Extending the MAD Portfolio Optimization Model to Incorporate Downside Risk Aversion*

Wojtek Michalowski
Carleton University, Decision Analysis Laboratory, Ottawa, Ontario K1S 5B6, Canada

Włodzimierz Ogryczak†
Warsaw University of Technology, Institute of Control & Computation Engineering, Nowowiejska 15/19, 00–665 Warsaw, Poland

A mathematical model of portfolio optimization is usually represented as a bicriteria optimization problem where a reasonable trade-off between expected rate of return and risk is sought. In a classical Markowitz model, the risk is measured by a variance, thus resulting in a quadratic programming model. As an alternative, the MAD model was developed by Konno and Yamazaki, where risk is measured by (mean) absolute deviation instead of a variance. The MAD model is computationally attractive, since it is easily transformed into a linear programming problem. An extension to the MAD model proposed in this paper allows us to measure risk using downside deviations, with the ability to penalize larger downside deviations. Hence, it provides for better modeling of risk averse preferences. The resulting m–MAD model generates efficient solutions with respect to second degree stochastic dominance, while at the same time preserving the simplicity and linearity of the original MAD model.

Keywords: Portfolio Optimization, Downside Risk, Linear Programming

1. INTRODUCTION

From the advent of Modern Portfolio Theory (MPT) which arouse of the work of Markowitz [16], the notion of investing in diversified portfolios has become one of the most fundamental concepts of portfolio management. The original Markowitz model was derived by using a representative investor belonging to the normative utility framework, which manifested itself in portfolio optimization techniques based on the mean-variance rule. This framework proved to be sufficiently rich to provide the main theoretical background for an analysis of the importance of diversification. It also gave rise to asset pricing models for security pricing, the best known being the Capital Asset Pricing Model (CAPM) [3].

The portfolio optimization problem considered in this paper follows the original Markowitz formulation, and is based on a single period model of investment where at the beginning of each period, an investor would allocate capital among various securities. Assuming that each security is represented by a variable, this is equivalent to assigning a nonnegative weight to each of the variables. During the investment period, a security would generate a certain (random) rate of

* The preliminary version of this work was distributed on-line as the IIASA Interim Report IR-98-041 (June 1998).
† *Correspondence to:* W.Ogryczak

Naval Research Logistics , Vol. 48, No. 3, pp. 185–200(2001)

return. The change of invested capital observed at the end of the period is measured by the weighted average of the individual rates of return. In mathematical terms, to select (optimal) weights reflecting the amount invested in each security, investors need to solve an optimization model consisting of a set of linear constraints, one of which should state that the weights must sum to one (thus reflecting the fact that portions of available total capital are invested in individual securities).

Following Markowitz [16], such a portfolio optimization problem is usually modeled as a bicriteria optimization problem where a reasonable trade–off between the expected rate of return and risk is sought. In the Markowitz model, the risk is measured by a variance from the mean rate of return, thus resulting in the formulation of a quadratic programming model. Following Sharpe [23], many attempts have been made to linearize the portfolio optimization problem (c.f., [27] and references therein). Recently, Konno and Yamazaki [13] presented a MAD portfolio optimization model where risk is measured by (mean) absolute deviation instead of variance. This model is computationally attractive as (for discrete random variables) it results in the solving of linear programming (LP) problems.

It can be argued that the variability of the rate of return above the mean should not be penalized since the investors are concerned with the underperformance rather than the overperformance of a portfolio. This led Markowitz [17] to propose downside risk measures such as (downside) semivariance to replace variance as the risk measure. Consequently, one can observe the growing popularity of the mean return – downside risk portfolio selection models [9]. The absolute deviation used in the MAD model to measure risk is taken as twice the downside semideviation. Therefore, the MAD model may be viewed as based on the downside risk, measured with mean deviation to the mean. Nevertheless, due to a symmetric characteristic of the risk measure used in the MAD model, it can be equally viewed as the "upside risk model".

An investor who uses the MAD model is assumed to have a constant trade-off for a unit deviation from the mean rate of return. This assumption does not allow for the distinction of risk associated with larger losses. The purpose of this paper is to account for such a distinction, and to present an extension to the MAD model that incorporates a "true" downside risk measure.

The Markowitz model has been criticized as not being consistent with axiomatic models of preferences for choice under risk because it does not rely on a relation of stochastic dominance (c.f., [28], [15]). The MAD model is consistent with second degree stochastic dominance, provided that the trade-off coefficient between risk and return is bounded by a certain constant [20]. The proposed extension of the MAD model retains consistency with stochastic dominance.

The paper is organized as follows. In the next section we discuss the original MAD model. Section 3 deals with the proposed extension of MAD, that incorpoates a true downside risk measure. The consistency of the resulting model with stochastic dominance is discussed in Section 4. The paper concludes with a discussion.

2. THE MAD MODEL

Let $J = \{1, 2, \ldots, n\}$ denote a set of securities considered for investment. The rate of return for each security $j \in J$ 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 securities' weights (decision variables) defining a portfolio. To represent a portfolio, the weights must satisfy a set of constraints that form a feasible set Q. The simplest way of defining a feasible set is by a requirement that the weights must sum to one, i.e.:

$$\{\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \ : \ \sum_{j=1}^{n} x_j = 1, \quad x_j \geq 0 \quad \text{for } j = 1, \ldots, n\} \tag{1}$$

An investor must usually consider some other requirements which are expressed as a set of additional side constraints. Hereafter, it is assumed that Q is a general LP feasible set given in a canonical form as a system of linear equations (including (1)) with nonnegative variables:

$$Q = \{\mathbf{x} = (x_1, x_2, \ldots, x_n)^T \ : \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \quad \mathbf{x} \geqq \mathbf{0}\} \tag{2}$$

where $\mathbf{A}$ is a given $p \times n$ matrix and $\mathbf{b} = (b_1, \ldots, b_p)^T$ is a given RHS vector. A vector $\mathbf{x} \in Q$ is called a *portfolio*.

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

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

Following Markowitz [16], the portfolio optimization problem is modeled as a mean–risk optimization problem where $\mu(\mathbf{x})$ is maximized and some risk measure $\varrho(\mathbf{x})$ is minimized. The important advantage of a mean–risk approach is that it provides for the possibility of trade-off analysis. Having assumed a trade-off coefficient λ between the risk and the mean, one may directly compare real values $\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x})$ and find the best portfolio by solving the optimization problem:

$$\max \ \{\mu(\mathbf{x}) - \lambda\varrho(\mathbf{x}) \ : \quad \mathbf{x} \in Q\} \tag{3}$$

This analysis is conducted by way of a so-called *critical line approach* [18], by solving a parametric problem (3) with changing $\lambda > 0$. Such an approach enables to select an appropriate value for the trade-off coefficient λ, and the corresponding optimal portfolio, through graphical analysis in the mean-risk image space.

It is clear that if the risk is measured by a variance:

$$\sigma^2(\mathbf{x}) = E\{(\mu(\mathbf{x}) - R_{\mathbf{x}})^2\} = \sum_{i=1}^{n} \sum_{j=1}^{n} \sigma_{ij} x_i x_j$$

where $\sigma_{ij} = E\{(R_i - \mu_i)(R_j - \mu_j)\}$ is the covariance of securities i and j, then the problem (3) has a quadratic objective function.

The practical use of problem (3) with a quadratic objective function as a tool for optimizing large portfolios is limited. In an attempt to analyze the reasons behind the limited popularity of the Markowitz model among investors, Konno and Yamazaki [13] identify as its shortcomings the need to solve a large scale quadratic programming problem and the investor's reluctance to rely on variance as a measure of risk [14]. An attempt to address these concerns involves considering an alternative risk measure defined as the (mean) *absolute deviation* from a mean:

$$\delta(\mathbf{x}) = E\{|R_{\mathbf{x}} - \mu(\mathbf{x})|\} = \int_{-\infty}^{+\infty} |\mu(\mathbf{x}) - \xi| \ P_{\mathbf{x}}(d\xi) \tag{4}$$

where $P_{\mathbf{x}}$ denotes a probability measure induced by the random variable $R_{\mathbf{x}}$ [21]. Absolute deviation (4) has already been considered by Edgeworth [2] in the context of regression analysis. Within this context, it was used in various areas of decision making resulting among other things in the goal programming formulation of LP problems [1]. Absolute deviation was also considered in the portfolio analysis ([24] and references therein), and recently an interesting approach has been proposed to solve stochastic optimization problems involving this risk measure [4]. Konno and Yamazaki [13] presented the complete portfolio optimization model based on absolute deviation as a risk measure — the so-called MAD model. Using data on stocks comprising the Nikei 225 index, they demonstrated that the MAD model generates results very similar to those obtained with the Markowitz model, while at the same time being easier to solve.

There is no universal risk measure equally good for all broad categories of risk and thus there is a need for caution while using one [26]. For example, when an investment situation involves minimal acceptable returns, then target semivariance and its extensions are considered to be good risk measures [17, 7]. However, when the mean expected return is used as a performance measure, one may then consider extending the above approach by using downside semideviations from the mean. Many authors have pointed out that the MAD model opens up opportunities for more specific modeling of the downside risk (c.f. [12, 5, 29]), because absolute deviation may be considered as a measure of downside risk. Namely, the mean absolute deviation $\delta(\mathbf{x})$ equals twice the (downside) absolute semideviation [20]:

$$\begin{aligned} \bar{\delta}(\mathbf{x}) &= E\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \\ &= E\{\mu(\mathbf{x}) - R_{\mathbf{x}} | R_{\mathbf{x}} \leq \mu(\mathbf{x})\} P\{R_{\mathbf{x}} \leq \mu(\mathbf{x})\} = \int_{-\infty}^{\mu(\mathbf{x})} (\mu(\mathbf{x}) - \xi) \; P_{\mathbf{x}}(d\xi) \end{aligned} \tag{5}$$

The absolute semideviation $\bar{\delta}(\mathbf{x})$ is well defined for any random variable $R_{\mathbf{x}}$ satisfying the condition $E\{|R_{\mathbf{x}}|\} < \infty$ which is true in the portfolio optimization context considered here. According to [13], the following parametric optimization problem is called the MAD model:

$$\max \; \{\mu(\mathbf{x}) - \lambda\bar{\delta}(\mathbf{x}) \; : \quad \mathbf{x} \in Q\} \tag{6}$$

The MAD model is not a true downside risk model, since the absolute semideviation represents both downside as well as upside mean deviations [11, 20]:

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

Therefore, it is equally appropriate to interpret (6) as an upside risk model. Assuming that the MAD model is being interpreted in a downside risk measure framework, one might notice that for any real number η it holds that:

$$\eta - E\{\max\{\eta - R_{\mathbf{x}}, 0\}\} = E\{\min\{R_{\mathbf{x}}, \eta\}\} \tag{8}$$

and

$$\mu(\mathbf{x}) - \lambda\bar{\delta}(\mathbf{x}) = (1-\lambda)\mu(\mathbf{x}) + \lambda(\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x})) = (1-\lambda)\mu(\mathbf{x}) + \lambda E\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}\}$$

This implies, that in the MAD model a convex combination of the original mean and the mean of underachievements (where all larger outcomes are replaced by the mean) is maximized. Therefore,

$0 < \lambda \leq 1$ represents reasonable trade-offs between the mean and the downside risk. The downside risk is measured only by the mean of downside deviations (see (5)), and thus the MAD model assumes a constant trade-off for a unit of downside deviation from the mean portfolio rate of return.

The proposed extension to the MAD model allows one to differentiate between downside and upside risks, and to penalize larger downside deviations. It thus provides for better modeling of risk averse preferences. Note that such an extension is in some ways equivalent to replacing $\bar{\delta}(\mathbf{x})$ with $\bar{\delta}_u(\mathbf{x})$ defined as:

$$\bar{\delta}_u(\mathbf{x}) = E\{u(\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\})\} \tag{9}$$

where u is some convex penalty function.

Simplicity and computational robustness are perceived as the most important advantages of the MAD model. Following [13], it is assumed that r_{jt} is the realization of the random variable R_j during the period t (where $t = 1, \dots, T$) that is available from historical data. It is also assumed that the expected value of R_j can be approximated by:

$$\mu_j = \frac{1}{T} \sum_{t=1}^{T} r_{jt}$$

Therefore, model (6) for a discrete set of realizations r_{jt} can be rewritten as the following LP [5]:

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda}{T} \sum_{t=1}^{T} d_t \tag{10}$$

subject to

$$\mathbf{x} \in Q \tag{11}$$

$$d_t \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \dots, T \tag{12}$$

$$d_t \geq 0 \quad \text{for } t = 1, \dots, T \tag{13}$$

The LP formulation (10)–(13) can be effectively solved, even for large number of securities. Moreover, a diversification of the optimal portfolio (i.e., a number of weights with nonzero values) is controlled by the number T. In the case when Q is given as (1), no more than $T + 1$ securities will be included in the optimal portfolio. The proposed extension to the MAD model, although increasing the problem size, maintains the LP formulation.

If the rates of return are multivariate normally distributed, then the MAD model is equivalent to the Markowitz model [13], although when the return distribution parameters are not known with certainty the former performs slightly worse due to larger estimation errors [25]. However, development of the MAD model does not ask for any specific type of return distributions, facilitating its application to portfolio optimization for mortgage-backed securities and other investment situations where the distribution of rate of return is known not to be symmetric [29].

Recently, the MAD model was further validated by Ogryczak and Ruszczyński [20]. They demonstrated that if the trade-off coefficient λ is bounded by 1, then the MAD model is partially consistent with second degree stochastic dominance [28]. The origins of stochastic dominance are in an axiomatic model of risk-averse preferences [6, 10, 22]. Since that time it has been widely used in economics and finance (see [15] for numerous references). A detailed and comprehensive discussion of stochastic dominance and its relation to downside risk measures is given in [19] and [20]

In stochastic dominance, uncertain returns (random variables) are compared by the pointwise comparison of some performance functions constructed from their distribution functions. Let $R_{\mathbf{x}}$ be a random variable which represents the rate of return for portfolio $\mathbf{x}$ and $P_{\mathbf{x}}$ denote the induced probability measure. The first performance function $F_{\mathbf{x}}^{(1)}$ is defined as the right-continuous cumulative distribution function:

$$F_{\mathbf{x}}^{(1)}(\eta) = F_{\mathbf{x}}(\eta) = P\{R_{\mathbf{x}} \leq \eta\} \quad \text{for real numbers } \eta.$$

The second performance function $F_{\mathbf{x}}^{(2)}$ is derived from the distribution function $F_{\mathbf{x}}$ as:

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

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

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

The corresponding strict SSD relation $\succ_{SSD}$ is defined as

$$R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''} \quad \Leftrightarrow \quad R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \text{and} \quad R_{\mathbf{x}''} \not\succeq_{SSD} R_{\mathbf{x}'}.$$

Thus, we say that portfolio $\mathbf{x}'$ *dominates* $\mathbf{x}''$ *under the SSD* ($R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$), if $F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta)$ for all η, with at least one strict inequality. A feasible portfolio $\mathbf{x}^0 \in Q$ is called *efficient under the SSD* if there is no $\mathbf{x} \in Q$ such that $R_{\mathbf{x}} \succ_{SSD} R_{\mathbf{x}^0}$.

The SSD relation is crucial for decision making under risk. If $R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$, then $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ within all risk-averse preference models where larger outcomes are preferred. Note that the SSD relation covers increasing and concave utility functions, while the first stochastic dominance is less specific as it covers all increasing utility functions [15], thus neglecting a risk averse attitude. It is therefore a matter of primary importance that a model for portfolio optimization be consistent with the SSD relation, which implies that the optimal portfolio is efficient under SSD.

The necessary condition for SSD relation is (c.f. [8]):

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad \mu(\mathbf{x}') \geq \mu(\mathbf{x}'')$$

Ogryczak and Ruszczyński [20] modified this relation to consider absolute semideviations, and proved the following proposition:

PROPOSITION 1 (Ogryczak and Ruszczyński) [20]: If $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$, then $\mu(\mathbf{x}') \geq \mu(\mathbf{x}'')$ and $\mu(\mathbf{x}') - \bar{\delta}(\mathbf{x}') \geq \mu(\mathbf{x}'') - \bar{\delta}(\mathbf{x}'')$, where the second inequality is strict whenever $\mu(\mathbf{x}') > \mu(\mathbf{x}'')$.

The assertion of Proposition 1 together with relation (8) leads to the following corollary (see [20], for details):

COROLLARY: Except for portfolios with identical mean and absolute semideviation, every portfolio $\mathbf{x} \in Q$ that is maximal by $\mu(\mathbf{x}) - \lambda\bar{\delta}(\mathbf{x})$ with $0 < \lambda \leq 1$ is efficient under SSD.

It follows from this Corollary that the unique optimal solution of model (6) with the trade-off coefficient $0 < \lambda \leq 1$ is efficient under SSD. In the case of multiple optimal solutions, one of them is efficient under SSD, but some of them may also be SSD dominated. Due to the Corollary, an optimal portfolio $\mathbf{x}' \in Q$ can be SSD dominated only by another optimal portfolio $\mathbf{x}'' \in Q$ such that $\mu(\mathbf{x}'') = \mu(\mathbf{x}')$ and $\bar{\delta}(\mathbf{x}'') = \bar{\delta}(\mathbf{x}')$. Although the MAD model is consistent with SSD for bounded trade-offs, it requires an additional specification if one wants to maintain the SSD efficiency for every optimal portfolio. The extension of the MAD model presented in this paper provides such a specification.

3. EXTENDED MAD MODEL

The MAD model does not properly account for the risk aversion attitude. In order to do so, one would need to differentiate between the various levels of downside deviations, and to penalize the "larger" ones. Konno [12] has already proposed such an extension of the MAD model for portfolio optimization. He considered additional mean deviations from some target rate of return predefined as being proportional to the mean rate of return. Within the framework of downside risk (and downside deviations) this may be interpreted as consideration of the following deviations:

$$\bar{\delta}_\kappa(\mathbf{x}) = E\{\max\{\kappa\ \mu(\mathbf{x}) - R_\mathbf{x}, 0\}\} \quad \text{for } 0 \leq \kappa \leq 1 \tag{14}$$

For $\kappa = 1$ one gets the $\bar{\delta}_1(\mathbf{x}) = \bar{\delta}(\mathbf{x})$, namely the absolute semideviation used in the original MAD model. One may attempt to augment the downside risk measure by penalizing additional deviations for several $\kappa < 1$. In terms of a penalty function (see (9)), this approach is equivalent to introducing a convex piecewise linear function with breakpoints proportional to the mean of $R_\mathbf{x}$.

Let us focus on the model with one additional downside deviation as Konno [12] did:

$$\max\ \{\mu(\mathbf{x}) - \lambda\bar{\delta}(\mathbf{x}) - \lambda_\kappa\bar{\delta}_\kappa(\mathbf{x}) \ :\ \ \mathbf{x} \in Q\} \tag{15}$$

where $\lambda > 0$ is the basic trade-off parameter and $\lambda_\kappa > 0$ is an additional parameter (a penalty for larger deviations). We refer to this model as κ–MAD.

Note that in the κ–MAD model one penalizes deviations that are relatively large with respect to the expected rate of return. However, the model performs correctly (i.e. $\kappa\mu(\mathbf{x}) \leq \mu(\mathbf{x})$) only in the case of a nonnegative mean. This is true for a typical portfolio optimization problem, but in general, one must be very cautious while trying to apply the κ–MAD model, especially because the deviations $\bar{\delta}_\kappa(\mathbf{x})$ are sensitive to any shift in the scale of measurement.

Konno [12] did not analyze the consistency of the κ–MAD model with stochastic dominance, and such a comprehensive analysis is also beyond the scope of this paper. Nevertheless, one can see that for SSD consistency a proper selection of the parameters in κ–MAD may be quite a difficult task. We illustrate this with a simple example.

Consider two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as:

$$P\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 1/(1+\varepsilon), & \xi = 0 \\ \varepsilon/(1+\varepsilon), & \xi = 1 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad P\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 1, & \xi = 0 \\ 0, & \text{otherwise} \end{cases} \tag{16}$$

where ε is an arbitrarily small positive number. Note that $R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$ and $\mu(\mathbf{x}') = \varepsilon/(1+\varepsilon)$, $\bar{\delta}(\mathbf{x}') = \varepsilon/(1+\varepsilon)^2$ while $\mu(\mathbf{x}'') = \bar{\delta}(\mathbf{x}'') = 0$. It is easy to show that $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ in the MAD model with any $0 < \lambda \leq 1$. Consider now κ–MAD with $\kappa = 0.5$ as suggested in [12]. Then

$\bar{\delta}_{0.5}(\mathbf{x}') = (0.5\varepsilon)/(1+\varepsilon)^2 = 0.5\bar{\delta}(\mathbf{x}')$. Hence, the objective function of the κ–MAD model for $R_{\mathbf{x}'}$ is $\mu(\mathbf{x}') - (\lambda + 0.5\lambda_{0.5})\bar{\delta}(\mathbf{x}')$ which means that $\lambda_{0.5}$ impacts only the value of trade-off coefficient λ. Thus, in case of $\lambda_{0.5} \geq 2(1-\lambda+\varepsilon)$, $R_{\mathbf{x}''}$ is preferred to $R_{\mathbf{x}'}$. This inconsistency of κ–MAD with SSD is rectified in the proposed extension of the MAD model.

Let us begin with the original MAD model, assuming that the trade-off coefficient (λ) has the value τ_1. Since the mean deviation is already considered in (6), it is quite natural to focus on this part of large deviations that exceed the mean deviation (later referred to as "surplus deviations"). Mean surplus deviation $E\{\max\{\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) - R_{\mathbf{x}}, 0\}\}$ needs to be penalized by a value, let's say τ_2, of a trade-off between the surplus deviation and a mean deviation, which leads to the maximization of:

$$\mu(\mathbf{x}) - \tau_1(\bar{\delta}(\mathbf{x}) + \tau_2 E\{\max\{\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}) - R_{\mathbf{x}}, 0\}\})$$

Consequently, because surplus deviations are again measured by their mean, one might wish to penalize "second level" surplus deviations exceeding that mean. This can be formalized as follows:

$$\max\ \{\mu(\mathbf{x}) - \sum_{i=1}^{m} (\prod_{k=1}^{i} \tau_k)\bar{\delta}_i(\mathbf{x}) \ :\quad \mathbf{x} \in Q\} \tag{17}$$

where $\tau_1 > 0, \ldots, \tau_m > 0$ are the assumed known trade-off coefficients and

$$\begin{aligned} \bar{\delta}_1(\mathbf{x}) &= \bar{\delta}(\mathbf{x}) = E\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \\ \bar{\delta}_i(\mathbf{x}) &= E\{\max\{\mu(\mathbf{x}) - \sum_{k=1}^{i-1} \bar{\delta}_k(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \quad \text{for } i = 2, \ldots, m \end{aligned}$$

It will be shown further that although formulated in a recursive manner, yet the problem (17) remains a linear program. By substitution

$$\lambda_i = \prod_{k=1}^{i} \tau_k \quad \text{for } i = 1, \ldots, m \tag{18}$$

one gets the model:

$$\max\ \{\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i\bar{\delta}_i(\mathbf{x}) \ :\quad \mathbf{x} \in Q\} \tag{19}$$

with parameters $\lambda_1 > 0, \ldots, \lambda_m > 0$. Hereafter, we will refer to the model (19) as the recursive m–level MAD model (or m–MAD, for short).

Recalling example (16) which was used to illustrate the drawbacks of the κ–MAD model, and applying the m–MAD model to these random variables, one gets: $\bar{\delta}_i(\mathbf{x}') = \varepsilon^i/(1+\varepsilon)^{i+1}$ and $\bar{\delta}_i(\mathbf{x}'') = 0$. Observe that for any $m \geq 1$ and $0 < \lambda_i \leq 1$:

$$\mu(\mathbf{x}') - \sum_{i=1}^{m} \lambda_i\bar{\delta}_i(\mathbf{x}') > \mu(\mathbf{x}'') - \sum_{i=1}^{m} \lambda_i\bar{\delta}_i(\mathbf{x}'')$$

which is consistent with the fact that $R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$. In fact, an important feature of the m–MAD model is its consistency with the SSD relation.

According to the Corollary, the MAD model is consistent with the SSD relation provided that the trade-off coefficient is positive and not greater than 1. Imposing this restriction on coefficients τ_i, and considering (18), one gets:

$$1 \geq \lambda_1 \geq \ldots \geq \lambda_m > 0. \tag{20}$$

Moreover, taking parameters λ_i satisfying (20), and due to (18), one gets that $0 < \tau_i \leq 1$ for $i = 1, \ldots, m$. Thus, one may expect the m–MAD model to be consistent with the SSD relation provided that the parameters λ_i satisfy (20). This will be demonstrated in Section 4.

The parameters λ_i in the m–MAD model represent the corresponding trade-offs for different perceptions of downside risk. Using (18) they can be easily derived from coefficients τ_i. If the specific trade-off coefficient λ is selected in the MAD model, then it is quite natural to use the same value at every level of the m–MAD model, thus assuming $\tau_i = \lambda$ for $i = 1, \ldots, m$. This gives $\lambda_1 = \lambda$, $\lambda_2 = \lambda^2, \ldots, \lambda_m = \lambda^m$.

One may consider the objective function of the form:

$$\mu(\mathbf{x}) - \lambda_1 \sum_{i=1}^{m} \frac{\lambda_i}{\lambda_1} \bar{\delta}_i(\mathbf{x})$$

which explicitly shows that λ_1 is the basic risk to mean trade-off (denoted by λ in the original MAD model), whereas the quotients λ_i/λ_1 define additional penalties for larger deviations. Specifically, in terms of the penalty function u, the objective function in the m–MAD model takes the form

$$\mu(\mathbf{x}) - \lambda_1 \; E\{u(\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\})\}$$

where u is the (distribution dependent) piecewise linear convex function defined (for nonnegative arguments) by breakpoints: $b_0 = 0$, $b_i = b_{i-1} + \bar{\delta}_i(\mathbf{x})$ for $i = 1, \ldots, m-1$ and the corresponding slopes $s_1 = 1$, $s_i = \sum_{k=1}^{i} \lambda_i/\lambda_1$ for $i = 1, \ldots, m$. The quotients λ_i/λ_1 represent the increment of the slope of u at the breakpoints b_{i-1}. In particular, while assuming $\lambda_m = \ldots = \lambda_2 = \lambda_1$ one gets the convex function u with slopes $s_i = i$. The original MAD model with the linear function u, may be considered as a limiting case of m–MAD with $\lambda_m = \ldots = \lambda_2 = 0$.

Let us return to the case where the rates of return are represented as random variables measured by a finite set of discrete realizations r_{jt} (for $j = 1, \ldots, n$ and $t = 1, \ldots, T$). Then, assuming that the parameters λ_i satisfy (20), the m–MAD model can be formulated as an LP problem. For instance, the 2–MAD model (i.e. m–MAD model with $m = 2$) is given as:

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda_1}{T} \sum_{t=1}^{T} d_{t1} - \frac{\lambda_2}{T} \sum_{t=1}^{T} d_{t2} \tag{21}$$

subject to

$$\mathbf{x} \in Q \tag{22}$$

$$d_{t1} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \ldots, T \tag{23}$$

$$d_{t2} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j - \frac{1}{T} \sum_{l=1}^{T} d_{l1} \quad \text{for } t = 1, \ldots, T \tag{24}$$

$$d_{t1} \geq 0, \; d_{t2} \geq 0 \quad \text{for } t = 1, \ldots, T \tag{25}$$

The above formulation differs from (10)–(13) by having an additional group of T deviational variables d_{t2} (while the original d_t are renamed d_{t1}) and a corresponding additional group of T inequalities (24) linking these variables together (similar to equations (12) in the MAD model).

A general m–MAD model can be formulated as LP with mT deviational variables and mT inequalities linking them. In order to maintain the sparsity of its LP formulation (which is convenient when searching for solutions of the large scale LPs), it is preferable to write the m–MAD as:

$$\max\ z_0 + \sum_{i=1}^{m} \lambda_i z_i \tag{26}$$

subject to

$$\mathbf{x} \in Q \tag{27}$$

$$z_0 - \sum_{j=1}^{n} \mu_j x_j = 0 \tag{28}$$

$$T z_i + \sum_{t=1}^{T} d_{ti} = 0 \quad \text{for } i = 1, \ldots, m \tag{29}$$

$$d_{ti} - \sum_{k=0}^{i-1} z_k + \sum_{j=1}^{n} r_{jt} x_j \geq 0 \quad \text{for } t = 1, \ldots, T;\ i = 1, \ldots, m \tag{30}$$

$$d_{ti} \geq 0 \quad \text{for } t = 1, \ldots, T;\ i = 1, \ldots, m \tag{31}$$

In the above formulation $\mu(\mathbf{x})$ and $\bar{\delta}_i(\mathbf{x})$ $(i = 1, \ldots, m)$ are explicitly represented using additional variables z_0 and $-z_i$ $(i = 1, \ldots, m)$, respectively. Therefore, additional $m+1$ constraints (28)–(29) need to be introduced to define these variables. A number of nonzero coefficients in (30) can be further reduced if repetitions of coefficients r_{jt} in several groups of inequalities (30) for various i are avoided. This could be accomplished by introducing additional variables $y_t = \sum_{j=1}^{n} r_{jt} x_j$; however, it would increase the size of the LP problem to be solved. Diversification of the optimal portfolio of the m–MAD model is controlled by the number mT.

An important shortcoming of the MAD model is that although it can be expressed in terms of the downside risk measure $\bar{\delta}(\mathbf{x}) = E\{\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}$, the measure itself is symmetric in the sense of the downside and upside risks (see (7)). The m–MAD model overcomes this flaw as it introduces a risk term that is a true asymmetric risk measure. The reason for this is that for asymmetric distributions $E\{\max\{\mu(\mathbf{x}) + \eta - R_{\mathbf{x}}, 0\} \neq E\{\max\{R_{\mathbf{x}} - \mu(\mathbf{x}) - \eta, 0\}$ for $\eta > 0$. To illustrate this aspect of the m–MAD model, let us consider two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as [12]:

$$P\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 0.2, & \xi = 0 \\ 0.1, & \xi = 1 \\ 0.4, & \xi = 2 \\ 0.3, & \xi = 7 \\ 0, & \text{otherwise} \end{cases} \quad \text{and} \quad P\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 0.3, & \xi = -1 \\ 0.4, & \xi = 4 \\ 0.1, & \xi = 5 \\ 0.2, & \xi = 6 \\ 0, & \text{otherwise} \end{cases}$$

where $R_{\mathbf{x}''} = \mu(\mathbf{x}') - (R_{\mathbf{x}'} - \mu(\mathbf{x}'))$. Note that for any η, the upside deviation for $R_{\mathbf{x}'}$ is equal to the corresponding downside deviation for $R_{\mathbf{x}''}$ (i.e. $E\{\max\{R_{\mathbf{x}'} - \eta, 0\} = E\{\max\{\eta - R_{\mathbf{x}''}, 0\}$). Since $\mu(\mathbf{x}') = \mu(\mathbf{x}'') = 3$, $\bar{\delta}(\mathbf{x}') = \bar{\delta}(\mathbf{x}'') = 1.2$ and $\sigma^2(\mathbf{x}') = \sigma^2(\mathbf{x}'') = 7.4$, both random variables are indistinguishable according to the Markowitz and MAD models. However, it appears that $R_{\mathbf{x}''}$ has a longer and "heavier" tail to the left of the mean, which can be demonstrated by comparing their $F^{(2)}$ functions for $\eta < 3$. It is possible to show that for any $m > 1$ and λ_i satisfying (20), $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ according to the m–MAD model, thus allowing us to differentiate among portfolios with different downside risk characteristics.

4. THE m–MAD MODEL AND STOCHASTIC DOMINANCE

The function $F_{\mathbf{x}}^{(2)}$, used to define the SSD relation (see Section 2) can also be presented as [20]:

$$F_{\mathbf{x}}^{(2)}(\eta) = \int_{-\infty}^{\eta} (\eta - \xi)\ P_{\mathbf{x}}(d\xi) = P\{R_{\mathbf{x}} \leq \eta\} E\{\eta - R_{\mathbf{x}} | R_{\mathbf{x}} \leq \eta\} = E\{\max\{\eta - R_{\mathbf{x}}, 0\}\}$$

thus expressing the expected shortage for each target return η. Hence, in addition to being the most general dominance relation for all risk-averse preferences, SSD is also an intuitive multidimensional (continuum-dimensional) risk measure. As shown by Ogryczak and Ruszczyński [20], the graph of $F_{x}^{(2)}$, referred to as the Outcome–Risk (O–R) diagram, appears to be particularly useful for comparing uncertain returns. The function $F_{\mathbf{x}}^{(2)}$ is continuous, convex, nonnegative and nondecreasing. The graph $F_{\mathbf{x}}^{(2)}(\eta)$ (Figure 1) has two asymptotes which intersect at the point $(\mu(\mathbf{x}), 0)$. Specifically, the η-axis is the left asymptote and the line $\eta - \mu(\mathbf{x})$ is the right asymptote. In the case of a deterministic (risk-free) return ($R_{\mathbf{x}} = \mu(\mathbf{x})$), the graph of $F_{\mathbf{x}}^{(2)}(\eta)$ coincides with the asymptotes, whereas any uncertain return with the same expected value $\mu(\mathbf{x})$ yields a graph above (precisely, not below) the asymptotes. The space between the curve $(\eta, F_{\mathbf{x}}^{(2)}(\eta))$, and its asymptotes represents the dispersion (and thereby the riskiness) of $R_{\mathbf{x}}$ in comparison to the deterministic return $\mu(\mathbf{x})$. It is therefore called the *dispersion space.*

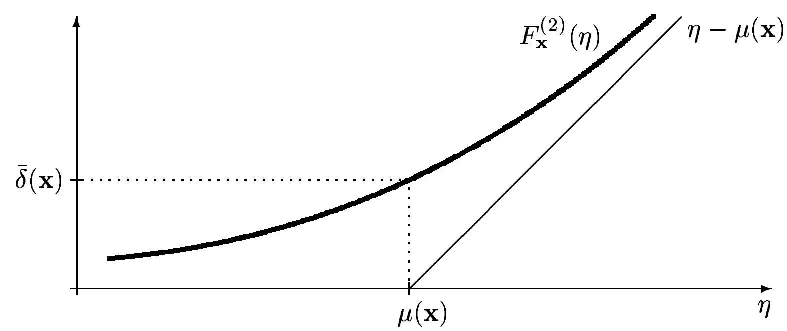

Figure 1. The O–R diagram and the absolute semideviation

The absolute semideviation $\bar{\delta}(\mathbf{x}) = F_{\mathbf{x}}^{(2)}(\mu(\mathbf{x}))$ turns out to be the maximal vertical diameter of the dispersion space [20]. Following the argument that only the dispersion related to underachievements should be considered as a measure of riskiness [17], one should rather focus on the downside dispersion space, that is, to the left of $\mu(\mathbf{x})$. Note that $\bar{\delta}(\mathbf{x})$ is the largest vertical diameter for both the entire dispersion space and the downside dispersion space.

Due to (8) and Proposition 1, it is possible to state that

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad E\{\min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\}\} \geq E\{\min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\}\} \tag{32}$$

Note that $P\{\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\} \leq \eta\}$ is equal to $P\{R_{\mathbf{x}} \leq \eta\}$ for $\eta < \mu(\mathbf{x})$ and equal to 1 for $\eta \geq \mu(\mathbf{x})$. The second performance function $F^{(2)}$ for the random variable $\min\{R_{\mathbf{x}}, \mu(\mathbf{x})\}$ coincides with

$F_{\mathbf{x}}^{(2)}(\eta)$ for $\eta \leq \mu(\mathbf{x})$ and takes the form of a straight line $\eta - (\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}))$ for $\eta > \mu(\mathbf{x})$. One may notice that

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad \min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\} \succeq_{SSD} \min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\} \tag{33}$$

which is a stronger relation than (32). From (33) it is possible to derive a stronger form of Proposition 1, namely:

PROPOSITION 2: If $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$, then $\min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\} \succeq_{SSD} \min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\}$ and $E\{\min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\}\} > E\{\min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\}\}$ whenever $\mu(\mathbf{x}') > \mu(\mathbf{x}'')$.

Let us define a sequence of random variables related to portfolio $\mathbf{x}$:

$$R_{\mathbf{x}}^{(0)} = R_{\mathbf{x}} \quad \text{and} \quad R_{\mathbf{x}}^{(i)} = \min\{R_{\mathbf{x}}^{(i-1)}, E\{R_{\mathbf{x}}^{(i-1)}\}\} \quad \text{for } i = 1, \ldots, m. \tag{34}$$

and the corresponding means:

$$\mu_i(\mathbf{x}) = E\{R_{\mathbf{x}}^{(i)}\} \quad \text{for } i = 0, 1, \ldots, m \tag{35}$$

where $\mu_0(\mathbf{x}) = \mu(\mathbf{x})$. Note that:

$$\mu_i(\mathbf{x}) = E\{\min\{R_{\mathbf{x}}^{(i-1)}, \mu_{i-1}(\mathbf{x})\}\} \leq \mu_{i-1}(\mathbf{x}) \quad \text{for } i = 1, \ldots, m.$$

Hence, $R_{\mathbf{x}}^{(i)} = \min\{R_{\mathbf{x}}, \mu_{i-1}(\mathbf{x})\}\}$ for $i = 1, \ldots, m$ and:

$$\mu_0(\mathbf{x}) = \mu(\mathbf{x}) \quad \text{and} \quad \mu_i(\mathbf{x}) = E\{\min\{R_{\mathbf{x}}, \mu_{i-1}(\mathbf{x})\}\} \quad \text{for } i = 1, \ldots, m.$$

Finally, because of (8), one gets $\mu_i(\mathbf{x}) = \mu_{i-1}(\mathbf{x}) - \bar{\delta}_i(\mathbf{x})$ for $i = 1, \ldots, m$. Thus:

$$\mu_i(\mathbf{x}) = \mu(\mathbf{x}) - \sum_{k=1}^{i} \bar{\delta}_k(\mathbf{x}) \quad \text{for } i = 1, \ldots, m \tag{36}$$

and

$$\bar{\delta}_i(\mathbf{x}) = E\{\max\{\mu_{i-1}(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} = F_{\mathbf{x}}^{(2)}(\mu_{i-1}(\mathbf{x})) \quad \text{for } i = 1, \ldots, m$$

The relations between $\mu_i(\mathbf{x})$ and $\bar{\delta}_i(\mathbf{x})$ may be illustrated on the O–R diagram as shown in Figure 2.

Note that, $\mu_i(\mathbf{x}) \leq \mu_{i-1}(\mathbf{x})$ for any $i \geq 1$. However, if $R_{\mathbf{x}}$ is lower bounded by a real number $l_{\mathbf{x}}$ (i.e. $P\{R_{\mathbf{x}} < l_{\mathbf{x}}\} = 0$), then $l_{\mathbf{x}} \leq \mu_i(\mathbf{x})$ for any $i \geq 0$. One may prove that in the case of $P\{R_{\mathbf{x}} = l_{\mathbf{x}}\} > 0$ if m tends to infinity then $\mu_m(\mathbf{x})$ converges to $l_{\mathbf{x}}$.

The objective function in model (19) can be expressed as:

$$\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}) = \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}) \tag{37}$$

where

$$\alpha_0 = 1 - \lambda_1 \ , \quad \alpha_i = \lambda_i - \lambda_{i+1} \quad \text{for } i = 1, \ldots, m-1 \quad \text{and} \quad \alpha_m = \lambda_m \tag{38}$$

Note that $\sum_{i=0}^{m} \alpha_i = 1$. Moreover, for the parameters λ_i satisfying (20), all α_i are nonnegative and the objective function (37) of m–MAD becomes a convex combination of means $\mu_i(\mathbf{x})$. Maximization of the means $\mu_i(\mathbf{x})$ is consistent with SSD.

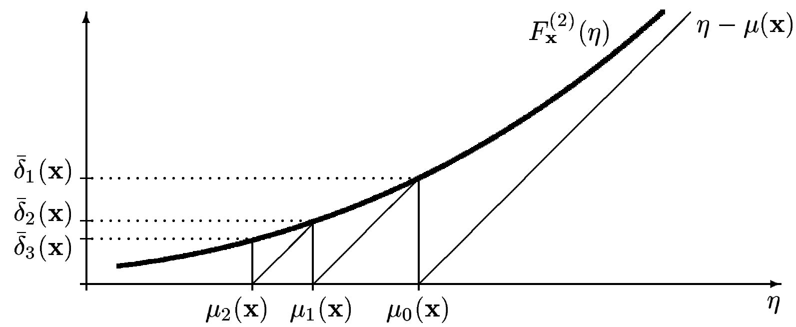

Figure 2. $\mu_i(\mathbf{x})$ and $\bar{\delta}_i(\mathbf{x})$ on the O–R diagram for $R_{\mathbf{x}}$

THEOREM 3: If $R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''}$, then $\mu_i(\mathbf{x}') \geq \mu_i(\mathbf{x}'')$ for all $i = 0, 1, \ldots, m$ and if any of these inequalities is strict ($\mu_{i_o}(\mathbf{x}') > \mu_{i_o}(\mathbf{x}'')$), then all subsequent inequalities are also strict ($\mu_i(\mathbf{x}') > \mu_i(\mathbf{x}'')$ for $i = i_o, \ldots, m$).

PROOF: According to (35), $\mu_i(\mathbf{x})$ ($i = 0, \ldots, m$) are the means of the corresponding random variables $R_{\mathbf{x}}^{(i)}$ defined in (34). By (recursive) application of Proposition 2 m times for defined random variables $R_{\mathbf{x}}^{(i)}$ (for $i = 0, \ldots, m-1$) one gets $R_{\mathbf{x}'}^{(i)} \succeq_{SSD} R_{\mathbf{x}''}^{(i)}$, and thereby $\mu_i(\mathbf{x}') \geq \mu_i(\mathbf{x}'')$ for all $i = 0, 1, \ldots, m$, as well as $\mu_i(\mathbf{x}') > \mu_i(\mathbf{x}'')$ whenever $\mu_{i-1}(\mathbf{x}') > \mu_{i-1}(\mathbf{x}'')$ for all $i = 1, \ldots, m$ ■

The assertion of Theorem 3, along with the relations (37)–(38), leads to the following theorem.

THEOREM 4: Except for portfolios characterized by identical mean and identical semideviations, every portfolio $\mathbf{x} \in Q$ that maximizes $\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x})$ with λ_i satisfying the condition (20) is efficient under SSD.

PROOF: According to (37)–(38) and (20), it follows that $\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}) = \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x})$ where all the coefficients α_i for $i = 0, \ldots, m$ are nonnegative, whereas α_m is strictly positive. Let $\mathbf{x}^0 \in Q$ maximize $\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x})$. This means that $\sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}^0) \geq \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x})$ for all $\mathbf{x} \in Q$. Suppose that there exists $\mathbf{x}' \in Q$ such that $R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}^0}$. Then from Theorem 3, $\mu_i(\mathbf{x}') \geq \mu_i(\mathbf{x}^0)$ for all $i = 0, \ldots, m$ and $\sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}') \geq \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}^0)$. The latter together with the fact that $\mathbf{x}^0$ is optimal, implies that $\sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}') = \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}^0)$ which means that $\mathbf{x}'$ must also be an optimal portfolio. Further, suppose that for some i_o ($0 \leq i_o \leq m$) there is $\mu_{i_o}(\mathbf{x}') > \mu_{i_o}(\mathbf{x}^0)$, then, according to Theorem 3, $\mu_m(\mathbf{x}') > \mu_m(\mathbf{x}^0)$. Since $\alpha_m > 0$, the latter leads to the conclusion that $\sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}') > \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}^0)$ which contradicts the assumption that $\mathbf{x}^0$ is optimal. Hence, $\mu_i(\mathbf{x}') = \mu_i(\mathbf{x}^0)$ for all $i = 0, \ldots, m$, and therefore $\mu(\mathbf{x}') = \mu(\mathbf{x}^0)$. Due to (36), it follows that $\bar{\delta}_i(\mathbf{x}') = \bar{\delta}_i(\mathbf{x}^0)$ for all $i = 1, \ldots, m$. ■

According to Theorem 4, a unique optimal portfolio of the m–MAD model with the trade-off coefficient λ_i satisfying (20), is efficient under SSD. In the case of multiple optimal solutions of (19) (similarly to the case of the original MAD model) some of them may be SSD dominated. Due to Theorem 4, an optimal portfolio $\mathbf{x}' \in Q$ can be SSD dominated only by another optimal portfolio $\mathbf{x}'' \in Q$ such that $\mu(\mathbf{x}'') = \mu(\mathbf{x}')$ and $\bar{\delta}_i(\mathbf{x}'') = \bar{\delta}_i(\mathbf{x}')$ for all $i = 1, \ldots, m$. This means that even if one generates an SSD dominated portfolio, then it has the same mean and downside risk as the dominating one.

5. DISCUSSION

The m–MAD model is well defined for any type of rate of return distribution, and it is not sensitive to the scale shifting with regard to the mean and deviations. It also is a "true" downside risk model as the risk measure used in m–MAD is not symmetric. Therefore, it comprehends the investor's (downside) risk aversion, and as demonstrated in the paper, it is robust given SSD efficiency. The computational robustness associated with the linear programming (LP) formulation for random returns defined by a finite number of scenarios (historical data) is the most important advantage of the MAD model. Although formulated in a recursive manner, the proposed m–MAD model, maintains the LP formulation. The number of LP constraints is increased by factor m (usually not greater than 3 or 4) and the resulting augmented LP can still be solved by a standard commercial LP solver.

Both the Markowitz and MAD models are powerful portfolio optimization tools, which do not impose a significant information burden on an investor for a given risk/return trade-off. This feature, considered to be an advantage in certain situations, may also be viewed as a shortcoming because it does not provide an investor with any process control mechanism. This is not the case with the m–MAD model. Application of this model allows an investor to control and fine-tune the portfolio optimization process through the ability to determine m trade-off parameters λ_i. Thus, an investor exhibiting (downside) risk aversion can, to some extent, control which securities enter his/her optimal portfolio through varying a penalty associated with the "larger" (downside) deviations from a mean return. Within such a framework, higher risk aversion is reflected in an investor's desire to exclude from a portfolio those securities that have potential "large" deviations, while a more risk neutral investment attitude will result in the acceptance of these securities. On the other hand, the modeling flexibility of m–MAD is also its possible shortcoming related to the selection of proper values for m and λ_i parameters. It is important to stress here that if a specific trade-off coefficient λ is selected in the original MAD model, then it is quite reasonable to use the same coefficient at every level of the m–MAD model, which results in: $\lambda_1 = \lambda$, $\lambda_2 = \lambda^2, \ldots,$ $\lambda_m = \lambda^m$. For computational reasons it becomes clear that a limited number of levels (small value of m) should be considered. However, for the trade-off $\lambda < 1$ it is very likely that, in the sequence $\{\lambda^i\}_{1,\ldots,m}$, the values of its elements approach 0 very quickly. Thus, in practice, it is not necessary to solve m–MAD for a value of m larger than 3 or 4.

The m–MAD model allows us to penalize larger downside deviations relative to the mean by using a piecewise linear penalty function. The penalty increase per unit deviation depends on two factors: the slope of a corresponding linear segment of the penalty function and a number of segments. The former is controlled through trade-off coefficients λ_i. The latter depends on the recursive scheme of penalizing deviation greater than the corresponding mean deviation, and thus it involves a self-scaling. One might consider a scaling parameter v and the scaled deviations such that $\bar{\delta}_{2(v)}(\mathbf{x}) = E\{\max\{\mu(\mathbf{x}) - v\bar{\delta}_1(\mathbf{x}) - R_\mathbf{x}, 0\}\}$. However, in order to maintain the SSD consistency of the m–MAD model, values of v must belong to $(0, 1]$ interval and more complicated restrictions on the trade-off coefficients must be satisfied. Such requirements together with an unnecessary modeling burden, in our opinion, overweight the possible advantages associated with relaxing the self-scaling component with a parameter v.

In this paper, we argue that a solution of the m–MAD model for a particular m, reflects an investor's specific (downside) risk aversion attitude. At the same time, by varying m and solving a sequence of the m–MAD models, it is possible to generate a set of optimal portfolios $\{\mathbf{x}^0(m)\}_{m=1,2,\ldots}$. For each of these portfolios, one can define some piecewise linear penalty function where $\mathbf{x}^0(m)$ is its optimal solution. Moreover, assuming that this process is applied to historical data, for every $\{\mathbf{x}^0(m)\}_{m=1,2,\ldots}$ it is possible to calculate the portfolio cumulative wealth index (pcwi). Therefore, a trajectory of $\mathbf{x}^0(m)$ plotted on the "pcwi scale" allows one to represent an investor's proneness to (downside) risk aversion as a function of the pcwi. Such combined representation (a penalty function and a function of pcwi) may prove to be sufficiently intuitive to serve as a useful tool in evaluating an investor's risk aversion attitude — information that is critical when designing an effective investment strategy. However, the specific computational and methodological issues associated with this representation and evaluation need to be investigated further and resolved prior to its practical application.

ACKNOWLEDGMENTS

The authors wish to thank Ms. Anne Burgess for editorial assistance and the associate editor and anonymous referees for their helpful comments and suggestions. Research conducted by W. Michalowski was supported by the grant from Natural Sciences and Engineering Research Council of Canada and was conducted while the author was with the DAS Project at International Institute for Applied Systems Analysis, Laxenburg, Austria.

REFERENCES

[1] Charnes, A., Cooper, W.W., Ferguson, R.O., "Optimal Estimation of Executive Compensation by Linear Programming," *Management Science*, vol. 2, pp. 138–151, 1955.

[2] Edgeworth, F.Y., "A New Method for Reducing Observations Relating to Several Quantities," *Philosophical Magazine (5th Ser.)*, vol. 24, pp. 222–223, 1887.

[3] Elton, E.J., Gruber, M.J., *Modern Portfolio Theory and Investment Analysis* (Third Edition), John Wiley & Sons, New York, 1987.

[4] Entriken, R., Yu, C., "Absolute Deviation as a Measure of Risk," Technical Report SOL 95-5, SOL, Stanford University, Stanford, CA, 1995.

[5] Feinstein, C.D., Thapa, M.N., "A Reformulation of a Mean–Absolute Deviation Portfolio Optimization Model," *Management Science*, vol. 39, pp. 1552–1553, 1993.

[6] Fishburn, P.C., *Decision and Value Theory*, John Wiley & Sons, New York, 1964.

[7] Fishburn, P.C., "Mean–Risk Analysis with Risk Associated with Below Target Returns", *American Economic Review*, vol. 67, pp. 116–126, 1977.

[8] Fishburn, P.C., "Stochastic Dominance and Moments of Distributions," *Mathematics of Operations Research*, vol. 5, pp. 94–100, 1980.

[9] Grootveld, H., Hallerbach, W., "Variance vs Downside Risk: Is There Really That Much Difference?", *European Journal of Operational Research*, vol. 114, pp. 304–319, 1999.

[10] Hanoch, G., Levy, H., "The Efficiency Analysis of Choices Involving Risk," *Rev. Economic Studies*, vol. 36, pp. 335–346, 1969.

[11] Kenyon, C.M., Savage, S., Ball, B., "Equivalence of Linear Deviation about Mean and Mean Absolute Deviation about the Mean Objective Functions," *Operations Research Letters*, vol. 24, pp. 181–185, 1999.

[12] Konno, H., "Piecewise Linear Risk Function and Portfolio Optimization," *Journal of the Operations Research Society of Japan*, vol. 33, pp. 139–156, 1990.

[13] Konno, H., Yamazaki, H., "Mean–Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market," *Management Science*, vol. 37, pp. 519–531, 1991.

[14] Kroll, Y., Levy, H., Markowitz, H.M., "Mean–Variance Versus Direct Utility Maximization," *Journal of Finance*, vol. 39, pp. 47–62, 1984.

[15] Levy, H., "Stochastic Dominance and Expected Utility: Survey and Analysis," *Management Science*, vol. 38, pp. 555–593, 1992.

[16] Markowitz, H.M., "Portfolio Selection," *Journal of Finance*, vol. 7, pp. 77–91, 1952.

[17] Markowitz, H.M., *Portfolio Selection: Efficient Diversification of Investments*, John Wiley & Sons, New York, 1959.

[18] Markowitz, H.M., *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford, 1987.

[19] Ogryczak, W., Ruszczyński, A., "On Stochastic Dominance and Mean–Semideviation Models," *Mathematical Programming*, vol. 89, pp. 217–232, 2001.

[20] Ogryczak, W., Ruszczyński, A., "From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures," *European Journal of Operational Research*, vol. 116, pp. 33–50, 1999.

[21] Pratt, J.W., Raiffa, H., Schlaifer, R., *Introduction to Statistical Decision Theory*, MIT Press, Cambridge, MA, 1995.

[22] Rothschild, M., Stiglitz, J.E., "Increasing Risk: I. A Definition," *Journal of Economic Theory*, vol. 2, pp. 225–243, 1969.

[23] Sharpe, W.F., "A Linear Programming Approximation for the General Portfolio Analysis Problem," *Journal of Financial and Quantitative Analysis*, vol. 6, pp. 1263–1275, 1971.

[24] Sharpe, W.F., "Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios," *Management Science*, vol. 18, pp. B1–B13, 1971.

[25] Simaan, Y., "Estimation Risk in Portfolio Selection: The Mean Variance Model versus the Mean Absolute Deviation Model," *Management Science*, vol. 43, pp. 1437–1446, 1997.

[26] Sortino, F.A., Forsey, H.J., "On the Use and Misuse of Downside Risk", *Journal of Portfolio Management*, Winter, pp. 35–42, 1996.

[27] Speranza, M.G., "Linear Programming Models for Portfolio Optimization", *Finance*, vol. 14, pp. 107–123, 1993.

[28] Whitmore, G.A, Findlay, M.C. (Eds.), *Stochastic Dominance: An Approach to Decision–Making Under Risk*, D.C.Heath, Lexington, MA, 1978.

[29] Zenios, S.A., Kang, P., "Mean–Absolute Deviation Portfolio Optimization for Mortgage-Backed Securities," *Annals of Operations Research*, vol. 45, pp. 433–450, 1993.

Received August, 1998
Revised July, 2000
Accepted 28 November, 2000