Control and Cybernetics

vol. 28 (1999) No. 4

A Recursive Procedure for Selecting Optimal Portfolio According to the MAD Model

by

Wojtek Michalowski and **Włodzimierz Ogryczak**

International Institute for Applied Systems Analysis, DAS Project, Laxenburg, A–2361, Austria
e-mail: wojtek@business.carleton.ca

Warsaw University, Department of Mathematics and Computer Science, Banacha 2, 02–097 Warsaw, Poland
e-mail: ogryczak@mimuw.edu.pl

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 easy to solve linear programming program. In this paper we present a recursive procedure which allows to identify optimal portfolio of the MAD model depending on investor's downside risk aversion.

Keywords: Portfolio Optimization, Downside Risk Aversion, Linear Programming

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.

Following the seminal work by Markowitz (1952), this 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 (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.

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). On the other hand, 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, 1999). 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 proposed procedure of recursive identification of optimal solution of the MAD model, such that (downside) risk aversion of an investor is accounted for. The paper concludes with a discussion.

2. The MAD model

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, which is equivalent to assigning a nonnegative weight to each variable representing a security. During the investment period, a security generates a certain (random) rate of return so at the end of the period, the change of capital invested is measured by the weighted average of the returns. In mathematical terms, for selecting security weights, an investor needs to solve a model consisting of a set of linear constraints, one of which state that the weights must sum to one.

Let $J = \{1, 2, \dots, n\}$ denotes a set of securities considered for an investment. For each security $j \in J$, its rate of return is represented by a random variable R_j with a given mean $\mu_j = E\{R_j\}$.

Further, let $\mathbf{x} = (x_j)_{j=1,2,\ldots,n}$ denotes 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.:

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

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:

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

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:

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

This analysis is conducted with a so-called *critical line approach* (Markowitz, 1987), by solving parametric problem (3) with changing values of trade-off coefficient $\lambda > 0$. 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) involves consideration of a quadratic objective function.

One may consider 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 $R_{\mathbf{x}}$ (Pratt et al., 1995). The absolute deviation was used in the portfolio analysis (Sharpe, 1971a, and references therein) and has been given official status as a recommended measure of dispersion by the Bank Administration Institute (1968). Konno and Yamazaki (1991) presented the complete portfolio optimization model based on the absolute deviation as a risk measure, so-called MAD model, and they validated it by experiments on the Tokyo stock market. The MAD model does not require any specific type of return distributions, which enabled 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.

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):

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

Hence, taking into account (5), the following parametric optimization problem will be called the MAD model:

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

Simplicity and computational robustness are perceived as the most important advantages of the MAD model.

Following Konno and Yamazaki (1991), r_{jt} is the realization of random variable R_j during period t (where $t = 1, \ldots, T$) which values are 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. Therefore:

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

and, according to (5),

$$\bar{\delta}(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} d_t$$

where d_t $(t = 1, \ldots, T)$ is the downside deviation for the realization of portfolio $\mathbf{x}$ during period t, i.e.:

$$d_t = \max \{ \sum_{j=1}^{n} \mu_j x_j - \sum_{j=1}^{n} r_{jt} x_j \, , \, 0 \} = \max \{ \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \, , \, 0 \} \tag{7}$$

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

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda}{T} \sum_{t=1}^{T} d_t \tag{8}$$

subject to

$$\mathbf{x} \in Q \tag{9}$$

$$d_t \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \ldots, T \tag{10}$$

$$d_t \geq 0 \quad \text{for } t = 1, \ldots, T \tag{11}$$

where inequalities (10)–(11) guarantee that the optimal values of variables d_t satisfy (7).

The LP formulation (8)–(11) 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.

Recently, the MAD model was further validated by Ogryczak and Ruszczyński (1999) 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 this concept 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 (1998, 1999).

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 $F_{\mathbf{x}}^{(1)}$ is defined as the right-continuous cumulative distribution function itself:

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

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 the *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 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 $F_{\mathbf{x}'}^{(2)}(\eta) \leq F_{\mathbf{x}''}^{(2)}(\eta)$ 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. Ogryczak and Ruszczyński (1999) showed partial consistency of the MAD model with the SSD relation in the sense that, 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. This implies that 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), some of them may be SSD dominated. Exactly, 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. The recursive extension of the MAD model presented in this paper provides such a specification.

3. Recursive optimization with the MAD model

The MAD model (6) measures downside risk but it does not properly account for risk aversion attitude. Absolute semideviation (according to definition (5)) averages deviations and treats as equivalent a situation with low probability large deviation and a situation with high probability small deviation. This can be illustrated with two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as:

$$P\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 0.5, & \xi = -20 \\ 0.5, & \xi = 20 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

and

$$P\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 0.01, & \xi = -1000 \\ 0.98, & \xi = 0 \\ 0.01, & \xi = 1000 \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

Note that $\mu(\mathbf{x}') = \mu(\mathbf{x}'') = 0$ and $\bar{\delta}(\mathbf{x}') = \bar{\delta}(\mathbf{x}'') = 10$. Hence, two random variables are identical from the viewpoint of the MAD model. Nevertheless, $R_{\mathbf{x}'} \succ_{SSD} R_{\mathbf{x}''}$ and $R_{\mathbf{x}'}$ is strictly preferred to $R_{\mathbf{x}''}$ within all risk-averse preference models.

In order to account for downside risk aversion attitude, one needs to differentiate between different levels of deviations, and to penalize "larger" ones. 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:

$$\max\ \{\mu(\mathbf{x}) - \sum_{i=1}^{m} (\prod_{k=1}^{i} \tau_k)\bar{\delta}_i(\mathbf{x}) \ : \quad \mathbf{x} \in Q\} \tag{14}$$

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

$$\lambda_i = \prod_{k=1}^{i} \tau_k \quad \text{for } i = 1, \ldots, m \tag{15}$$

one gets the model:

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

where $\lambda_1 > 0, \ldots, \lambda_m > 0$ are the model parameters. Hereafter, we will refer to the problem (16) as the recursive m–level MAD model (or m–MAD for short).

Note that two random variables (12) and (13), which were identical from the viewpoint of the original MAD model, are no longer identical from the perspective of the m–MAD model (for any $m > 1$).

The parameters λ_i in the m–MAD model represent corresponding trade-offs for different perceptions of downside risk. Using formula (15), they can be easily derived from known trade-off coefficients τ_i. If specific value of 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 $\lambda_1 = \lambda$, $\lambda_2 = \lambda^2, \ldots, \lambda_m = \lambda^m$.

Recall that 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, due to formula (15), one gets:

$$1 \geq \lambda_1 \geq \ldots \geq \lambda_m > 0. \tag{17}$$

Vice versa, having defined coefficients λ_i satisfying (17), due to formula (15), one gets $0 < \tau_i \leq 1$ for $i = 1, \ldots, m$. Thus the m–MAD model is consistent with the SSD relation provided that the coefficients λ_i satisfy the condition (17). The issues of SSD efficiency of the m–MAD solutions are discussed in Michalowski and Ogryczak (1998).

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 the coefficients λ_i satisfy the condition (17), 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:

$$\max \sum_{j=1}^{n} \mu_j x_j - \frac{\lambda_1}{T} \sum_{t=1}^{T} d_{t1} - \frac{\lambda_2}{T} \sum_{t=1}^{T} d_{t2} \tag{18}$$

subject to

$$\mathbf{x} \in Q \tag{19}$$

$$d_{t1} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j \quad \text{for } t = 1, \ldots, T \tag{20}$$

$$d_{t2} \geq \sum_{j=1}^{n} (\mu_j - r_{jt}) x_j - \frac{1}{T} \sum_{l=1}^{T} d_{l1} \quad \text{for } t = 1, \ldots, T \tag{21}$$

$$d_{t1} \geq 0, \ d_{t2} \geq 0 \quad \text{for } t = 1, \ldots, T \tag{22}$$

The above formulation differs from (8)–(11) 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 (21) linking these variables together (similar to equations (10) 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:

$$\max \; z_0 + \sum_{i=1}^{m} \lambda_i z_i \tag{23}$$

subject to

$$\mathbf{x} \in Q \tag{24}$$

$$z_0 - \sum_{j=1}^{n} \mu_j x_j = 0 \tag{25}$$

$$T z_i + \sum_{t=1}^{T} d_{ti} = 0 \quad \text{for } i = 1, \ldots, m \tag{26}$$

$$d_{ti} - \sum_{k=0}^{i-1} z_k + \sum_{j=1}^{n} r_{jt} x_j \geq 0 \quad \text{for } t = 1, \ldots, T; \; i = 1, \ldots, m \tag{27}$$

$$d_{ti} \geq 0 \quad \text{for } t = 1, \ldots, T; \; i = 1, \ldots, m \tag{28}$$

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 (25)–(26) need to be introduced to define these variables. A number of nonzero coefficients in (27) can be further reduced if repetitions of coefficients r_{jt} in several groups of inequalities (27) 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.

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):

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

and

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

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 as well as the MAD models. It turns out, however, that $R_{\mathbf{x}''}$ has a longer and "heavier" tail to the left of the mean which can be demonstrated by comparing third moments of the random variables or their $F^{(2)}$ functions:

$$F^{(2)}_{\mathbf{x}'}(\eta) = \begin{cases} 0, & \eta \in (-\infty, 0] \\ 0.2\eta, & \eta \in (0, 1] \\ 0.3(\eta - 1) + 0.2, & \eta \in (1, 2] \\ 0.7(\eta - 2) + 0.5, & \eta \in (2, 7] \\ \eta - 3, & \eta \in (7, \infty) \end{cases}$$

and

$$F^{(2)}_{\mathbf{x}''}(\eta) = \begin{cases} 0, & \eta \in (-\infty, -1] \\ 0.3(\eta + 1), & \eta \in (-1, 4] \\ 0.7(\eta - 4) + 1.5, & \eta \in (4, 5] \\ 0.8(\eta - 5) + 2.2, & \eta \in (5, 6] \\ \eta - 3, & \eta \in (6, \infty] \end{cases}$$

One may notice that neither $R_{\mathbf{x}'}$ dominates $R_{\mathbf{x}''}$ nor $R_{\mathbf{x}''}$ dominates $R_{\mathbf{x}'}$ under the SSD rules but $F^{(2)}_{\mathbf{x}'}(\eta) \leq F^{(2)}_{\mathbf{x}''}(\eta)$ for all $\eta \leq 3$ and the inequality is strict for all $-1 < \eta < 3$. Thus $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ in a downside risk aversion context. Simple arithmetic shows that

$$\begin{array}{ll} \bar{\delta}_2(\mathbf{x}') = 0.44, & \bar{\delta}_2(\mathbf{x}'') = 0.84 \\ \bar{\delta}_3(\mathbf{x}') = 0.308, & \bar{\delta}_3(\mathbf{x}'') = 0.588 \\ \bar{\delta}_4(\mathbf{x}') = 0.2156, & \bar{\delta}_4(\mathbf{x}'') = 0.4116 \\ \bar{\delta}_i(\mathbf{x}') = 0.2(3 - \sum_{k=1}^{i-1} \bar{\delta}_k(\mathbf{x}')), & \bar{\delta}_i(\mathbf{x}'') = 0.3(4 - \sum_{k=1}^{i-1} \bar{\delta}_k(\mathbf{x}'')) \quad \text{for } i \geq 5 \end{array}$$

Hence, $\bar{\delta}_i(\mathbf{x}') \leq \bar{\delta}_i(\mathbf{x}'')$ for $i < 10$ whereas $\bar{\delta}_i(\mathbf{x}') > \bar{\delta}_i(\mathbf{x}'')$ for $i \geq 10$. Nevertheless, for any $m > 1$ and λ_i satisfying (17),

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

and $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ according to the corresponding m–MAD model.

The m–MAD model allows to penalize larger downside deviations, thus providing for better modeling of the risk avert preferences. Observe that the objective function of the m–MAD model can be written in 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. Thus our extension of the MAD model is in some manner equivalent to introduction of a convex dis-utility function u of downside deviations. Specifically, 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 $\lambda_m = \ldots = \lambda_2 = 0$.

Similar extension of the MAD model for portfolio optimization was already proposed by Konno (1990) who considered a convex piecewise linear function with breakpoints proportional to the mean of $R_{\mathbf{x}}$. The comprehensive analysis of this approach is beyond the scope of this paper. However, we illustrate with a small example that a proper selection of slope parameters may prove to be quite a difficult task. Consider two finite random variables $R_{\mathbf{x}'}$ and $R_{\mathbf{x}''}$ defined as:

$$P\{R_{\mathbf{x}'} = \xi\} = \begin{cases} 1/(1+\varepsilon), & \xi = 0 \\ \varepsilon/(1+\varepsilon), & \xi = 1 \\ 0, & \text{otherwise} \end{cases} \tag{29}$$

and

$$P\{R_{\mathbf{x}''} = \xi\} = \begin{cases} 1, & \xi = 0 \\ 0, & \text{otherwise} \end{cases} \tag{30}$$

where ε is 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 $\mu(\mathbf{x}'') = \bar{\delta}(\mathbf{x}'') = 0$. Simple arithmetic shows that $R_{\mathbf{x}'}$ is preferred to $R_{\mathbf{x}''}$ in the MAD model with any $0 < \lambda \leq 1$.

Consider function u with (one) breakpoint $b_1 = 0.5\mu(\mathbf{x})$ as Konno (1990) did. This results in a model involving maximization of the objective function

$$\mu(\mathbf{x}) - \lambda_1 \bar{\delta}(\mathbf{x}) - \lambda_2 E\{\max\{0.5\mu(\mathbf{x}) - R_{\mathbf{x}}, 0\}\} \tag{31}$$

where $\lambda_1 > 0$ is the basic trade-off parameter and $\lambda_2 > 0$ is an additional parameter (a penalty for larger deviations). Then $E\{\max\{0.5\mu(\mathbf{x}') - R_{\mathbf{x}'}, 0\}\} = (0.5\varepsilon)/(1+\varepsilon)^2 = 0.5\bar{\delta}(\mathbf{x}')$. Hence, the objective function (31) for $R_{\mathbf{x}'}$ is $\mu(\mathbf{x}') - (\lambda_1 + 0.5\lambda_2)\bar{\delta}(\mathbf{x}')$ which means that λ_2 only increases the value of trade-off coefficient λ_1. It is easy to see that in the case of $\lambda_2 \geq 2(1 - \lambda_1 + \varepsilon)$, $R_{\mathbf{x}''}$ has larger value of the objective (31) than $R_{\mathbf{x}'}$.

While applying the m–MAD model to compare the random variables (29) and (30), 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, important feature of the m–MAD model is its consistency with the SSD relation.

4. 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 are amplified by a fact that it maintains 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: $\lambda_1 = \lambda$, $\lambda_2 = \lambda^2, \ldots,$ $\lambda_m = \lambda^m$. 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.

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.

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

FEINSTEIN, C.D. and 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.

HANOCH, G. and 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. and YAMAZAKI, H. (1991) Mean–absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science*, **37**, 519–531.

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.

MICHALOWSKI, W. and OGRYCZAK, W. (1998) Extending the MAD portfolio optimization model to incorporate downside risk aversion. Interim Report IR–98–41, International Institute for Applied Systems Analysis (IIASA), Laxenburg.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (1998) On stochastic dominance and mean–semideviation models. Research Report RRR 7–98, RUTCOR, Rutgers University, Piscataway NJ.

OGRYCZAK, W. and RUSZCZYŃSKI, A. (1999) From stochastic dominance to mean–risk models: Semideviations as risk measures. *European Journal of Operational Research*, **116**, 33–50.

PRATT, J.W., RAIFFA, H. and SCHLAIFER, R. (1995) *Introduction to Statistical Decision Theory*. MIT Press, Cambridge, MA.

ROTHSCHILD, M. and STIGLITZ, J.E. (1970) 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. and FINDLAY, M.C. (EDS.) (1978) *Stochastic Dominance: An Approach to Decision–Making Under Risk*. D.C.Heath, Lexington, MA.

ZENIOS, S.A. and KANG, P. (1993) Mean–absolute deviation portfolio optimization for mortgage-backed securities. *Annals of Operations Research*, **45**, 433–450.