $\alpha = 0$, the number of selected securities for the Minimax model ranges between 25 and 28 securities, while with $\alpha = 1$ the corresponding range is 25–27. The number of selected securities only takes into account the stocks with a share larger than or equal to 10^{-6}. On average, we have observed that, when the required return increases, the diversification decreases (with the lowest diversification achieved by the MSP). The Markowitz model provides the ranges with the largest upper limits but it may also result in extremely low diversified portfolios (lower limits from 3 to 5). Some other models (like Minimax and CVaR(0·1)) have larger lower limits showing a more stable diversification. The single security portfolio in Period B for the model CVaR(0·5) corresponds to the MRP and represents an exception, with respect to the diversification, when compared to the portfolios selected by the same model in the remaining three periods. By comparing Table 4 with Table 5 we can conclude that both the MAD and the Markowitz models generate the corresponding MSP with the largest mean return but with the lowest diversification.

Table 6 shows the ranges for the minimum and the maximum share held by stocks for each model in the four periods when $\alpha = 0$ and $\alpha = 1$, respectively. Notice that the minimum shares are scaled with 10^{-4}, while the maximum ones are scaled with 10^{-2} (i.e. expressed in %). On average, efficient portfolios with respect to the mean–safety measures consist of securities with larger minimum share. This is not always the case for the securities with maximum share. Moreover, the Markowitz model and MAD seem to be the models which generate more frequently portfolios with a huge maximum share. This is the case for all the four periods, if we exclude Period B. This suggests that these models might require the introduction of artificial bounding on the maximum share to guarantee a necessary diversification.

In order to compare the structure of the various portfolios, we have analysed the ranking

TABLE 5 *Diversification of optimal portfolios*

	Min. risk models ($\alpha = 0$)				Max. safety models ($\alpha = 1$)			
	A	B	C	D	A	B	C	D
Minimax	25–28	38–44	41–46	20–21	25–27	38–40	41	20
MAD	3–37	2–50	12–42	9–33	3	2	12	9
2-MAD(0·4)	9–36	6–51	15–45	17–30	9	6	15	17
2-MAD(1)	21–38	24–50	31–46	21–33	21	24	31	21
GMD	28–35	29–41	35–48	19–30	28	29	35	21
CVaR(0·1)	30–34	38–44	41–46	19–21	30–31	38–40	44	19
CVaR(0·5)	33–38	1–39	37–46	25–34	33	37	37	25
Markowitz	5–37	5–63	3–63	4–34	5	5	3	4

TABLE 6 *Minimum and maximum shares over the four periods*

	Period A		Period B		Period C		Period D	
	Min $\times 10^{-4}$	Max $\times 10^{-2}$	Min $\times 10^{-4}$	Max $\times 10^{-2}$	Min $\times 10^{-4}$	Max $\times 10^{-2}$	Min $\times 10^{-4}$	Max $\times 10^{-2}$
$\alpha = 0$								
Minimax	0·23–15·7	10·2–13·8	0·11–11·8	12·3–13·6	1·94–12·2	17·8–19·6	6·28	11·7
MAD	0·07–85·4	12·4–81·9	0·10–690	16·3–93·7	0·46–65·0	24·9–32·6	4·99	26·3
2-MAD(0·4)	0·01–34·0	12·4–32·9	0·06–159	16·2–94·3	0·15–31·2	24·1–28·2	0·58	18·5
2-MAD(1)	0·82–19·3	11·8–15·8	0·10–20·8	17·7–92·3	0·60–18·3	18·3–27·3	4·16	20·2
GMD	0·11–6·27	10·2–15·2	0·07–8·31	13·0–91·9	0·11–12·3	8·70–28·9	2·08–102	10·6–26·7
CVaR(0·1)	0·52–66·5	9·4–11·5	0·11–11·8	12·3–13·6	0·44–12·2	15·8–19·5	55·7	13·7
CVaR(0·5)	0·11–1·68	12·3–17·0	0·44–10^4	15·2–100	0·28–11·6	7·30–31·9	0·39	21·5
Markowitz	0·01–101	13·8–80·9	0·01–102	13·1–54·1	0·01–0·09	23·0–88·4	0·01–507	20·5–76·2
$\alpha = 1$								
Minimax	2·61–15·7	10·2–12·9	0·11–11·8	12·3–13·6	2·30	17·8	1·51	12·5
MAD	85·4	81·9	690	93·1	65·0	28·1	2·40	35·4
2-MAD(0·4)	34·0	32·9	159	31·6	31·2	25·0	6·21	17·8
2-MAD(1)	19·3	14·2	20·8	17·9	0·60	18·3	3·01	12·1
GMD	6·27	11·8	8·31	14·9	5·80	8·70	42·5	10·6
CVaR(0·1)	0·70–15·5	10·3–10·7	0·11–11·8	12·3–13·6	0·44	15·8	50·5	11·5
CVaR(0·5)	0·82–1·18	12·3–12·5	1·59	15·2	11·6	7·30	28·6	12·0
Markowitz	101	80·9	102	54·1	0·09	88·4	507	76·2

of the securities which have been selected the most. Table 7 shows the ranking of the first four securities with the largest share in the portfolios selected by the different models over Period A, when the required return is set to 17·5% per year. In each cell the name of the security and its share are given.

Similar tables have been built for the remaining three periods and for different levels of the required rates of return. The ranking of the securities selected by a specific model may vary for different levels of the required return, even within the same period. Nevertheless, some core of the top ranked securities remains quite stable. For instance, with respect to Period A, when $\alpha = 0$ and for all the required rates of return, the maximum share security selected by the Minimax model is always Bpcomin. The only exception is for $\mu_0 = 20\%$; in such case the maximum share is held by Poligraf. As far as the MAD model is concerned, the maximum share security is Bayer for $\mu_0 = 17{\cdot}5\%$ while in all the other cases it is Bpcomin, and similarly for the models 2-MAD(0·4), 2-MAD(0·1) and

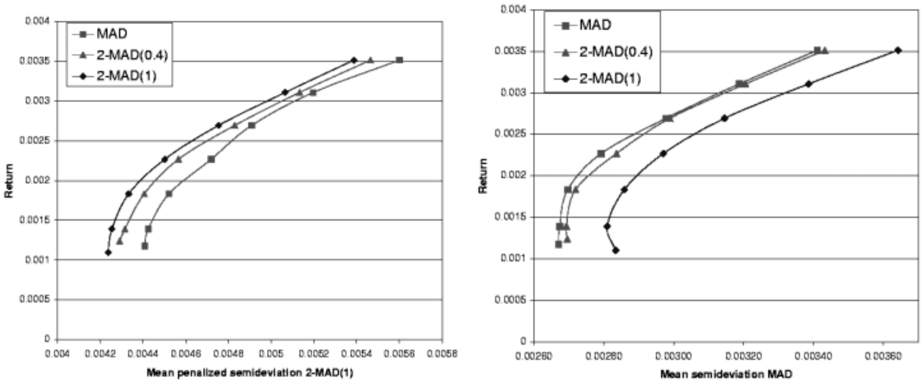
TABLE 7 *Ranking of the first four securities over Period A; required return* 17·5% *per year*

	Security 1		Security 2		Security 3		Security 4	
Minimax	Bpcomin	(0·102)	Poligraf	(0·101)	Pininfrr	(0·0956)	Cbarlett	(0·0809)
MAD	Bayer	(0·128)	Bpcomin	(0·110)	Cbarlett	(0·0974)	Bpintra	(0·0711)
2-MAD(0·4)	Bayer	(0·130)	Bpcomin	(0·106)	Cbarlett	(0·099)	Bpintra	(0·0822)
2-MAD(1)	Cbarlett	(0·119)	Bayer	(0·099)	Bpcomin	(0·098)	Bpintra	(0·076)
GMD	Bayer	(0·111)	Bpcomin	(0·110)	Cbarlett	(0·090)	Bpintra	(0·084)
CVaR(0·1)	Bpcomin	(0·105)	Cbarlett	(0·104)	Bplodi	(0·077)	Poligraf	(0·073)
CVaR(0·5)	Bayer	(0·124)	Cbarlett	(0·111)	Bpcomin	(0·095)	Premudar	(0·062)
Markowitz	Bayer	(0·141)	Crvaltel	(0·096)	Bpintra	(0·082)	Saesgetp	(0·076)

TABLE 8 *Ranking of the four top securities over Period A; required return* 17·5% *per year*

	Bayer		Cbarlett		Bpcomin		Poligraf	
Minimax	no		0·089	(4)	0·102	(1)	0·101	(2)
MAD	0·128	(1)	0·0974	(3)	0·110	(2)	0·018	(18)
2-MAD(0·4)	0·130	(1)	0·099	(3)	0·106	(2)	0·022	(15)
2-MAD(1)	0·098	(2)	0·119	(1)	0·098	(3)	0·031	(12)
GMD	0·111	(1)	0·09	(3)	0·11	(2)	0·0314	(13)
CVaR(0·1)	0·0487	(10)	0·104	(2)	0·105	(1)	0·073	(4)
CVaR(0·5)	0·124	(1)	0·111	(2)	0·095	(3)	0·0197	(16)
Markowitz	0·141	(1)	0·0325	(13)	0·0578	(7)	0·0579	(6)

CVaR(0·5). In 2-MAD(0·1) Cbarlett is the first security only for portfolios with $\mu_0 \geqslant 15\%$, while in CVaR(0·5) Bpcomin is the maximum share security for $\mu_0 \leqslant 12{\cdot}5\%$. For all these models, one of the last ranked securities is Simint. This security has a large mean return over Period A, thus also a small investment may imply a large share of the total portfolio return. In the MSP for the model 2-MAD(1), Simint is ranked at the 11th position, but it is responsible for 81·92% of the total porfolio return. In the portfolio selected by the GMD model the maximum share security is Bpintra when $\alpha = 1$ and Bayer or Bpcomin, depending on the bound on the expected return, when $\alpha = 0$. The model CVaR(0·1) generates, on average, portfolios very similar to those selected by the Minimax model. Finally, portfolios selected by the Markowitz model usually contain a large number of securities with small shares; such securities often cause marginal contributions to the portfolio return. For all the required levels of the rate of return, the Markowitz model chooses Bayer as the first security. In Table 8 we analyse the portfolio position (ranking) of the four securities (namely, Bayer, Cbarlett, Bpcomin and Poligraf) with the largest share out of all the portfolios selected by the different models in Period A for a required return equal to 17·5% per year. The value of the share and the ranking position of the security (if selected) in each portfolio are given. According to Table 8 we can conclude that, with respect to the three most important securities in the portfolio ranking, the models tend to produce similar results. These three top securities cover, however, only about 30% of the total investment.



(a) (b)

FIG. 3. Period A - Efficient frontiers for the MAD-type models.

As an additional insight in models comparison, the efficient frontiers, found by the models over the different periods, can be analysed. In particular, it may be interesting to compare the efficient frontiers obtained by the different models in the same period. In this case the analysis strictly depends on the mean/risk space used to compare the models. With respect to Period A, Fig. 3 shows the efficient frontiers for the models MAD, 2-MAD(1) and 2-MAD(0·4), respectively. In particular, panel (a) of the figure represents the frontiers in the space mean/2-MAD(1), while panel (b) represents those in the space mean/MAD.

For each frontier we have plotted the sequence of points corresponding to the portfolios selected by each model for the seven target required rates of return. Since in Fig. 3(a) the frontiers are represented in the mean/2-MAD(1) space, in order to plot the frontier for the MAD model it has been necessary to compute the value of the 2-level penalized mean semideviation z_2. For the model 2-MAD(0·4), we have simply summed up the values of z_1 and z_2 already available. On the other hand, in Fig. 3(b) the frontiers for the models 2-MAD(1) and 2-MAD(0·4) have been represented by plotting only their z_1 value (i.e. the mean absolute semideviation) as a component of the risk. Notice that the relative position of the frontiers in the two figures is reversed. Hence, the relative position of the different frontiers may be misleading in terms of models comparison.

4.3 *Out-of-sample analysis*

In a real-life environment, model comparisons is usually done by means of ex-post analysis. Several approaches can be used in order to compare models. One of the most commonly applied methods is based on the representation of the ex-post returns of the selected portfolios over a given period and on their comparison against a required level of the return. Unfortunately, the portfolio performances are usually affected by the market trend which makes it very difficult to draw some uniform conclusions. This can be easily seen by comparing the behaviour of the portfolios selected by the same model over the different periods. As an example we show in Figs 4 and 5 the out-of-sample behaviour of the portfolios selected by the different mean–risk models ($\alpha = 0$) with the required

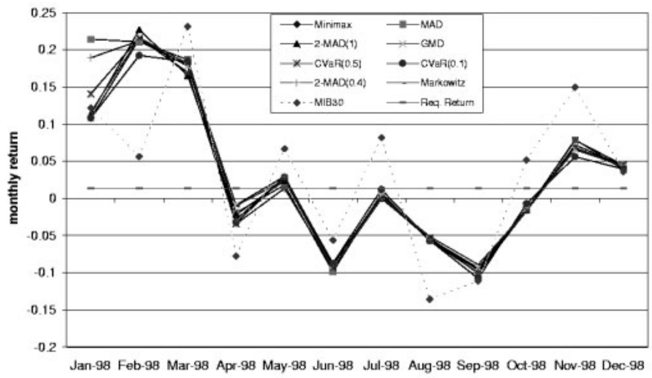


FIG. 4. Ex-post portfolios performances: Period C, required return 17·5% and $\alpha = 0$.

return 17·5% per year in Periods C and D, respectively. In order to take into account the overall market performance, in the two figures the Milan Stock Market Index (MIB30) has been added. This index consists of the 30 most important securities (so-called Blue Chips) quoted at Milan Stock Exchange. Notice that, in both the periods, it may occur that the returns of the selected portfolios are lower than the required level and that it usually happens when the whole market has a negative trend. By comparing the two figures it emerges that, depending on the period, the models may generate portfolios performing very similarly (Period C) or very differently (Period D). In particular, in Period D, it is evident that Minimax and CVaR(0·1) are the unique models which provide highly positive returns when the market index is decreasing, thus confirming their extreme modelling of the downside risk aversion.

Certainly, the models have been applied directly to the original historical data treated as future returns scenarios thus losing the trend information. Possible application of some forecasting procedures prior to the portfolio optimization models, we consider, seems to be an interesting direction for future research. For references on scenarios generation see Carino *et al.* (1998), while on index tracking applications of optimization models see Worzel *et al.* (1994); Consiglio & Zenios (2001); Jobst & Zenios (2003).

We have decided to use some performance criteria to compare different models in the out-of-sample period. For this purpose, we have computed the following nine ex-post parameters:

- the number of times the mean portfolio return is above the required one (symbol #);
- the minimum, average and maximum portfolio return ($r_{\min}$, r_{av} and $r_{\max}$, respectively);
- the standard deviation (std) and the semi-standard deviation (s-std);
- the mean absolute deviation (MAD) and the mean downside semideviation (s-MAD);
- the maximum downside deviation (D-DEV).

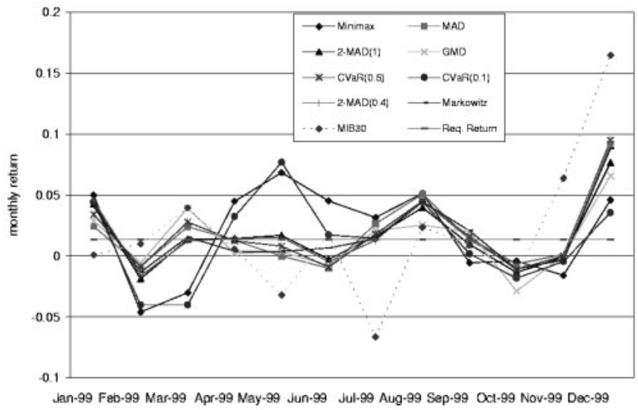


FIG. 5. Ex-post portfolios performances: Period D, required return 17·5% and $\alpha = 0$.

The minimum, maximum and average ex-post portfolio returns have been converted from a monthly to yearly basis. All the dispersion measures (std, s-std, MAD, s-MAD and D-DEV) have been computed with respect to a given target return, thus allowing their direct comparison in the different models. Actually, we focus our analysis on the case of 17·5% target (yearly) return.

In Tables 9–11 we present the average value of each criterion, over the four periods, for several models. In each table we also add a line for the MIB30. This allows a direct comparison of the market index performance with that of the portfolios selected by the other models.

Table 9 shows the ex-post average performances of the optimal portfolios for the corresponding risk minimization models ($\alpha = 0$) with the required return equal to 17·5%. Similarly, Table 10 presents those of the portfolios selected by the models when the required return is equal to 10% per year, and Table 11 those for the MSPs. Recall that all three tables share a common required return level, equal to 17·5%, used as a target return for the ex-post dispersion measures.

Let us start with the analysis of the MIB30 performance. In Tables 9 and 10 the MIB30 index has an average return larger than that of all the other portfolios; at the same time the minimum return is very low and the dispersion measures are impressively larger than those of the other models. While in Table 9 the MIB30 maximum return is the largest, in Table 10 the models MAD, 2-MAD(0·4) and CVaR(0·5) show larger values for such parameter.

In Table 11 the models MAD and 2-MAD(0·4) have an average return larger than the MIB30, but show dispersion measures (s-std, s-MAD and D-DEV) similar in value to those of the MIB30. In general, the MIB30 is the most unstable portfolio with the largest downside deviation. We can easily conclude that any model is preferable to a direct investment in the market index.

Table 9 shows that for all the models the average portfolio returns exceed the required level of 17·5% per year. Exactly, the average returns are varying from 22·46% for the

TABLE 9 *Ex-post criteria average values: required return equal to* 17·5% *and* $\alpha = 0$

	#	$r_{\min}$	r_{av}	$r_{\max}$	std	s-std	MAD	s-MAD	D-DEV
Minimax	5·75	−96·39%	23·93%	271·02%	0·0466	0·0281	0·0385	0·0173	0·0649
MAD	5·5	−77·29%	29·63%	340·91%	0·0477	0·0243	0·0366	0·0144	0·0553
2-MAD(0·4)	5·75	−78·23%	28·41%	341·89%	0·0470	0·0243	0·0365	0·0148	0·0572
2-MAD(1)	6	−77·56%	24·37%	363·28%	0·0431	0·0241	0·0336	0·0146	0·0571
GMD	5·75	−76·47%	24·77%	315·29%	0·0424	0·0241	0·0336	0·0145	0·0567
CVaR(0·1)	5·5	−92·42%	22·46%	277·66%	0·0455	0·0277	0·0362	0·0166	0·0628
CVaR(0·5)	6	−78·70%	26·30%	357·98%	0·0456	0·0242	0·0352	0·0147	0·0559
Markowitz	6	−74·31%	27·02%	337·48%	0·0446	0·0233	0·0341	0·0140	0·0550
MIB30	6·75	−192·21%	35·50%	583·21%	0·0728	0·0419	0·0585	0·0234	0·1027

TABLE 10 *Ex-post criteria average values: required return equal to 10% and* $\alpha = 0$

	#	$r_{\min}$	r_{av}	$r_{\max}$	std	s-std	MAD	s-MAD	D-DEV
Minimax	6	−84·05%	27·75%	334·19%	0·0483	0·0275	0·0398	0·0168	0·0616
MAD	5·25	−66·23%	35·00%	687·13%	0·0569	0·0230	0·0391	0·0142	0·0512
2-MAD(0·4)	5·75	−74·09%	36·03%	701·71%	0·0578	0·0232	0·0392	0·0139	0·0550
2-MAD(1)	5·75	−66·57%	30·24%	487·45%	0·0509	0·0233	0·0361	0·0141	0·0523
GMD	5·5	−79·14%	29·47%	386·20%	0·0497	0·0248	0·0370	0·0149	0·0587
CVaR(0·1)	6	−80·83%	26·07%	340·05%	0·0473	0·0272	0·0376	0·0162	0·0599
CVaR(0·5)	5·75	−66·88%	35·20%	611·25%	0·0550	0·0229	0·0385	0·0138	0·0515
Markowitz	5·75	−73·31%	32·96%	412·81%	0·0517	0·0238	0·0383	0·0144	0·0543
MIB30	6·75	−192·21%	35·50%	583·21%	0·0728	0·0419	0·0585	0·0234	0·1027

CVaR(0·1) model to 29·63% for the MAD model. In contrast to what one might expect the results obtained for models solved with the required return equal to 10% per year (Table 10) turn out to be better than those with 17·5% per year. Portfolios average returns over the four periods for all the models in Table 10 are larger (from 26·07% for CVaR(0·1) to 36·03% for 2-MAD(0·4)) than those in Table 9. Similarly, with respect to the average minimum and maximum returns the portfolios built with the required return 10% outperform those found for 17·5%. Moreover, downside dispersion measures are, on average, lower in Table 10 than in Table 9 (both computed with respect to the target return 17·5%). Although, if we count the number of times (on average) the portfolio return is higher than the target return, then the results in Table 9 are better than Table 10 with the only exceptions of the Minimax and CVaR(0·1) portfolios. In conclusion, the above observations suggest that the required return (bound μ_0) may not be the best way to control ex-post performance.

The better ex-post performances of the portfolios built with a relaxed bound on the required return (see Table 11) suggest possible advantages of the maximum safety portfolios. The latter are built with the expected return maximized as a part of the objective function rather than bounded by a strict constraint. Indeed, for MAD, 2-MAD(0·4) and the Markowitz models, ex-post average returns of the MSPs (Table 11) are even higher than the corresponding performances of the portfolios found with the required return of 10% (Table 10). On the other hand, for the Minimax, CVaR and GMD models, the average returns of the MSPs are worse than those of the portfolios built with the required return

TABLE 11 *Ex-post criteria average values: MSPs*

	#	$r_{\min}$	r_{av}	$r_{\max}$	std	s-std	MAD	s-MAD	D-DEV
Minimax	5	−99·09%	17·18%	244·82%	0·0456	0·0305	0·0380	0·0194	0·0658
MAD	7·5	−244·02%	75·03%	1947·24%	0·0974	0·0420	0·0756	0·0221	0·1088
2-MAD(0·4)	6·5	−159·14%	40·84%	1084·85%	0·0726	0·0361	0·0541	0·0200	0·0885
2-MAD(1)	6	−115·47%	25·53%	527·25%	0·0550	0·0308	0·0414	0·0181	0·0730
GMD	6	−114·13%	21·82%	303·09%	0·0464	0·0280	0·0358	0·0166	0·0706
CVaR(0·1)	5	−97·38%	18·34%	231·75%	0·0459	0·0301	0·0385	0·0191	0·0648
CVaR(0·5)	6·25	−86·87%	25·76%	320·68%	0·0457	0·0254	0·0346	0·0146	0·0612
Markowitz	6·5	−526·37%	33·53%	714·77%	0·0923	0·0587	0·0751	0·0326	0·1567
MIB30	6·75	−192·21%	35·50%	583·21%	0·0728	0·0419	0·0585	0·0234	0·1027

TABLE 12 *Best performance: required return equal to* 17·5% *per year and* $\alpha = 0$

	#	$r_{\min}$	r_{av}	$r_{\max}$	std	s-std	MAD	s-MAD	D-DEV
Minimax	1 (D)		1 (D)		1 (C)		1 (C)		
MAD	1 (B)	1 (D)	1 (C)				1 (B)	1 (A)	1 (D)
2-MAD(0·4)	1 (B)					1 (D)		1 (D)	
2-MAD(1)	2 (B,D)					1 (A)	1 (D)		
GMD	1 (B)	2 (B,C)			2 (B,D)			1 (C)	2 (B,C)
CVaR(0·1)	1 (D)	1 (A)			1 (A)		1 (A)		1 (A)
CVaR(0·5)	1 (B)								
Markowitz	1 (B)					2 (B,C)		1 (B)	
MIB30	3 (A,B,C)		2 (A,B)	4					

equal to 17·5% (Table 9). The MSPs, on average, are characterized by a larger gap between minimum and maximum return as well as larger downside dispersion measures when compared to the portfolios built with the required return equal to 10 or 17·5%. Moreover, the MSPs with the largest average and maximum returns (for the Markowitz, MAD and 2-MAD(0·4) models) are simultaneously characterized by the largest downside deviations, thus generating very unstable results.

Finally, to compare the behaviour of the different models for $\alpha = 0$ and the required return equal to 17·5% per year, in Table 12 we show the number of times, out of the four periods, a given model has found the best performance for each parameter used in the ex-post comparison. The corresponding periods in which the result was achieved are given in parentheses. For instance, the Minimax model found a portfolio with the largest average return only once in Period D. In the first column the total number of entries is greater than four since the highest average value for the corresponding parameter is equally reached by different models in the same period. Notice that the MIB30 index has always the largest maximum return and in two out of four periods also the best average return, however in no periods does it succeed in minimizing a measure of dispersion.

Table 12 can be used as a valid means for ex-post model comparison and may represent a useful tool as support for investors' decisions. Similar results are also available for other levels of the required rate of return.

Concluding remarks

The classical Markowitz model uses variance as the risk measure, thus resulting in a quadratic optimization problem. Several alternative risk measures were introduced thereafter 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 these 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 basic SSD shortfall criteria. This has allowed us to classify the models with respect to the use of risk measures or the corresponding safety measures.

Theoretical properties, although crucial for understanding the modelling 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 have been tested in a real-life decision environment. The second part of this paper has presented a comprehensive experimental study comparing practical performances of the LP solvable portfolio optimization models on real-life stock market data. The efficient frontiers representation is mainly useful for evaluating a portfolio's relative position in a given mean/risk space and not for direct model comparison. Therefore, the experimental analysis has focused on average properties and performances of the models. This allows us to draw several interesting conclusions, some of which may deserve further research.

First of all, our analysis has shown that although the LP solvable models allow one to avoid multiple marginal shares within the optimal portfolio, they usually provide a reasonable diversification. Actually, for various datasets and varying values of the required return bound, our experiments show that many of the LP solvable models provide a more stable diversification than that given by the Markowitz model (see Table 5). In terms of average ex-post performances (Tables 9–11), the MAD type models, similar to the Markowitz one, generate the portfolios with the largest returns but also entailing the largest risk of underachievement (expressed with various downside measures). On the other hand, the GMD and CVaR(0·5) models demonstrate quite good average returns with relatively low risk of underachievement. This suggests further detailed research on a proper parameter selection within the CVaR and the m-MAD models. One may also try to take advantage of the LP models' simplicity by combining the risk criteria of different models to achieve better overall performances.

Further, our analysis shows that a historical period may affect average returns and that all the models (including the Markowitz one) are preferable to a direct investment in the market index. Moreover, the level of the required return does not seem to represent the best way to control ex-post performances, as a lower level may result in higher achievements (Tables 9 and 10). Therefore, the LP solvable models as well as all the mean–risk models, deserve further work on their operational implementations to improve their capabilities to adjust to the investor's preferences and to control effectively the portfolio performances. In our experiments, the models have been applied directly to the original historical data treated as equally probable scenarios of the future return while possible application of some scenario generation procedures (see Carino *et al.*, 1998; Klaassen, 1998; Zenios & McKendall, 1993; Mulvey *et al.*, 2000; Jobst & Zenios, 2004) seems to be a necessary first step toward the operational implementations. Note that the LP solvable models themselves allow one to consider scenarios with different probabilities although the experiments have been limited to the equally probable scenarios. Nevertheless, further work on better ways to control the portfolio selection process within the mean–risk modelling environment remains an important direction for future research.

Acknowledgements

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

REFERENCES

ACERBI, C. & TASCHE, D. (2002) On the coherence of expected shortfall. *J. Banking Finance*, **26**, 1491–1507.

ARTZNER, P., DELBAEN, F., EBER, J.-M. & HEATH, D. (1999) Coherent measures of risk. *Math. Fin.*, **9**, 203–228.

BAUMOL, W. J. (1964) An expected gain-confidence limit criterion for portfolio selection. *Manage. Sci.*, **10**, 174–182.

BAWA, V. S. (1978) Safety-first, stochastic dominance and optimal portfolio choice. *J. Fin. Quant. Anal.*, **13**, 252–271.

BAWA, V. S. (1982) Stochastic dominance: A research bibliography. *Manage. Sci.*, **28**, 698–712.

BONAGLIA, M., CHIARINI, N., MANSINI, R. & SPERANZA, M. G. (2002) An exact approach for the portfolio selection problem with transaction costs and rounds. *Technical Report 2002-07-30* Department of Electronics for Automation, University of Brescia.

CARINO, D. R., MYERS, D. H. & ZIEMBA, W. T. (1998) Concepts, technical issues and uses of the Russel–Yasuda Kasai financial planning model. *Oper. Res.*, **46**, 450–463.

CHEKHLOV, A., URYASEV, S. & ZABARANKIN, M. (2000) Portfolio optimization with drawdown constraints. *Research Report 2000-5* Center for Applied Optimization, DISE, University of Florida.

CHIODI, L., MANSINI, R. & SPERANZA, M. G. (2003) Semi-absolute deviation rule for mutual funds portfolio selection. *Ann. Oper. Res.*, to appear.

CONSIGLIO, A. & ZENIOS, S. A. (2001) Integraded simulation and optimisation models for tracking international fixed income indices. *Math. Program.*, **89**, 311–339.

EMBRECHTS, P., KLÜPPELBERG, C. & MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance*. New York: Springer.

FEINSTEIN, C. D. & THAPA, M. N. (1993) A reformulation of a mean–absolute deviation portfolio optimization model. *Manage. Sci.*, **39**, 1552–1553.

FISHBURN, P. C. (1977) Mean-risk analysis with risk associated with below target returns. *Amer. Econ. Rev.*, **67**, 116–126.

ILOG INC. (1997) *Using the CPLEX Callable Library*. Incline Village: ILOG Inc., CPLEX Div.

JOBST, N. J., HORNIMAN, M. H., LUCAS, C. & MITRA, G. (2001) Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints. *Quant. Fin.*, **1**, 1–13.

JOBST, N. J. & ZENIOS, S. A. (2003) Tracking bond indices in an integrated market and credit risk environment. *Quantitative Finance*, **3**, 117–135.

JOBST, N. J. & ZENIOS, S. A. (2004) On the simulation of portfolios of interest rate and credit risk sensitive securities. *Eur. J. Oper. Res.*, to appear.

KELLERER, K., MANSINI, R. & SPERANZA, M. G. (2000) Selecting portfolios with fixed costs and minimum transaction lots. *Ann. Oper. Res.*, **99**, 287–304.

KENYON, C. M., SAVAGE, S. & BALL, B. (1999) Equivalence of linear deviation about mean and mean absolute deviation about the mean objective functions. *Oper. Res. Lett.*, **24**, 181–185.

KLAASSEN, P. (1998) Financial asset-pricing theory and stochastic programming models for asset-liability management: A synthesis. *Manage. Sci.*, **44**, 31–48.

KONNO, H. (1990) Piecewise linear risk function and portfolio optimization. *J. Oper. Res. Soc. Japan*, **33**, 139–156.

KONNO, H. & YAMAZAKI, H. (1991) Mean–absolute deviation portfolio optimization model and its application to Tokyo stock market. *Manage. Sci.*, **37**, 519–531.

KONNO, H. & WIJAYANAYAKE, A. (2001) Portfolio optimization problem under concave transaction costs and minimal transaction unit constraints. *Math. Program.*, **89**, 233–250.

KRUSKAL, W. H. & TANUR, J. M. (eds) (1978) *International Encyclopedia of Statistics*. London: Macmillan.

LEVY, H. (1992) Stochastic dominance and expected utility: Survey and analysis. *Manage. Sci.*, **38**, 555–593.

MANSINI, R. & SPERANZA, M. G. (1999) Heuristic algorithms for the portfolio selection problem with minimum transaction lots. *Eur. J. Oper. Res.*, **114**, 219–233.

MANSINI, R., OGRYCZAK, W. & SPERANZA, M. G. (2002) LP solvable models for portfolio optimization: a classification and computational comparison. *Technical Report 2002-06-29* Department of Electronics for Automation, University of Brescia.

MARKOWITZ, H. M. (1952) Portfolio selection. *J. Fin.*, **7**, 77–91.

MARKOWITZ, H. M. (1959) *Portfolio Selection: Efficient Diversification of Investments*. New York: Wiley.

MICHALOWSKI, W. & OGRYCZAK, W. (2001) Extending the MAD portfolio optimization model to incorporate downside risk aversion. *Nav. Res. Log.*, **48**, 185–200.

MITRA, G., KYRIAKIS, T., LUCAS, C. & PIRBAI, M. (2003) A review of portfolio planning: Models and systems. *Advances in Portfolio Construction and Implementation*. (S. Satchell & A. Scowcroft, eds). Butterworth-Heinemann Finance, pp. 1–39.

MULVEY, J. M., GOULD, G. & MORGAN, C. (2000) An asset and liability management system for Towers Perrin–Tillinghast. *Interfaces*, **30**, 96–114.

NAZARETH, J. L. (1987) *Computer Solutions of Linear Programs*. New York: Oxford University Press.

OGRYCZAK, W. (2000) Multiple criteria linear programming model for portfolio selection. *Ann. Oper. Res.*, **97**, 143–162.

OGRYCZAK, W. & RUSZCZYŃSKI, A. (1999) From stochastic dominance to mean–risk models: Semideviations as risk measures. *Eur. J. Oper. Res.*, **116**, 33–50.

OGRYCZAK, W. & RUSZCZYŃSKI, A. (2001) On stochastic dominance and mean-semideviation models. *Math. Program.*, **89**, 217–232.

OGRYCZAK, W. & RUSZCZYŃSKI, A. (2002) Dual stochastic dominance and related mean–risk models. *SIAM J. Optimization*, **13**, 60–78.

ROCKAFELLAR, R. T. (1970) *Convex Analysis*. Princeton, MA: Princeton University Press.

ROCKAFELLAR, R. T. & URYASEV, S. (2000) Optimization of conditional value-at-risk. *J. Risk*, **2**, 21–41.

ROTHSCHILD, M. & STIGLITZ, J. E. (1969) Increasing risk: I. A definition. *J. Econ. Theory*, **2**, 225–243.

ROY, A. D. (1952) Safety-first and the holding of assets. *Econometrica*, **20**, 431–449.

SHALIT, H. & YITZHAKI, S. (1994) Marginal conditional stochastic dominance. *Manage. Sci.*, **40**, 670–684.

SHARPE, W. F. (1971a) A linear programming approximation for the general portfolio analysis problem. *J. Fin. Quant. Anal.*, **6**, 1263–1275.

SHARPE, W. F. (1971b) Mean-absolute deviation characteristic lines for securities and portfolios. *Manage. Sci.*, **18**, B1–B13.

SIMAAN, Y. (1997) Estimation risk in portfolio selection: the mean variance model and the mean-absolute deviation model. *Manage. Sci.*, **43**, 1437–1446.

SPERANZA, M. G. (1993) Linear programming models for portfolio optimization. *Finance*, **14**, 107–123.

WHITMORE, G. A. & FINDLAY, M. C. (eds) (1978) *Stochastic Dominance: An Approach to Decision-Making under Risk*. Lexington, KT: Heath.

WORZEL, K. J., VASSIADOU-ZENIOU, C. & ZENIOS, S. A. (1994) Integrated simulation and optimization models for tracking fixed-income indices. *Oper. Res.*, **42**, 223–233.

YITZHAKI, S. (1982) Stochastic dominance, mean variance, and Gini's mean difference. *Amer. Econ. Rev.*, **72**, 178–185.

YOUNG, M. R. (1998) A minimax portfolio selection rule with linear programming solution. *Manage. Sci.*, **44**, 673–683.

ZAGST, R. (2002) *Interest Rate Management*. New York: Springer.

ZENIOS, S. A. & KANG, P. (1993) Mean-absolute deviation portfolio optimization for mortgage-backed securities. *Ann. Oper. Res.*, **45**, 433–450.

ZENIOS, S. A. & MCKENDALL, R. A. (1993) Computing price scenarios of mortgage-backed securities using massively parallel computing. *Modeling Reality and Personal Modelling*. (R. Flavell, ed.). New York: Springer.

Appendix

THEOREM 1 Let $f(\mathbf{x})$ be a convex function of portfolio $\mathbf{x}$ and $\mathbf{x}^f \in \mathcal{P}$ be its global minimizer, i.e. $f(\mathbf{x}^f) \leqslant f(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{P}$. The bounded minimization problem

$$\min \{ f(\mathbf{x}) : \ \mu(\mathbf{x}) \geqslant \mu_0, \ \mathbf{x} \in \mathcal{P} \} \tag{A1}$$

has the following properties:

- if $\mu_0 \leqslant \mu(\mathbf{x}^f)$, then $\mathbf{x}^f$ is an optimal solution to (A1);

- if $\mu_0 \geqslant \mu(\mathbf{x}^f)$, then the optimal solution of the fixed return problem

$$\min \{ f(\mathbf{x}) : \ \mu(\mathbf{x}) = \mu_0, \ \mathbf{x} \in \mathcal{P} \} \tag{A2}$$

  is also an optimal solution to (A1).

Proof. If $\mu_0 \leqslant \mu(\mathbf{x}^f)$, then $\mathbf{x}^f$ is a feasible solution to (A1) and, as a global minimizer, it is optimal.

Let $\mu_0 \geqslant \mu(\mathbf{x}^f)$ and let $\mathbf{x}^0$ be an optimal solution to the corresponding fixed return problem (A2). Consider portfolio $\bar{\mathbf{x}} \in \mathcal{P}$ such that $\mu(\bar{\mathbf{x}}) > \mu_0$ and let us define $\bar{\mathbf{x}}^0 = (1-\lambda)\mathbf{x}^f + \lambda\bar{\mathbf{x}}$ with $\lambda = (\mu(\mathbf{x}^0) - \mu(\mathbf{x}^f))/(\mu(\bar{\mathbf{x}}) - \mu(\mathbf{x}^f))$. By the convexity of set

$\mathcal{P}$, portfolio $\bar{\mathbf{x}}^0$ is feasible and, due to the convexity of function $f(\mathbf{x})$, one gets $f(\bar{\mathbf{x}}^0) \leqslant (1-\lambda)f(\mathbf{x}^f) + \lambda f(\bar{\mathbf{x}}) \leqslant f(\bar{\mathbf{x}})$, since $\mathbf{x}^f$ is the global minimizer. Moreover, $\mu(\bar{\mathbf{x}}^0) = \mu_0$ and therefore $f(\mathbf{x}^0) \leqslant f(\bar{\mathbf{x}}^0) \leqslant f(\bar{\mathbf{x}})$ which proves the optimality of $\mathbf{x}^0$ for problem (A1). □

THEOREM 2 Let $\varrho(\mathbf{x})$ be a convex risk measure and $\mathbf{x}^s \in \mathcal{P}$ be the maximum safety portfolio, i.e. an optimal solution to problem (25). The maximum safety bounded problem (26) has the following properties:

- if $\mu_0 \leqslant \mu(\mathbf{x}^s)$, then the maximum safety portfolio $\mathbf{x}^s$ is an optimal solution to (26);
- if $\mu_0 \geqslant \mu(\mathbf{x}^s)$, then the optimal solution to the corresponding problem of risk minimization under fixed return (24) is the optimal solution to both bounded problems: the corresponding minimum risk problem (23) and the maximum safety problem (26).

Proof. The theorem follows from Theorem 1 applied to $f(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(x)$ which is a convex function. The case of $\mu_0 \leqslant \mu(\mathbf{x}^s)$ is obvious as the minimization of $f(\mathbf{x})$ is equivalent to the safety maximization. For the second case one needs to notice that $\mathbf{x}^0$, the optimal solution to problem (24), due to restriction $\mu(\mathbf{x}) = \mu_0$, is also an optimal solution to the fixed return problem (A2) with the performance function $f(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(\mathbf{x})$. □

THEOREM 3 Let $\varrho(\mathbf{x}) \geqslant 0$ be a convex, positively homogeneous and shift-independent (dispersion type) risk measure. If the measure satisfies additionally the SSD consistency

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad \mu(\mathbf{x}') - \varrho(\mathbf{x}') \geqslant \mu(\mathbf{x}'') - \varrho(\mathbf{x}'')$$

then the corresponding performance function $f(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(\mathbf{x})$ fulfils the coherence axioms (Artzner *et al.*, 1999).

Proof. The axioms are: translation invariance, positive homogeneity, subadditivity, monotonicity ($R_{\mathbf{x}'} \geqslant R_{\mathbf{x}''} \Rightarrow f(\mathbf{x}') \leqslant f(\mathbf{x}'')$), and relevance ($R_{\mathbf{x}} \leqslant 0$, $R_{\mathbf{x}} \neq 0 \Rightarrow f(\mathbf{x}) < 0$). The composite objective $-\mu(\mathbf{x}) + \bar{\delta}(\mathbf{x})$ does satisfy the first three axioms by assumed properties of $\varrho(\mathbf{x})$. Moreover, due to the consistency with stochastic dominance, it also satisfies monotonicity and relevance, because $R_{\mathbf{x}'} \geqslant R_{\mathbf{x}''} \Rightarrow R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$. □

THEOREM 4 Let $\varrho(\mathbf{x}) \geqslant 0$ be a convex, positively homogeneous and shift-independent (dispersion type) risk measure. If the measure additionally meets the risk scaling bound

$$R_{\mathbf{x}} \geqslant 0 \quad \Rightarrow \quad \varrho(\mathbf{x}) \leqslant \mu(\mathbf{x}) \tag{A3}$$

then the corresponding performance function $f(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(\mathbf{x})$ fulfils the coherence axioms (Artzner *et al.*, 1999).

Proof. By assumed properties of $\varrho(\mathbf{x})$, the performance function $f(\mathbf{x}) = \varrho(\mathbf{x}) - \mu(\mathbf{x})$ does satisfy the axioms of translation invariance, positive homogeneity, and subadditivity. Further, if $R_{\mathbf{x}'} \geqslant R_{\mathbf{x}''}$, then $R_{\mathbf{x}'} = R_{\mathbf{x}''} + (R_{\mathbf{x}'} - R_{\mathbf{x}''})$ and $R_{\mathbf{x}'} - R_{\mathbf{x}''} \geqslant 0$. Hence, the subadditivity together with the risk scaling bound (A3) imply that the performance function $f(\mathbf{x})$ satisfies also the axioms of monotonicity and relevance. □