IIASA

International Institute for Applied Systems Analysis • A-2361 Laxenburg • Austria
Tel: +43 2236 807 • Fax: +43 2236 71313 • E-mail: info@iiasa.ac.at • Web: www.iiasa.ac.at

INTERIM REPORT IR-98-041 / June 1998

Extending the MAD Portfolio Optimization Model to Incorporate Downside Risk Aversion

Wojtek Michalowski(michalow@iiasa.ac.at)
Wlodzimierz Ogryczak(Wlodzimierz.Ogryczak@mimuw.edu.pl)

Approved by
Pekka Korhonen (korhonen@iiasa.ac.at)
Leader, *Decision Analysis and Support Project (DAS)*

Abstract

The 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 proposed where risk is measured by (mean) absolute deviation instead of a variance. The MAD model is computationally attractive, since it is transformed into an easy to solve linear programming program. In this paper we present an extension to the MAD model allowing to account for downside risk aversion of an investor, and at the same time preserving simplicity and linearity of the original MAD model.

Keywords: Portfolio Optimization, Downside Risk Aversion, Linear Programming

About the Authors

Wojtek Michalowski is a Senior Research Scholar with the Decision Analysis and Support Project at IIASA.

Wlodzimierz Ogryczak is Associate Professor of Operations Research at the Department of Mathematics and Computer Science, Warsaw University.

Contents

Extending the MAD Portfolio Optimization Model to Incorporate Downside Risk Aversion

Wojtek Michalowski(michalow@iiasa.ac.at)
Wlodzimierz Ogryczak(Wlodzimierz.Ogryczak@mimuw.edu.pl)

1 Introduction

Since the advent of the Modern Portfolio Theory (MPT) arising from the work of Markowitz (1952), the notion of investing in diversified portfolios has become one of the most fundamental concepts of portfolio management. While developed as a financial economic theory in conditional-normative framework, the MPT has spawned a variety of applications and provided background for further theoretical models. The original Markowitz model was derived using a representative investor belonging to the normative utility framework, which manifested in portfolio optimization techniques based on the mean-variance rule. This framework proved to be sufficiently rich to provide the main theoretical background for the analysis of importance of diversification. It also gave rise to asset pricing models for security pricing, the most known among them being the Capital Asset Pricing Model (CAPM) (Elton and Gruber, 1987). A reliance on the MPT let to the notion that the best managed portfolio is the one which is most widely diversified and such a portfolio may be created through passive buy-and-hold investment strategy.

The portfolio optimization problem considered in this paper follows the original Markowitz formulation and is based on a single period model of investment. At the beginning of a period, an investor allocates 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 generates a certain (random) rate of return. The change of capital invested observed at the end of the period is measured by the weighted average of the individual rates of return. In mathematical terms, for selecting weights reflecting an amount invested in each security, an investor needs to solve a 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 into individual securities).

Following the seminal work by Markowitz (1952), such a portfolio optimization problem is usually modeled as a bicriteria optimization problem where a reasonable trade–off between expected rate of return and risk is sought. In the Markowitz model the risk is measured by a variance from mean rate of return, thus resulting in a formulation of a quadratic programming model. Following Sharpe (1971), many attempts have been made to linearize the portfolio optimization problem (c.f., Speranza, 1993 and references therein). Lately, Konno and Yamazaki (1991) proposed the MAD portfolio optimization model where risk is measured by (mean) absolute deviation instead of variance. The model is computationally attractive as (for discrete random variables) it results in solving linear programming (LP) problems.

There is an argument that the variability of rate of return above the mean should not be penalized since an investor worries rather about underperformance of a portfolio than its overperformance. This led Markowitz (1959) to propose downside risk measures such as (downside) semivariance to replace variance as the risk measure. The absolute deviation used in the MAD model to measure risk is taken as twice the downside semideviation. Therefore, the MAD model is, in fact, based on the downside risk measured with mean deviation to the mean. However, an investor who uses this model is assumed to have constant dis-utility (a term "dis-utility" is used here to emphasize a fact that an investor is a "utility minimizer") for a unit deviation from the mean portfolio 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 risk attitude and to present an extension to the MAD model which incorporates downside risk aversion.

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., Whitmore and Findlay, 1978; Levy, 1992). However, the MAD model is consistent with the second degree stochastic dominance, provided that the trade-off coefficient between risk and return is bounded by a certain constant (Ogryczak and Ruszczyński, 1997). The proposed extension of the MAD model retains consistency with the 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, enabling to incorporate (downside) risk aversion of an investor. Consistency of a resulting model with the stochastic dominance is discussed in Section 4. The paper concludes with a discussion.

2 The MAD model

Let $J = \{1, 2, \ldots, n\}$ denotes 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 securities' weights (decision variables) defining a portfolio. To represent a portfolio, the weights must satisfy a set of constraints which 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 usually needs to consider some other requirements 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 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 $R_{\mathbf{x}} = \sum_{j=1}^{n} R_j x_j$ which represents portfolio's 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 (1952), 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. An important advantage of mean–risk approaches is a 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 with a so-called *critical line approach* (Markowitz, 1987), by solving parametric problem (3) with changing $\lambda > 0$. Such an approach allows to select appropriate value of the trade-off coefficient λ and the corresponding optimal portfolio through a graphical analysis in the mean-risk image space.

It is clear that if the risk is measured by 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 problem (3) results in having a quadratic objective function.

Despite the fact that problem (3) is seldom used as a tool for optimizing large portfolios, this model is widely recognized as a starting point for the MPT (c.f., Elton and Gruber, 1987). In an attempt to analyze reasons behind limited popularity of the Markowitz's model among investors, Konno and Yamazaki (1991) summarized its shortcomings as:

a) a necessity to solve a large scale quadratic programming problem;

b) investor's reluctance to rely on variance as a measure of risk (Kroll at al., 1984).
   The Markowitz model is known to be valid (and consistent with the stochastic dominance) in the case of normal distribution of returns but, becomes doubtful in case of other return distributions, especially nonsymmetric ones;

c) possible existence of too many weights with nonzero values in the optimal solution of (3), thus making the resulting portfolio over-diversified and hardly implementable.

Konno and Yamazaki (1991) proposed an alternative model where they use (mean) absolute deviation from a mean as a risk measure. It is defined as:

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

where $P_{\mathbf{x}}$ denotes a probability measure induced by the random variable $R_{\mathbf{x}}$ (Pratt et al., 1995). When $\delta(\mathbf{x})$ is used as a risk measure $\varrho(\mathbf{x})$ in model (3), it gives a so-called MAD portfolio optimization model.

The absolute deviation (4) was already considered by Edgeworth (1887) in the context of regression analysis. Within this context it was used in various areas of decision making resulting among others in the goal programming formulation of LP problems (Charnes et al., 1955). The absolute deviation as a measure was also considered in the portfolio analysis (Sharpe, 1971a, and references therein) and has been recommended by the Bank Administration Institute (1968) as a measure of dispersion. The MAD model is based on the absolute semideviation as a risk measure and Konno and Yamazaki validated this model using the Tokyo stock exchange data (Konno and Yamazaki, 1991).

Many authors pointed out that the MAD model opens up opportunities for more specific modeling of the downside risk (Konno, 1990; Feinstein and Thapa, 1993), because absolute deviation may be considered as a measure of the downside risk, (observe that $\delta(\mathbf{x})$ equals twice the (downside) absolute semideviation):

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

Hence, the following parametric optimization problem will be called the MAD model:

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

Simplicity and computational robustness are perceived as the most important advantages of the MAD model. According to Konno and Yamazaki (1991), r_{jt} is the realization of random variable R_j during period t (where $t = 1, \ldots, T$) which is available from the historical data or from some future projection. It is also assumed that the expected value of the random variable can be approximated by the average derived from these data. Then:

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

Therefore, MAD model (6) can be rewritten (Feinstein and Thapa, 1993) as the following LP:

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda}{T}\sum_{t=1}^{T} d_t \tag{7}$$

subject to

$$\mathbf{x} \in Q \tag{8}$$

$$d_t \geq \sum_{j=1}^{n} (\mu_j - r_{jt})x_j \quad \text{for } t = 1, \ldots, T \tag{9}$$

$$d_t \geq 0 \quad \text{for } t = 1, \ldots, T \tag{10}$$

The LP formulation (7)–(10) can be effectively solved even for large number of securities. Moreover, a number of securities included in the optimal portfolio (i.e. a number of weights with nonzero values) is controlled by number T. In the case when Q as given by (1), no more than $T + 1$ securities will be included in the optimal portfolio.

The MAD model is clearly a downside risk model. Note that for any real number η it holds that:

$$\eta - E\{\max\{\eta - R_{\mathbf{x}}, 0\}\} = E\{\min\{R_{\mathbf{x}}, \eta\}\} \tag{11}$$

Hence

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

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. However, the downside risk is measured just by the mean of downside deviations (see (5)), and thus the MAD model assumes a constant dis-utility of an investor for a unit of the downside deviation from the mean portfolio rate of return.

An extension to the MAD model should allow to penalize larger downside deviations, thus providing for better modeling of the risk avert preferences. Observe that such an extension is in some manner equivalent to introduction of a convex dis-utility function u, resulting in replacing (5) with:

$$\bar{\delta}_u(\mathbf{x}) = E\{u(\max\{\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\})\} \tag{12}$$

Certainly, to preserve a linearity of the model, function u must be piecewise linear.

If the rates of return are multivariate normally distributed, then the MAD model is equivalent to the Markowitz model (Konno and Yamazaki, 1991). However, the MAD model does not require any specific type of return distributions, what facilitated its application to portfolio optimization for mortgage-backed securities (Zenios and Kang, 1993) and other classes of investments where distribution of rate of return is known to be not symmetric.

Recently, the MAD model was further validated by Ogryczak and Ruszczyński (1997) who demonstrated that if the trade-off coefficient λ is bounded by 1, then the model is partially consistent with the second degree stochastic dominance (Whitmore and Findlay, 1978). Origins of a stochastic dominance are in an axiomatic model of risk-averse preferences (Fishburn, 1964; Hanoch and Levy, 1969; Rothschild and Stiglitz, 1970). Since that time it has been widely used in economics and finance (see Levy, 1992 for numerous references). Detailed and comprehensive discussion of a stochastic dominance and its relation to the downside risk measures is given in Ogryczak and Ruszczyński (1997, 1998).

In the stochastic dominance approach uncertain prospects (random variables) are compared by 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 $F^{(1)}_{\mathbf{x}}$ is defined as the right-continuous cumulative distribution function itself:

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

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

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

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

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

The corresponding strict dominance 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 rules* ($R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$), if $F^{(2)}_{\mathbf{x}'}(\eta) \leq F^{(2)}_{\mathbf{x}''}(\eta)$ for all η, with at least one inequality strict. A feasible portfolio $\mathbf{x}^0 \in Q$ is called *efficient under the SSD rules* 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. 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 the SSD rules.

The necessary condition for the SSD relation is (c.f. Fishburn, 1980):

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

Ogryczak and Ruszczyński (1997) modified this relation to consider absolute semideviations, and proved the following proposition:

Proposition 1 *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 (11) lead to the following corollary (see Ogryczak and Ruszczyński, 1997, for details):

Corollary 1 *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 the SSD rules.*

It follows from Corollary 1 that the unique optimal solution of the MAD problem (model (6)) with the trade-off coefficient $0 < \lambda \leq 1$ is efficient under the SSD rules. In the case of multiple optimal solutions of model (6), one of them is efficient under SSD rules, but also some of them may be SSD dominated. Due to Corollary 1, 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 the SSD for bounded trade-offs, it requires additional specification if one wants to maintain the SSD efficiency for every optimal portfolio. An extension of the MAD model presented in this paper provides such a specification.

3 Extended MAD model

The MAD model (6) measures downside risk but it does not properly account for risk aversion attitude. In order to do so, one needs to differentiate between different levels of deviations, and to penalize “larger” ones. Such an extension of the MAD model for portfolio optimization was already proposed by Konno (1990) who considered additional mean deviations from some target rate of return predefined as proportional to the mean rate of return. Within the framework of downside risk (and downside deviations) this may be interpreted as an introduction to the following deviations:

$$\bar{\delta}_\kappa(\mathbf{x}) = E\{\max\{\kappa\, \mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \quad \text{for } 0 \leq \kappa \leq 1 \tag{13}$$

For $\kappa = 1$ one gets the $\bar{\delta}_1(\mathbf{x}) = \bar{\delta}(\mathbf{x})$, i.e. the absolute semideviation used in the original MAD model. One may try to augment the downside risk measure by penalizing additional deviations for several $\kappa < 1$. In terms of dis-utility function of deviations (see (12)), this approach is equivalent to introduction of 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 (1990) did:

$$\max \{\mu(\mathbf{x}) - \lambda\bar{\delta}(\mathbf{x}) - \lambda_\kappa \bar{\delta}_\kappa(\mathbf{x}) \ : \quad \mathbf{x} \in Q\} \tag{14}$$

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 the κ–MAD.

Note that in the κ–MAD model one penalizes deviations which are relatively large with respect to the expected rate of return (larger that $(1-\kappa)\mu(\mathbf{x})$). However, the model behaves correctly (defines target outcomes smaller than the mean) only in the case of nonnegative mean. This is true for a typical portfolio optimization problem, but in general, one needs to be very cautious while trying to apply the κ–MAD model to other types of outcomes. Especially, because the deviations $\bar{\delta}_\kappa(\mathbf{x})$ are sensitive to any shift of the scale of outcomes.

Konno (1990) did not analyze the consistency of the κ–MAD model with the stochastic dominance. Such a comprehensive analysis is beyond the scope of this paper. Nevertheless, one can see that for the SSD consistency, a proper selection of the parameters in κ–MAD may be quite a difficult task. We illustrate this with a small example. Consider two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as:

$$P\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 1/(1+\varepsilon), & \xi = 0 \\ \varepsilon/(1+\varepsilon), & \xi = 1 \\ 0, & \text{otherwise} \end{cases}, P\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 1, & \xi = 0 \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

where ε is arbitrarily a 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 $\mu(\mathbf{x}'') = \bar{\delta}(\mathbf{x}'') = 0$. Simple arithmetic shows 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 by Konno (1990). 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 only $\lambda_{0.5}$ increases the trade-off coefficient λ. It is easy to see that in the case of $\lambda_{0.5} \geq 1 - \lambda + \varepsilon$, $R_{\mathbf{x}''}$ is preferred $R_{\mathbf{x}'}$. This inconsistency of κ–MAD is overcome in the proposed extension of the MAD model.

Lets start with the original MAD model (6) assuming that the trade-off coefficient (λ) has value τ_1. Since the mean deviation is already considered in (6), it is quite natural to focus on this part of large deviations which 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 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 may wish to penalize the "second level" surplus deviations exceeding that mean.

This can be formalized as follows:

$$\max\ \{\mu(\mathbf{x}) - \sum_{i=1}^{m} (\prod_{k=1}^{i} \tau_k)\bar{\delta}_i(\mathbf{x}) \ :\quad \mathbf{x} \in Q\} \tag{16}$$

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

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

By substitution

$$\lambda_i = \prod_{k=1}^{i} \tau_k \quad \text{for } i = 1, \ldots, m \tag{17}$$

one gets the model:

$$\max\ \{\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}) \ :\quad \mathbf{x} \in Q\} \tag{18}$$

where $\lambda_1 > 0, \ldots, \lambda_m > 0$ are the model parameters. Hereafter, we will refer to the problem (18) as the recursive m–level MAD model (or m–MAD for short).

The parameters λ_i in the m–MAD model represent corresponding trade-offs for different perceptions of downside risk. Using (17), they can be easily derived from trade-off coefficients τ_i. If specific trade-off coefficient λ is selected in the MAD model, then it is quite natural to use the same value for the whole 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 a dis-utility function of downside deviations (see (12) in Section 2), 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 linear function u, may be considered as a limiting case of m–MAD with $\lambda_m = \ldots = \lambda_2 = 0$.

Lets consider the case when the mean rates of return of securities are derived from a finite set of (historical) data r_{jt} (for $j = 1, \ldots, n$ and $t = 1, \ldots, T$). Then, assuming that $1 \geq \lambda_1 \geq \ldots \geq \lambda_m > 0$, the m–MAD model can be formulated as an LP problem. For instance, 2–MAD model (i.e. m–MAD model with $m = 2$) is given as:

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda_1}{T} \sum_{t=1}^{T} d_{t1} - \frac{\lambda_2}{T} \sum_{t=1}^{T} d_{t2} \tag{19}$$

subject to

$$\mathbf{x} \in Q \tag{20}$$

$$d_{t1} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \ldots, T \tag{21}$$

$$d_{t2} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j - \frac{1}{T} \sum_{l=1}^{T} d_{l1} \quad \text{for } t = 1, \ldots, T \tag{22}$$

$$d_{t1} \geq 0, \; d_{t2} \geq 0 \quad \text{for } t = 1, \ldots, T \tag{23}$$

The above formulation differs from (7)–(10) by having an additional group of T deviational variables d_{t2} (while the original d_t are renamed to d_{t1}) and corresponding additional group of T inequalities (22) linking these variables together (similar to equations (9) in the MAD model).

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

$$\max \; z_0 + \sum_{i=1}^{m} \lambda_i z_i \tag{24}$$

subject to

$$\mathbf{x} \in Q \tag{25}$$

$$z_0 - \sum_{j=1}^{n} \mu_j x_j = 0 \tag{26}$$

$$T z_i + \sum_{t=1}^{T} d_{ti} = 0 \quad \text{for } i = 1, \ldots, m \tag{27}$$

$$d_{ti} - \sum_{k=0}^{i-1} z_s + \sum_{j=1}^{n} r_{jt} x_j \geq 0 \quad \text{for } t = 1, \ldots, T; \; i = 1, \ldots, m \tag{28}$$

$$d_{ti} \geq 0 \quad \text{for } t = 1, \ldots, T; \; i = 1, \ldots, m \tag{29}$$

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 (26)–(27) need to be introduced to define these variables.

A number of nonzero coefficients in (28) can be further reduced if repetitions of coefficients r_{jt} in several groups of inequalities (28) for various t are avoided. This can 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.

Recall the pair of random variables in (15) used to show drawbacks of the κ–MAD model. While applying the m–MAD model one gets: $\bar{\delta}_i(\mathbf{x}') = \varepsilon^i/(1+\varepsilon)^i$ and $\bar{\delta}_i(\mathbf{x}'') = 0$. It is easy to show that for any $m \geq 1$ and $0 < \lambda_i \leq 1$:

$$\mu(\mathbf{x}') - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}') > 0 = \mu(\mathbf{x}'') - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}'')$$

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. This will be demonstrated in the next section.

To illustrate how the m–MAD model introduces downside risk aversion into the original MAD, consider two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as (Konno, 1990):

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

Note that $\mu(\mathbf{x}') = \mu(\mathbf{x}'') = 3$, $\bar{\delta}(\mathbf{x}') = \bar{\delta}(\mathbf{x}'') = 1.2$ and $\sigma^2(\mathbf{x}') = \sigma^2(\mathbf{x}'') = 7.4$. Hence, two random variables are identical from the viewpoint of Markowitz's as well as the MAD models. It turns out, however, that $R_{\mathbf{x}''}$ has a longer tail to the left of the mean which can be demonstrated by comparing third moments of the random variables or their $F^{(2)}$ functions for $\eta < 3$. Simple arithmetic shows that for any $m > 1$ and λ_i satisfying $1 \geq \lambda_1 \geq \ldots \geq \lambda_m > 0$, $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ according to the m–MAD model.

4 The m–MAD model and stochastic dominance

Function $F^{(2)}_{\mathbf{x}}$, used to define the SSD relation (see Section 2) can also be presented as (Ogryczak and Ruszczyński, 1997):

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

thus expressing the expected shortage for each target outcome η. Hence, in addition to being the most general dominance relation for all risk-averse preferences, SSD is a rather intuitive multidimensional (continuum-dimensional) risk measure. As shown by Ogryczak and Ruszczyński (1997), the graph of $F^{(2)}_{x}$, referred to as the Outcome–Risk (O–R) diagram, appears to be particularly useful for comparing uncertain prospects. The function $F^{(2)}_{\mathbf{x}}$ is continuous, convex, nonnegative and nondecreasing. The graph $F^{(2)}_{\mathbf{x}}(\eta)$ (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) outcome ($R_{\mathbf{x}} = \mu(\mathbf{x})$), the graph of $F^{(2)}_{\mathbf{x}}(\eta)$ coincides with the asymptotes, whereas any uncertain outcome with the same expected value $\mu(\mathbf{x})$ yields a graph above (precisely, not below) the asymptotes. The space between the curve $(\eta, F^{(2)}_{\mathbf{x}}(\eta))$, and its asymptotes represents the dispersion (and thereby the riskiness) of $R_{\mathbf{x}}$ in comparison to the deterministic outcome of $\mu(\mathbf{x})$. It is called the *dispersion space*.

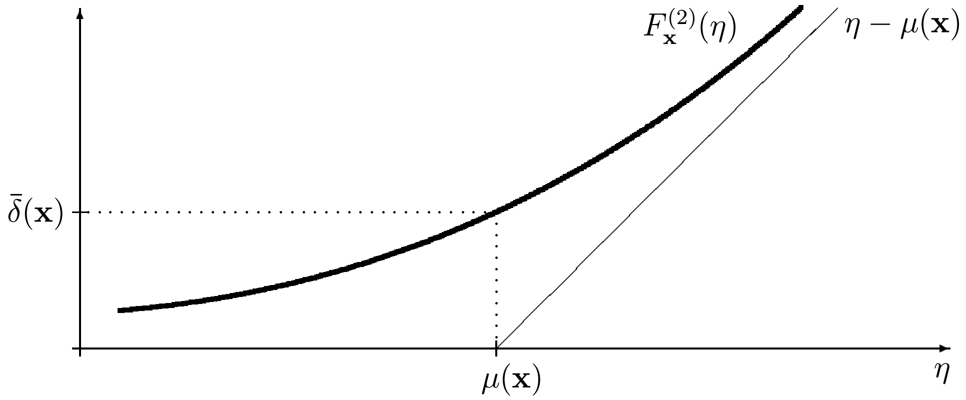

Figure 1: The O–R diagram and the absolute semideviation

Both size and shape of the dispersion space are important for complete description of the riskiness. Nevertheless, it is quite natural to consider some "size parameters" as summary characteristics of riskiness. The absolute semideviation $\bar{\delta}(\mathbf{x}) = F^{(2)}_{\mathbf{x}}(\mu(\mathbf{x}))$ turns out to be the maximal vertical diameter of the dispersion space (Ogryczak and Ruszczyński, 1997). According to arguments that only the dispersion related to underachievements should be considered as a measure of riskiness (Markowitz, 1959), one should 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. Thus $\bar{\delta}(\mathbf{x})$ appears to be a reasonable linear measure of the risk related to the representation of a random variable $R_{\mathbf{x}}$ by its expected value $\mu(\mathbf{x})$.

Due to (11) and Proposition 1, it is possible to state that

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad E\{\min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\}\} \geq E\{\min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\}\} \tag{30}$$

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 $F^{(2)}_{\mathbf{x}}(\eta)$ for $\eta \leq \mu(\mathbf{x})$ and takes the form of a straight line $\eta - (\mu(\mathbf{x}) - \bar{\delta}(\mathbf{x}))$ for $\eta > 1$. One may notice that

$$R_{\mathbf{x}'} \succeq_{SSD} R_{\mathbf{x}''} \quad \Rightarrow \quad \min\{R_{\mathbf{x}'}, \mu(\mathbf{x}')\} \succeq_{SSD} \min\{R_{\mathbf{x}''}, \mu(\mathbf{x}'')\} \tag{31}$$

which is stronger relation than (30). ¿From (31) 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 $\mathbf{x}$:

$$R_{\mathbf{x}}^{(0)} = R_{\mathbf{x}} \quad \text{and} \quad R_{\mathbf{x}}^{(i)} = \min\{R_{\mathbf{x}}^{(i-1)}, E\{R_{\mathbf{x}}^{(i-1)}\}\} \quad \text{for } i = 1, \ldots, m. \tag{32}$$

and the corresponding means:

$$\mu_i(\mathbf{x}) = E\{R_{\mathbf{x}}^{(i)}\} \quad \text{for } i = 0, 1, \ldots, m \tag{33}$$

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, due to (11), one gets $\mu_i(\mathbf{x}) = \mu_{i-1}(\mathbf{x}) - \bar{\delta}_i(\mathbf{x})$ for $i = 1, \ldots, m$. Thus:

$$\mu_i(\mathbf{x}) = \mu(\mathbf{x}) - \sum_{k=1}^{i} \bar{\delta}_k(\mathbf{x}) \quad \text{for } i = 1, \ldots, m \tag{34}$$

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 of the m–MAD model (18) can be expressed as:

$$\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x}) = \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}) \tag{35}$$

where

$$\alpha_0 = 1 - \lambda_1 \ , \quad \alpha_i = \lambda_i - \lambda_{i+1} \quad \text{for } i = 1, \ldots, m-1 \quad \text{and} \quad \alpha_m = \lambda_m \tag{36}$$

Note that $\sum_{i=0}^{m} \alpha_i = 1$. Moreover, if

$$1 \geq \lambda_1 \geq \ldots \geq \lambda_m > 0, \tag{37}$$

then all α_i are nonnegative and the objective function (35) of the m–MAD becomes a convex combination of means $\mu_i(\mathbf{x})$. Maximization of the means $\mu_i(\mathbf{x})$ is consistent with the SSD rules. Thus, the following theorem is true:

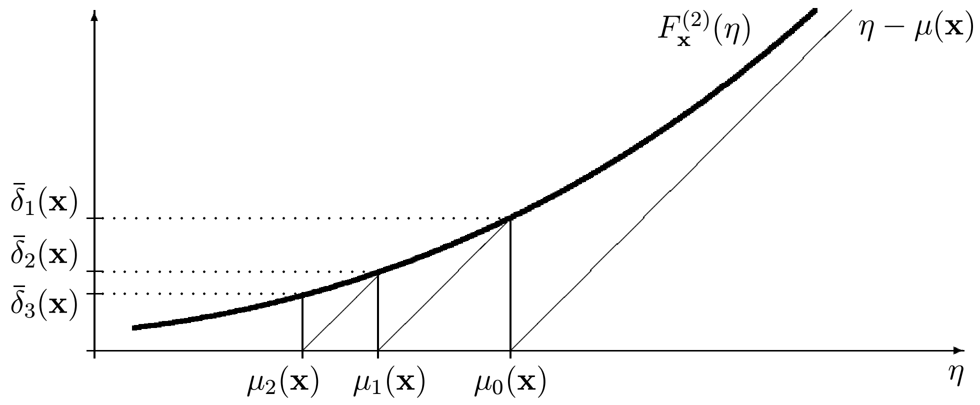

Figure 2: $\mu_i(\mathbf{x})$ and $\bar{\delta}_i(\mathbf{x})$ on the O–R diagram for $R_{\mathbf{x}}$

Theorem 1 *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 (33), $\mu_i(\mathbf{x})$ ($i = 0, \ldots, m$) are means of the corresponding random variables $R_{\mathbf{x}}^{(i)}$ defined in (32). 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 1 together with the relations (35)–(36) lead to the following theorem.

Theorem 2 *Except for the portfolios characterized by identical mean and all similar semideviations, every portfolio $\mathbf{x} \in Q$ that maximizes $\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x})$ with $0 < \lambda_m \leq \ldots \leq \lambda_1 \leq 1$ is efficient under the SSD rules.*

Proof. According to (35)–(36) and the requirement (37) 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 non-negative whereas α_m is strictly positive. Let $\mathbf{x}^0 \in Q$ maximizes $\mu(\mathbf{x}) - \sum_{i=1}^{m} \lambda_i \bar{\delta}_i(\mathbf{x})$. 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 1, $\mu_i(\mathbf{x}') \geq \mu_i(\mathbf{x}^0)$ for all $i = 0, \ldots, m$ and it follows that $\sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}') \geq \sum_{i=0}^{m} \alpha_i \mu_i(\mathbf{x}^0)$. The latter together with a 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 solution. 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 1 it holds that $\mu_m(\mathbf{x}') > \mu_m(\mathbf{x}^0)$. 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 (34) it follows that $\bar{\delta}_i(\mathbf{x}') = \bar{\delta}_i(\mathbf{x}^0)$ for all $i = 1, \ldots, m$. $\square$

According to Theorem 2, a unique optimal solution of the m–MAD problem (problem (18)) with the trade-off coefficient λ_i satisfying the requirement (37) is efficient under the SSD rules. In the case of multiple optimal solutions of (18) (similarly to the case of the original MAD model) some of them may be SSD dominated. Due to Theorem 2, 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 is quite similar in terms of a downside risk to 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 regards to the mean and deviations. Moreover, it allows to account for investor's (downside) risk aversion, and as demonstrated in the paper, it is robust considering the SSD efficiency. These advantages of the m–MAD model simultaneously maintain simplicity and linearity associated with the original MAD approach.

Both the Markowitz and MAD models are powerful portfolio optimization tools which for a given risk/return trade-off do not impose a significant information burden on an investor. This feature, considered as an advantage in certain situations, may be also 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 proposed here. 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 optimal portfolio through varying a penalty associated with "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 which have potential "large" deviations, while a more risk neutral investment attitude will result in accepting those securities. On the other hand, the modeling opportunities of the m–MAD constitute at the same time its possible drawback related to the selection of proper values for m and λ_i parameters. It is important to stress here, that if specific trade-off coefficient λ is selected in the original MAD model, then it is quite natural to use the same coefficient in the whole m–MAD model gives: $\lambda_1 = \lambda$, $\lambda_2 = \lambda^2$,..., $\lambda_m = \lambda^m$. For computational reasons it is clear that a rather small value of m should be considered. It turns out that there would be no reason to consider larger values of m even if it would be computationally acceptable. For the trade-off $\lambda < 1$ it is very likely that small values of m will have a corresponding λ_m close to 0.

In this paper we argue that a solution of the m–MAD model for a particular m, corresponds to specific (downside) risk aversion attitude of an investor. 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 $\{\mathbf{x}^0(m)\}_{m=1,2,...}$. Assuming that this process is applied to historical data, for every $\{\mathbf{x}^0(m)\}_{m=1,2,...}$ 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 to represent investors proneness to (downside) risk aversion as a function of the pcwi instead of often difficult to interpret risk or value functions. Such a representation may prove to be intuitive enough to serve as a useful tool in evaluating an investor's risk aversion attitude - information quite important when designing effective investment strategy. However, specific computational and methodological issues associated with this representation and evaluation need to be further investigated and resolved prior to its practical application.

References

Bank Administration Institute (1968), *Measuring the Investment Performance of Pension Funds for the Purpose of Inter-Fund Comparison*, Bank Administration Institute, Park Ridge, Ill.

Charnes, A., Cooper, W.W., Ferguson, R.O. (1955), "Optimal Estimation of Executive Compensation by Linear Programming", *Management Science*, **2**, 138–151.

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

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

Feinstein, C.D., Thapa, M.N. (1993), "A Reformulation of a Mean–absolute Deviation Portfolio Optimization Model", *Management Science*, **39**, 1552–1553.

Fishburn, P.C. (1964), *Decision and Value Theory*, John Wiley & Sons, New York.

Fishburn, P.C. (1980), "Stochastic Dominance and Moments of Distributions", *Mathematics of Operations Research*, **5**, 94–100.

Hanoch, G., Levy, H. (1969), "The Efficiency Analysis of Choices Involving Risk", *Rev. Economic Studies*, **36**, 335–346.

Konno, H. (1990), "Piecewise Linear Risk Function and Portfolio Optimization", *Journal of the Operations Research Society of Japan*, **33**, 139–156.

Konno, H., Yamazaki, H. (1991), "Mean–Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market", *Management Science*, **37**, 519–531.

Kroll, Y., Levy, H., Markowitz, H.M. (1984), "Mean–Variance Versus Direct Utility Maximization", *Journal of Finance*, **39**, 47–62.

Levy, H. (1992), "Stochastic Dominance and Expected Utility: Survey and Analysis", *Management Science*, **38**, 555–593.

Markowitz, H.M. (1952), "Portfolio Selection", *Journal of Finance*, **7**, 77–91.

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

Markowitz, H.M. (1987), *Mean–Variance Analysis in Portfolio Choice and Capital Markets*, Blackwell, Oxford.

Ogryczak, W., Ruszczyński, A. (1997), "From Stochastic Dominance to Mean–Risk Models: Semideviations as Risk Measures", Interim Report IR–97–027, IIASA, Laxenburg (*European Journal of Operational Research*, in print).

Ogryczak, W., Ruszczyński, A. (1998), "On Stochastic Dominance and Mean–Semideviation Models", Research Report RRR 7–98, RUTCOR, Rutgers University, Piscataway NJ.

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

Rothschild, M., Stiglitz, J.E. (1969), "Increasing Risk: I. A Definition", *Journal of Economic Theory*, **2**, 225–243.

Sharpe W.F. (1971), "A Linear Programming Approximation for the General Portfolio Analysis Problem", *Journal of Financial and Quantitative Analysis*, **6**, 1263–1275.

Sharpe, W.F. (1971a), "Mean-Absolute Deviation Characteristic Lines for Securities and Portfolios", *Management Science*, **18**, B1–B13.

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*, D.C.Heath, Lexington, MA.

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