ELSEVIER Journal of Banking & Finance 23 (1999) 85–103

Journal of BANKING & FINANCE

A linear model for tracking error minimization

Markus Rudolf *, Hans-Jürgen Wolter, Heinz Zimmermann

Swiss Institute of Banking and Finance, University of St. Gallen, Merkurstrasse 1, CH-9000 St. Gallen, Switzerland

Received 10 April 1997; accepted 3 June 1998

Abstract

This article investigates four models for minimizing the tracking error between the returns of a portfolio and a benchmark. Due to linear performance fees of fund managers, we can argue that linear deviations give a more accurate description of the investors' risk attitude than squared deviations. All models have in common that absolute deviations are minimized instead of squared deviations as is the case for traditional optimization models. Linear programs are formulated to derive explicit solutions. The models are applied to a portfolio containing six national stock market indexes (USA, Japan, UK, Germany, France, Switzerland) and the tracking error with respect to the MSCI (Morgan Stanley Capital International Index) world stock market index is minimized. The results are compared to those of a quadratic tracking error optimization technique. The portfolio weights of the optimized portfolio and its risk/return properties are different across the models which implies that optimization models should be targeted to the specific investment objective. Finally, it is shown that linear tracking error optimization is equivalent to expected utility maximization and lower partial moment minimization. © 1999 Elsevier Science B.V. All rights reserved.

JEL classification: C63; G11

Keywords: Tracking error; MAD; Mean absolute deviation model; MinMax model; Quadratic tracking error

* Corresponding author. Tel.: 41 71 220 3066; fax: 41 71 223 6563; e-mail: markus.rudolf@unisg.ch

PII: S 0 3 7 8 - 4 2 6 6 (9 8) 0 0 0 7 6 - 4

1. Introduction

An important problem arising in portfolio optimization by mutual fund managers or pension funds is the implementation of passive investment strategies. This means that the objective of many investors is to track a certain benchmark return as close as possible by minimizing the sum of the squared deviations of returns on a replicating portfolio from a benchmark ("mean square model"), i.e. the tracking error volatility. The problem of minimizing the volatility of the tracking error is solved by Roll (1992). Choosing quadratic tracking error measures is common in the financial practice, because they reveal a number of desirable statistical properties. In contrast to the quadratic tracking error definition, Clarke et al. (1994) define the tracking error as the absolute "difference between the managed portfolio return and the benchmark portfolio return". This definition is due to the fact that from a practitioners point of view, quadratic objective functions are difficult to interpret. Investors are in many cases faced with investment objectives where linear or absolute deviations between portfolio and benchmark returns are more relevant or have a more intuitive interpretation. This fact was already stressed by Sharpe (1971) who suggests a linear programming approximation for portfolio optimization. More recently, Konno and Yamazaki (1991) and Speranza (1993) developed a portfolio optimization model based on mean absolute deviations instead of the volatility of the portfolio returns. However, there is no model which minimizes the tracking error defined as the linear deviations between the returns on a portfolio and a benchmark.

Linear tracking error models have several advantages compared to quadratic models. Portfolio managers are rewarded by linear performance fees (see Kritzman, 1987) based on the return difference between the portfolio and the benchmark. Furthermore, a portfolio manager attempts to avoid extreme deviations between the portfolio return and the benchmark to prevent his mandate from being revoked. Therefore, portfolio managers typically think in terms of linear and not quadratic deviations from a benchmark. We use four alternative definitions of the tracking error (TE) in this study. For both of these attitudes of portfolio managers, possible tracking error definitions are developed in this paper. They have in common that the TE is based on linear objective functions where absolute deviations between portfolio and benchmark returns are used. Speranza (1993) shows that portfolio weights under mean variance models according to Markowitz (1959) and mean absolute deviation models are identical if portfolio returns are normally distributed.

The paper is structured as follows. In Section 2, the classical quadratic techniques are reviewed. In Section 3, four linear optimization models are introduced. Section 4 contains the formulations of the linear programs for all models. An application of this approach is given in Section 5 of this paper where the tracking error between six national stock market indices and a world stock market index is minimized. In Section 6, linear tracking error measures are analyzed in the framework of expected utility maximization. There, the relationship between utility theory, lower partial moments, and tracking error optimization is investigated. A summary follows in Section 7.

2. Mean square optimization

In finance, mean square problems in linear models arise in different forms, such as portfolio replicating strategies: A portfolio is selected such that its returns track or replicate those on a pre-determined benchmark. Let Y be the vector of continuously compounded benchmark returns, X the matrix of continuously compounded returns on n assets, and β the portfolio weights to be determined, this problem may be represented by:

$$\varepsilon = Y - X\beta, \quad Y \in \mathbb{R}^T,\ X \in \mathbb{R}^{T\times n},\ \beta \in \mathbb{R}^n,\ \varepsilon \in \mathbb{R}^T, \tag{1}$$

where n is the number of assets and T the number of observations.

The sum of squared deviations between portfolio and benchmark returns, $\varepsilon'\varepsilon$, is traditionally called *tracking error variance* (Roll, 1992). The asset weights, β, are selected such that the tracking error is minimized, i.e.

$$\min_{\beta}\ \varepsilon'\varepsilon \equiv \min_{\beta}\ (Y - X\beta)'(Y - X\beta), \tag{2}$$

where "′" denotes the transposition of a matrix.

Eq. (2) represents a quadratic optimization problem, and the vector of asset weights is given by

$$\beta = \left(X'X\right)^{-1}X'Y. \tag{3}$$

The mean square model is popular because of its computational simplicity. Furthermore, the estimator β for the portfolio weights reveals the BLUE properties, i.e. it is the best linear unbiased estimator. In addition to Eq. (2), a set of linear restrictions can be included, i.e.

$$A\beta \geqslant b, \quad A \in \mathbb{R}^{k\times n},\ \beta \in \mathbb{R}^n,\ b \in \mathbb{R}^k \quad (k \text{number of restrictions}). \tag{4}$$

Examples of such restrictions are short selling restrictions (weights of the portfolio must be non-negative), or the condition that the sum of weights characterizing the portfolio must add up to unity. Furthermore, Roll (1992) imposes restrictions on the "average gain over benchmark return". Note that the BLUE properties of the estimator are lost if inequality restrictions are imposed.

3. Minimizing the tracking error: Four alternative definitions

Two alternative definitions of the tracking error are examined in the remainder of this paper. Both provide a more immediate interpretation of the optimized value of the objective function compared to the mean square approach. They have in common that, instead of squared deviations, absolute deviations between the benchmark and portfolio returns are minimized. It is easy to imagine that this reflects the investment objective of many investors more adequately than minimizing quadratic tracking error (TE) functions as described by Eq. (2).

The *first* model minimizes mean absolute deviations (MAD). The portfolio weights are determined such that the sum of the absolute deviations between the benchmark returns and the portfolio returns (which is the tracking error: $1'|X\beta - Y|$) is minimized, i.e.

$$\min_{\beta} \; 1'(|X\beta - Y|), \quad \text{where } 1' \equiv (1, \ldots, 1) \in \mathbb{R}^T. \tag{5}$$

By measuring the tracking error according to Eq. (5), the measurement unit of the objective function is percentage, whereas the dimension of the mean square objective function (Eq. (2)) is "squared percentages".

In the *second* model, the portfolio weights are determined such that the maximum deviation between portfolio and benchmark returns is minimized. This is called "MinMax" model and it represents a "worst case" protection strategy. The objective function of the MinMax optimization problem is

$$\min_{\beta} \left(\max_{t} \; |X_t\beta - Y_t| \right), \tag{6}$$

where X_t represents the row t of matrix X and Y_t the tth element of vector Y. Since in contrast to the MinMax model, in the MAD model outliers are averaged, the MinMax model is less robust against outliers than the MAD model. Furthermore, since the return deviations are squared in quadratic models, large deviations get a higher weight in quadratic models than in the MAD model. Therefore, the MAD estimator is less sensitive against outliers than the mean square models. This is consistent with the findings of Amemiya (1985, p. 73).

In addition to the MAD and MinMax model, two variants are subsequently examined. For many investors, an alternative perception of risk may be that the return on the portfolio is below the return on the benchmark portfolio. This is called "downside risk" of an investment (see Harlow, 1991). Under this perspective, tracking error minimization is restricted to the negative deviations between portfolio and benchmark returns. In case of MAD, this means that the sum of absolute deviations is minimized subject to the restriction that the portfolio return is below the benchmark return. This model is called the "mean absolute downside deviation model" (MADD). [1] Or in case of the MinMax model, the maximum negative deviation is minimized. [2] This is called the "downside MinMax model" (DMinMax).

To summarize this section, the following equations describe the four different tracking error definitions which will be used in the subsequent sections:

$$\mathrm{TE_{MAD}} = \min_{\beta}\ 1'(|X\beta - Y|), \tag{7a}$$

$$\mathrm{TE_{MADD}} = \min_{\beta}\ 1'\left(\left|\overline{X}\beta - \overline{Y}\right|\right), \quad \text{where}\ \overline{X}_t\beta < \overline{Y}_t, \tag{7b}$$

$$\mathrm{TE_{MinMax}} = \max_{t}\ |X\beta - Y|, \tag{7c}$$

$$\mathrm{TE_{DMinMax}} = \max_{t}\ \left|\overline{X}\beta - \overline{Y}\right|, \quad \text{where}\ \overline{X}_t\beta < \overline{Y}_t. \tag{7d}$$

The matrix $\overline{X}$ and the vector $\overline{Y}$ contain only those lines where the benchmark return is below the portfolio return.

4. Linear programs

Each model described in the previous section can be characterized by a linear program:

(a) *MinMax problem*: The basic MinMax problem was stated in Eq. (6). Let $z \geqslant 0$ be an upper boundary of the absolute deviations:

$$z \geqslant |X_t\beta - Y_t|, \quad t \in \{1, \ldots, T\}.$$

The following cases can be considered for each t. First, the portfolio return is higher than the benchmark return,

$$\text{Case 1}: \qquad z \geqslant X_t\beta - Y_t \geqslant 0 \iff X_t\beta - z \leqslant Y_t$$

and in the second case, the portfolio return $X_t\beta$ is below the benchmark return Y_t,

$$\text{Case 2}: \qquad -z \leqslant X_t\beta - Y_t \leqslant 0 \iff X_t\beta + z \geqslant Y_t.$$

The upper boundary z is now minimized. Notice that for $X_t\beta - Y_t \geqslant 0$, the case 1 inequality implies the second one. On the other hand, if case 2 holds, the case

[1] These tracking error definitions can be seen in accordance to the goal of portfolio managers to maximize the performance participation.

[2] Which avoids a loss of the mandate due to extremely low performances in a specific time period.

2 inequality implies the first one. Thus, the MinMax problem can be stated as follows:

$$\min_z z \quad \text{s.t. } X_t\beta - z \leqslant Y_t, \;\; X_t\beta + z \geqslant Y_t. \tag{8}$$

(b) *Downside MinMax problem*: The previous program is restricted to observations satisfying $X_t\beta \leqslant Y_t$ where the portfolio underperforms the benchmark. Therefore only the restrictions according to case 2 are relevant, and the respective linear program can be derived from Eq. (8):

$$\min_z z \quad \text{s.t. } X_t\beta + z \geqslant Y_t. \tag{9}$$

(c) *Mean Absolute Deviations (MAD)*: Let $z_t^+ \geqslant 0$ be a positive deviation and $z_t^- \geqslant 0$ the absolute value of a negative deviation between the portfolio and benchmark returns. It then follows that:

$$\begin{aligned} X_t\beta - Y_t > 0 \iff X_t\beta - z_t^- = Y_t, \\ X_t\beta - Y_t < 0 \iff X_t\beta + z_t^+ = Y_t. \end{aligned} \tag{10}$$

The objective function is

$$\min \sum_{t=1}^{T} \left(z_t^+ + z_t^-\right). \tag{11}$$

Notice that either z_t^- is positive and z_t^+ is zero, or z_t^- is zero and z_t^+ positive, implying that a positive deviation leads to $z_t^- = 0$ and a negative deviation implies $z_t^+ = 0$. The two equations can thus be aggregated to one restriction:

$$X_t\beta + z_t^+ - z_t^- = Y_t. \tag{12}$$

(d) *Mean Absolute Downside Deviation (MADD)*: If investors are concerned about negative deviations between the portfolio and the benchmark, the z_t^+ are dropped from the optimization. The problem can then be stated as

$$\min \sum_{t=1}^{T} z_t^- \quad \text{s.t. } X_t\beta + z_t^+ \geqslant Y_t. \tag{13}$$

Eqs. (8), (9), (11)–(13) provide only the basic description of the optimization problems. The linear programs considered here reveal a considerable complexity: Let T be the number of observation periods. Then the initial tableau in the Downside MinMax, the MAD and the MADD problem has $T+1$ rows; in the MinMax model it has even $2T+1$ rows. However, with modern personal computers and optimization software this can easily be solved. Further restrictions can be added, e.g. non-negative portfolio weights or the restriction

Table 1
Risk/return characteristics of MSCI total return indices in terms of US$

Index	Whole observation period: March 1987 to April 1996			In-the-sample period: March 1987 to April 1992			Out-of-the-sample period: May 1992 to April 1996		
	μ^a	σ^b	β^c	μ^a	σ^b	β^c	μ^a	σ^b	β^c
MSCI USA	9.24	13.91	0.68	6.89	17.43	0.72	11.82	7.30	0.48
MSCI Japan	2.94	26.61	1.45	−3.77	29.63	1.37	12.00	22.14	1.79
MSCI UK	6.80	19.16	1.07	5.61	22.06	1.05	4.85	13.21	1.10
MSCI Germany	7.44	20.41	0.85	5.91	23.72	0.87	9.28	15.33	0.82
MSCI France	6.43	20.50	0.92	5.68	23.50	0.90	5.73	15.76	1.01
MSCI Switzerland	12.85	17.33	0.86	4.13	19.11	0.86	23.34	14.29	0.78
MSCI World	6.74	14.26	1.00	2.97	17.17	1.00	10.93	9.19	1.00

[a] Average return in % p.a.
[b] Standard deviation in % p.a.
[c] Beta to MSCI World.

that the sum of portfolio weights is unity. Finally, expected return restrictions of the portfolio can be implemented.

5. An empirical application

The models of Section 4 are illustrated by an empirical example. It is assumed that an investor wants to optimize an internationally diversified portfolio. His objective is to minimize the tracking error between his portfolio and the MSCI world stock market index which will be used as the benchmark. His portfolio consists of stocks from the following countries: USA, Japan, United Kingdom, Germany, France, and Switzerland. Total returns on national MSCI indexes are used, and their statistical characteristics are displayed in Table 1. In addition to the full sample period (March 1987 to April 1996) two subperiods are considered: From March 1987 to April 1992 (62 months), and from May 1992 to April 1996 (48 months). The optimized portfolios are based on the data of the first subperiod. The persistency of the portfolio weights is tested with the data of the second subperiod.

The returns are calculated in US$ and are based on monthly observations from March 1987 to April 1996 which gives a total of 110 observations per series. The analysis is executed with each of the four tracking error models presented in Section 4. Short selling is excluded throughout the analysis. [3] It is furthermore required that the portfolio weights add up to unity. In addition,

[3] Rudolf (1994) shows without short sale restriction how to construct the model and some empirical results. The results are similar. However, for small markets negative fractions can be observed which complicates the implementation and may contradict to legal restrictions.

Table 2
Optimized portfolio weights based on the in-the-sample period (March 1987 to April 1992) referring to different optimization models, all figures in %

Model	USA	JAP	UK	D	F	CH	Value of objective function
MinMax	35.21	30.73	18.09	12.07	0.00	3.90	1.18
DminMax	30.98	28.32	23.52	6.25	0.00	10.92	1.17
MAD	41.00	35.89	12.74	8.57	1.79	0.00	21.81 (0.35)[a]
MADD	41.78	34.32	13.99	7.25	2.66	0.00	11.03 (0.75)[b]
Quadratic TE	39.58	34.74	13.00	7.44	2.30	2.95	0.45[c]

[a] Sum of absolute deviations (average of absolute deviations in brackets).
[b] Sum of absolute downside deviations (average of absolute downside deviations in brackets), 29 observations of the portfolio reveal returns below the MSCI World benchmark.
[c] Square root of the average of the squared deviations between the returns on the MSCI world stock market index and the portfolio.

the quadratic model is used to estimate the vector of portfolio weights (see Eqs. (2) and (4), respectively) where the sum of squared deviations is minimized with restricted short sales. The structure of the optimal portfolio and the value of optimized objective function will be compared between the five models. The results are summarized in Table 2.

It is a well-known observation that the portfolio weights associated to US and Japanese stock market have the biggest explanatory power with respect to the world stock market index; the reason is that the MSCI index family is value weighted, and the US and Japanese stock markets are the highest capitalized worldwide. Among the European stock markets, United Kingdom is dominant. The MAD and quadratic tracking error results are remarkably similar; this is due to the fact that both models minimize (absolute and, respectively, squared) sums of positive and negative deviations. This is consistent with the perceptions of Konno and Yamazaki (1991) who prove that the optimization results are the same whether linear or quadratic objective functions are used when the joint probability distribution of the benchmark and portfolio returns is exactly normal. Due to the non-normality of the returns, slight differences between the MAD and the quadratic tracking error model occur. It can be observed that the portfolio weights substantially differ between the MAD/quadratic model and the MinMax/DMinMax model, which is not surprising due to the differences of the four tracking error definitions. The fraction of the US market for instance differs by 10%.

In Table 3, average returns, standard deviations, and betas of the portfolios are displayed. Given the previous results, it is surprising how similar the risk characteristics are. Except for the MAD and the quadratic tracking error model, average returns are, however, considerably different. The Sharpe and Treynor ratio clearly show that the risk/return characteristics of the DMin-

Table 3
Risk/return characteristics of optimized portfolios, all figures in %

Model	Average return[a]	Standard deviation[a]	Beta[b]	Sharpe ratio[c]	Treynor ratio[d]
MinMax	3.62	17.42	1.01	0.04	0.61
DminMax	3.72	17.55	1.01	0.04	0.71
MAD	3.28	17.52	1.01	0.02	0.28
MADD	3.48	17.41	1.01	0.03	0.48
Quadratic TE	3.30	17.46	1.01	0.02	0.30
MSCI World Stock Market	2.97	17.17	1.00	0.03	−0.03

[a] Annualized values in %.
[b] Beta of each portfolio to the MSCI world stock market index.
[c] Ecess return (average return minus 3%) to volatility ratio.
[d] Excess return to beta ratio in %.
Observation period: March 1987 to April 1992 (62 months).

Max, MADD, and MinMax models are superior to the quadratic tracking error portfolio as well to the market portfolio. However, it is the aim of this paper to minimize the tracking error of the portfolio and not to maximize the (mean/variance) based performance.

So far, the results have revealed that the linear optimization models provide quite different portfolios than the classical quadratic optimization model. The advantage of the linear models is, however, that the value of the objective

Table 4
Values of objective function of optimized portfolios with different optimization models, all figures in %

Portfolio	Model					Number of downside deviations
	Quadratic	MinMax	Downside MinMax	MAD[a]	MADD[b]	
Quadratic	**0.45**	1.42	1.42	22.06 (0.36)	11.40 (0.38)	30
MinMax	0.61	**1.18**	1.18	31.00 (0.50)	15.10 (0.56)	27
DMinMax	0.74	1.81	**1.17**	37.99 (0.61)	18.33 (0.63)	29
MAD	0.46	1.56	1.56	**21.81** (0.35)	11.36 (0.38)	30
MADD	0.47	1.39	1.39	22.14 (0.36)	**11.03** (0.38)	29

[a] Sum of absolute deviations (average of absolute deviations in brackets).
[b] Sum of absolute downside deviations (average of absolute downside deviations in brackets).
Observation period: March 1987 to April 1992 (62 months).

function provides an intuitive and immediate interpretation. It is easier for an investor to determine his attitude towards risk if he can express the tracking error in terms of absolute deviations from the benchmark rather than squared deviations. A comparison of the optimized values of the objective functions is provided by Table 4.

The minimum values of the objective function across the five models are in bold print. The lowest tracking error of a portfolio with respect to the benchmark using the quadratic TE model is, not surprisingly, provided by the quadratic TE portfolio. However, more interesting insights emerge from the objective functions of the four alternative models. For example, if an investor is concerned about the maximum absolute downside deviation (DMinMax), the minimum "risk" he takes is 1.18% by holding the downside MinMax portfolio. This is lower than the Quadratic TE portfolio which has a total absolute downside deviation (DMinMax) of 1.42% from the benchmark. If the investment objective is the minimum sum of absolute downside deviations (MADD), the total deviation can be slightly reduced from 11.40% to 11.03% if the Quadratic TE portfolio would be substituted by the MADD portfolio.

Figs. 1 and 2 show the deviations between the returns of portfolios and the returns on the benchmark for each month. The results of the Quadratic TE portfolio are represented by gray bars. In Fig. 1 the deviations of the MinMax portfolio returns from the benchmark are compared to those of the Quadratic TE portfolio. The highest deviation between the portfolio and the benchmark (the MSCI world stock index) occurs in December 1988. The distance is 1.18%. The largest deviation between the Quadratic TE portfolio and the benchmark is 1.42% in April 1992. Fig. 2 displays the deviations generated by the DMinMax and the Quadratic TE portfolios. The downside deviations can be slightly reduced. The highest negative deviation between the Downside MinMax portfolio and the benchmark occurs in December 1988 with 1.17% whereas the Quadratic TE portfolio has a deviation of 1.42% in April 1992.

Out-of-the-sample tests are performed next; the results are displayed in Tables 5 and 6. The "old" portfolio is the optimum tracking error portfolio based on return data of the period March 1987 to April 1992, whereas the "new" portfolio is calculated with return data between May 1992 to April 1996. The tracking errors are shown in the last column of Table 5. These figures are compared to the tracking errors of the out-of-the-sample optimum portfolios. For example, in the out-of-the-sample period May 1992 to April 1996, the "new" optimum MinMax portfolio reveals a tracking error of 1.09%, whereas the "old" MinMax portfolio exhibits an error of 1.93%, which is substantially more. This is due to the significant change of the portfolio fractions: While the weight of the US has increased, the weights of Japan and Germany have decreased. Similar observations can be made for the other strategies. Table 6 shows the risk/return characteristics of the "new" in-the-sample and the "old" out-of-the-sample portfolios in the out-of-the-sample

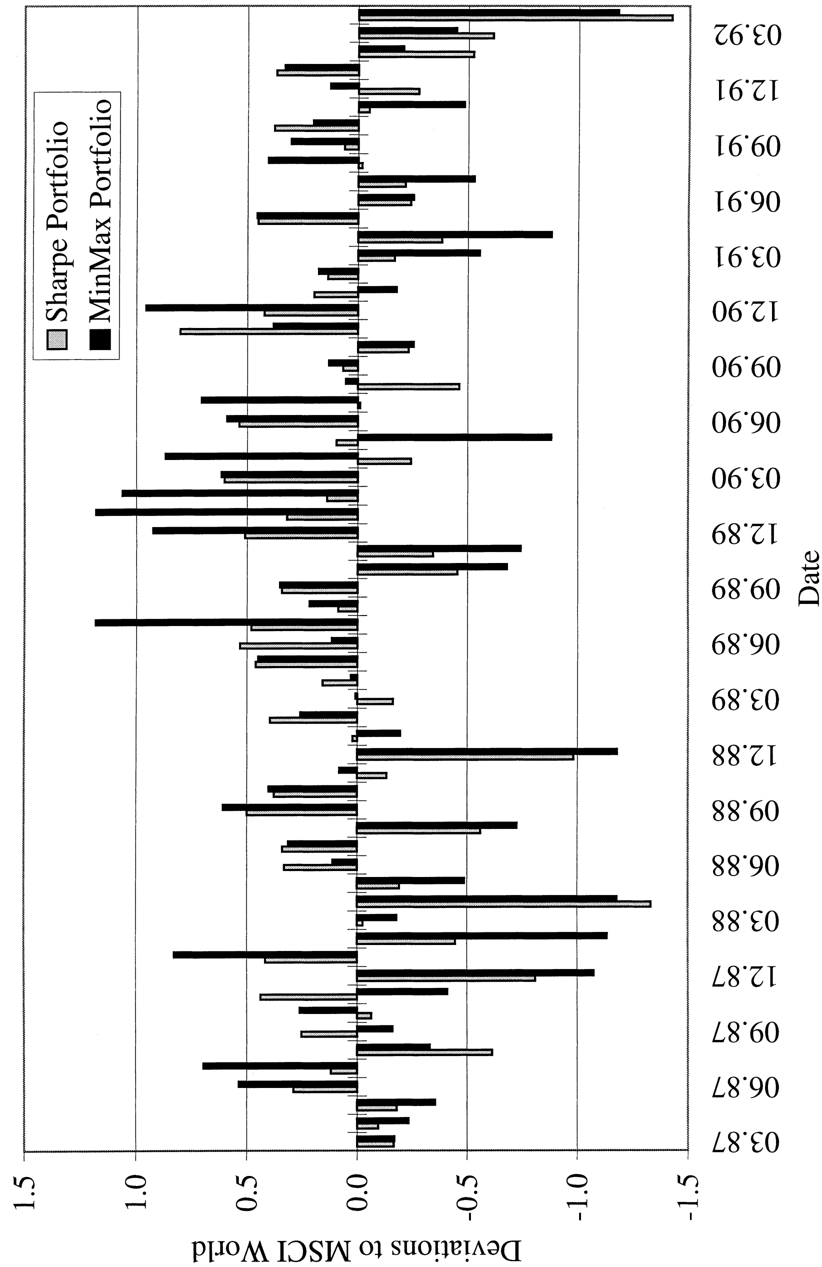

Fig. 1. A linear model for tracking error minimization.

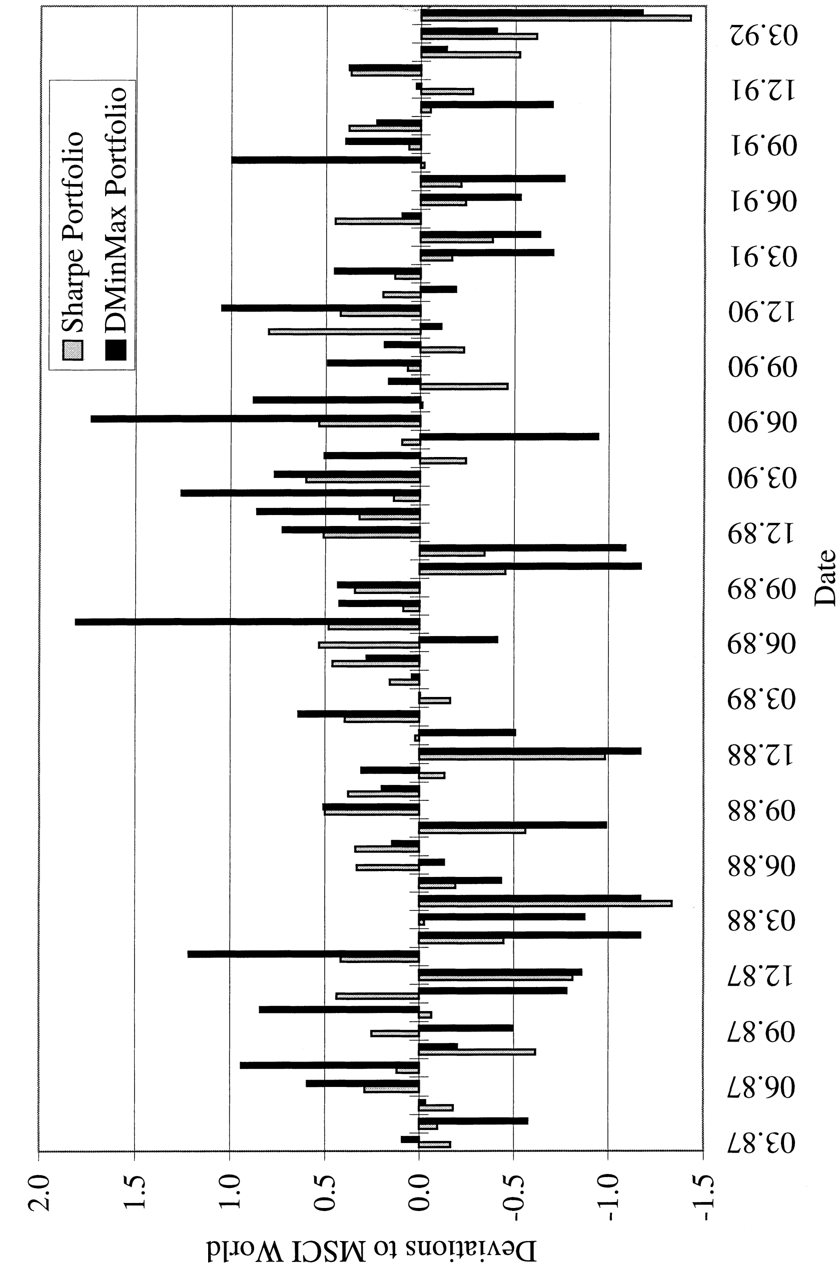

Fig. 2. A linear model for tracking error minimization.

Table 5
Optimized portfolio weights based on the out-of-sample period (May 1992 to April 1996) referring to different optimization models, all figures in %

Model	USA	JAP	UK	D	F	CH	Objective function of the "new" portfolio[a]	Objective function of the "old" portfolio[b]
MinMax	42.21	23.61	16.50	2.97	7.82	6.89	1.09	1.93
DMinMax	45.23	24.50	0.00	2.25	11.29	16.73	0.96	2.13
MAD	48.50	26.23	10.71	8.60	4.91	1.05	13.67 (0.22)[c]	26.44 (0.55)[c]
MADD	47.83	27.14	2.79	11.50	3.73	7.02	6.02 (0.27)[d]	12.77 (0.56)[e]
Quadratic TE	45.95	25.53	13.21	8.81	3.98	2.52	0.44[f]	0.74[f]

[a] The values are calculated by portfolio fractions based on March 1987 to April 1992 (62 months, in-the-sample period) and the return data from the in-the-sample period.
[b] The values are calculated by portfolio fractions based on March 1987 to April 1992 (62 months, in-the-sample period) and the return data from the out-of-sample period.
[c] Sum of absolute deviations (average of absolute deviations in brackets).
[d] Sum of absolute downside deviations (average of absolute downside deviations in brackets), 22 downside deviations of the portfolio from the MSCI World returns.
[e] Absolute downside deviations (average of absolute downside deviations in brackets), 23 downside deviations of the portfolio from the MSCI World returns.
[f] Squareroot of the average of the squared deviations between the returns on the MSCI world stock market index and the portfolio.

period (May 1992 to April 1996). While the average returns are virtually identical, the risk measures are higher for the in-the-sample portfolios. Obviously, suboptimal tracking error portfolios also lead to inefficient mean variance portfolios. Since expected utility theory is the theoretical base of mean variance portfolio selection, it is important to show that linear tracking error models are consistent with expected utility maximization. This is presented in Section 6.

6. Tracking error, utility theory and lower partial moments

In this section we demonstrate the consistency between our tracking error definitions (see Eq. (7)), expected utility and lower partial moments. According to Harlow (1991) the nth lower partial moment is defined by

$$\mathrm{LPM}(n, r^*, p) = \sum_{\varepsilon_k \leqslant r^*} (r^* - \varepsilon_k)^n p(\varepsilon_k) \geqslant 0, \tag{14}$$

where r^* is the threshold return, ε_k the difference between the portfolio return and the benchmark (see Eq. (1)), and $\mathrm{p}(\varepsilon_k)$ the probability of ε_k. By

Table 6
Risk/return characteristics of in- and out-of-the-sample optimized portfolios

Model	Portfolio based on the in-the-sample period – "new" portfolio			Portfolio based on the out-of-the-sample period – "old" portfolio		
	Average return[a]	Standard dev.[a]	Beta[b]	Average return[a]	Standard dev.[a]	Beta[b]
MinMax	11.56	10.10	1.06			0.96
DMinMax	11.47	9.90	1.05	11.82	11.44	8.94
MAD	12.27	10.05	1.06	13.50	8.77	0.92
MADD	11.20	10.19	1.07	11.32	8.88	0.95
Quadratic TE	11.16	10.03	1.06	12.31	8.73	0.93

[a] Annualized values in %.
[b] Beta of each portfolio to the MSCI world stock market index.
Observation period: May 1992 to April 1996.

choosing $r^* = 0$, the first lower partial moment can be derived from Eq. (14) as

$$\mathrm{LPM}(1, 0, p) = \sum_{\varepsilon_k \leqslant 0} (-\varepsilon_k) p(\varepsilon_k), \tag{15}$$

which corresponds to the definition of the MADD tracking error in Eq. (7b). A slight modification of Eq. (15) allows us to express the MAD problem in terms of LPMs:

$$\begin{aligned} \mathrm{TE_{MAD}} &= \sum_{\varepsilon_k} |0 - \varepsilon_k| p(\varepsilon_k) = -\sum_{\varepsilon_k < 0} \varepsilon_k p(\varepsilon_k) + \sum_{\varepsilon_k \geqslant 0} \varepsilon_k p(\varepsilon_k) \\ &= \sum_{\varepsilon_k < 0} (-\varepsilon_k)[p(-\varepsilon_k) + p(\varepsilon_k)] = \sum_{\varepsilon_k < 0} (-\varepsilon_k) p^*(-\varepsilon_k) \\ &= \mathrm{LPM}(1, 0, p^*), \end{aligned} \tag{16}$$

where $p^*(\varepsilon_k) \equiv p(-\varepsilon_k) + p(\varepsilon_k)$.

Eq. (16) allows us to reinterpret $\mathrm{TE_{MAD}}$ as the first lower partial moment with threshold return zero and a new probability p^*. Minimizing MAD and MADD is equal to minimizing the first lower partial moment using the same probability p^*. As Bawa (1975, 1978) showed, minimizing the first lower partial moment LPM(1, 0, p^*) is equivalent to expected utility maximization under risk aversion.

The specific definition in Eq. (16) of p^* may be interpreted in the following way: For investors using MAD or MADD, positive and negative deviations are equivalent. Therefore, they consider the probability of the absolute distance $|\varepsilon_k|$. Following the definition in Eq. (16), this probability equals p^*.

We now show that MinMax and DMinMax rules can as well be related to lower partial moments. Figs. 3 and 4 depict the probability distributions of two

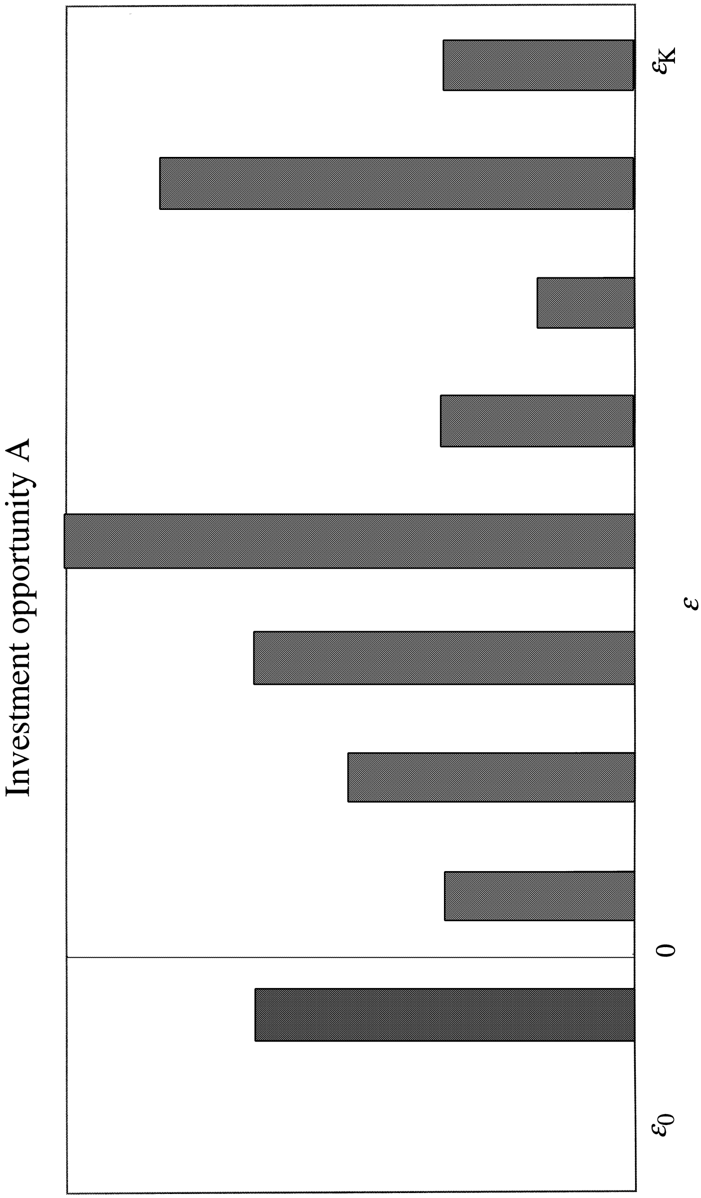

Fig. 3. A linear model for tracking error minimization.

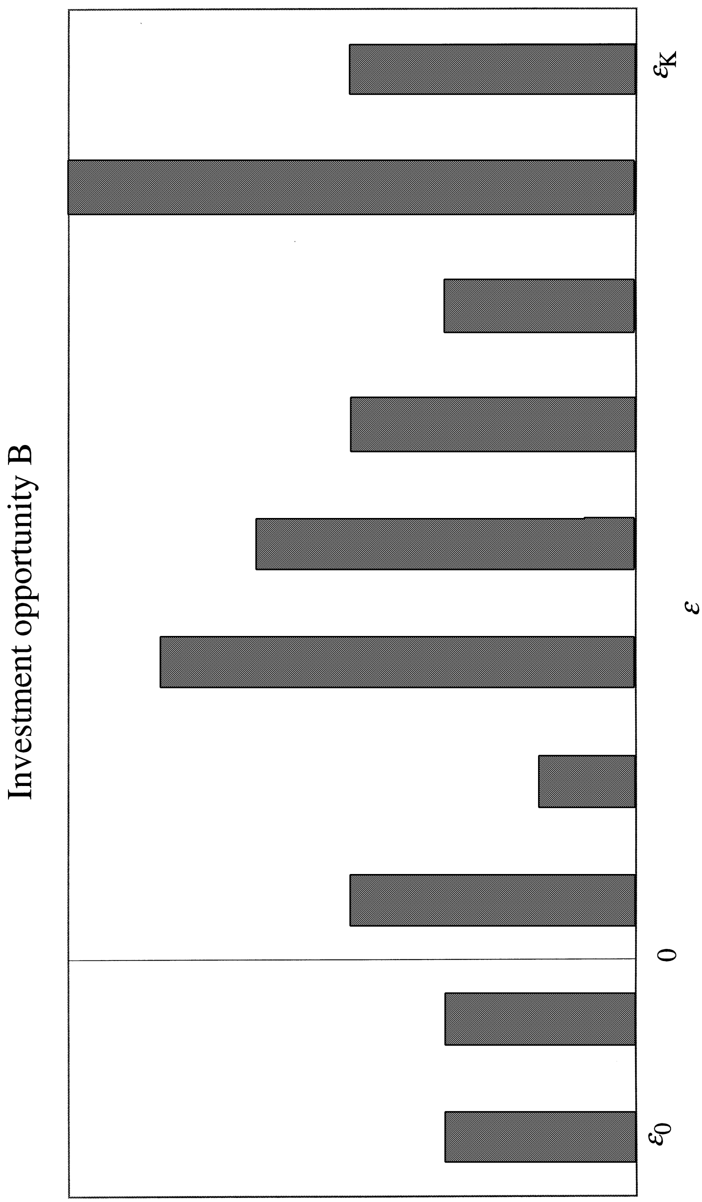

Fig. 4. A linear model for tracking error minimization.

investment alternatives, A and B. There it is assumed that ε_0 is the smallest realization of the return deviation ε for both investment opportunities, and that the probability for ε_0 in investment opportunity A, $p_A(\varepsilon_0)$, equals zero. Since B has a higher tracking error, $\varepsilon_0 < 0$, with positive probability, a DMinMax investor would prefer investment A to investment B. We now show that minimizing a specific LPM leads to the same decision. First, let us define a standardized LPM of infinite order as

$$\mathrm{sLPM}(\infty, p) = \lim_{n\to\infty} \frac{\mathrm{LPM}(n, 0, p)}{(-\varepsilon_0)^n}. \tag{17}$$

The MADD and the LPM criterion lead to the same result (A is preferred to B) if the difference between the standardized lower partial moments, i.e. $\mathrm{sLPM_B}(\infty, p) - \mathrm{sLPM_A}(\infty, p)$, is positive. This is true because

$$\begin{aligned}&\mathrm{sLPM_B}(\infty, p) - \mathrm{sLPM_A}(\infty, p)\\ &= \lim_{n\to\infty}\left[\sum_{\varepsilon_0<\varepsilon_k<0}(p_\mathrm{B}(\varepsilon_k) - p_\mathrm{A}(\varepsilon_k))\underbrace{\left(\frac{-\varepsilon_k}{-\varepsilon_0}\right)^n}_{<1}\right] + p_\mathrm{B}(\varepsilon_0)\left(\frac{-\varepsilon_0}{-\varepsilon_0}\right)^n \\ &= p_\mathrm{B}(\varepsilon_0) > 0.\end{aligned} \tag{18}$$

As a result, the DMinMax criterion leads to the same decision as LPM minimization.

In analogy to the previous case, it can be shown that the MinMax tracking error is equivalent to LPM minimization. We distinguish two cases: First, let the maximum negative deviation be larger than the maximum positive deviation between the benchmark and portfolio return, i.e. $|\varepsilon_0| \geqslant |\varepsilon_k|$. Define

$$p^*(\varepsilon_k) \equiv p(-\varepsilon_k) + p(\varepsilon_k), \quad \varepsilon_0 < \varepsilon_k < 0 \text{ with } p_B^*(\varepsilon_0) > 0, \quad p_A^*(\varepsilon_0) = 0. \tag{19}$$

It follows that

$$\begin{aligned}&\mathrm{sLPM_B}(\infty, p^*) - \mathrm{sLPM_A}(\infty, p^*)\\ &= \lim_{n\infty}\left[\sum_{\varepsilon_0<\varepsilon_k<0}(p^*_\mathrm{B}(\varepsilon_k) - p^*_\mathrm{A}(\varepsilon_k))\underbrace{\left(\frac{-\varepsilon_k}{-\varepsilon_0}\right)^n}_{<1}\right] + p^*_\mathrm{B}(\varepsilon_0)\left(\frac{-\varepsilon_0}{-\varepsilon_0}\right)^n \\ &= p^*_\mathrm{B}(\varepsilon_0) > 0.\end{aligned} \tag{20}$$

A positive difference implies that investment alternative A is preferred to alternative B under the LPM rule. The second case is $|\varepsilon_0| \leqslant |\varepsilon_k|$. The same result occurs after redefining

$$\mathrm{sLPM}(\infty) = \lim_{n\infty} \frac{\mathrm{LPM}(n,0)}{\varepsilon_K^n}. \tag{21}$$

This proves the equivalence between MinMax decision rules with rules based on LPM. However, there is no correspondence between LPMs of infinite order and expected utility maximization.

7. Summary

This paper is motivated by the fact that in practice, tracking error is defined as linear deviation between the returns of a portfolio and a benchmark. This is in contrast to the quadratic tracking error definition as it is typically defined in academic studies. We show that linear tracking error models are consistent with expected utility maximization. This article investigates four linear models for minimizing the tracking error between the returns of a portfolio and a benchmark:

1. The MinMax model, where the maximum absolute deviation between portfolio and benchmark returns is minimized.
2. The Downside MinMax model, where the maximum negative deviation is minimized.
3. The Mean Absolute Deviation (MAD) model, where the sum of absolute deviations is minimized.
4. The Mean Absolute Downside deviations (MADD) model, where the sum of negative deviations is minimized.

All models have in common that absolute deviations are minimized instead of squared deviations as is the case in the TEV (tracking error volatility) model by Roll (1992). Linear programs are formulated to derive explicit solutions. The models are applied to a portfolio containing six national stock market indexes (USA, JAP, UK, D, F, CH) and the tracking error with respect to the MSCI world stock market index is minimized. The results are compared to a quadratic tracking error model. The portfolio weights of the optimized portfolio and its risk/return properties are different across the models which implies that optimization models should be targeted to the specific investment objective. It is shown that the tracking errors of the five optimal portfolios (see Table 2) when evaluated under the different tracking error measures according to Eqs. (7a)–(7d) substantially differ. This illustrates the relevance of different investment objectives and the implied objective function in the portfolio optimization process.

Acknowledgements

The authors appreciate the comments from Paul Alapat, Alfred Bühler, Joshua Coval, Thomas Kraus, Ulrich Niederer, Walter Wasserfallen, Rudi Zagst, William Ziemba and a referee of this journal.

References

Amemiya, T., 1985. Advanced Econometrics. Harvard University Press, Cambridge, MA.

Bawa, V.S., 1975. Optimal rules for ordering uncertain prospects. Journal of Financial Economics 2, 95–121.

Bawa, V.S., 1978. Safety first, stochastic dominance, and optimal portfolio choice. Journal of Financial and Quantitative Analysis 13, 255–271.

Clarke, R.C., Krase, S., Statman, M., 1994. Tracking errors, regret, and tactical asset allocation. The Journal of Portfolio Management 20 (Spring), 16–24.

Harlow, W.V., 1991. Asset allocation in a downside risk framework. Financial Analysts Journal 47 (Sept./Oct.), 28 – 40.

Konno, H., Yamazaki, H., 1991. Mean absolute deviation portfolio optimization model and its application to Tokyo stock market. Management Science 37, 519–531.

Kritzman, M.P., 1987. Incentive fees: Some problems and some solutions. Financial Analysts Journal 43 (January/February), 21–26.

Markowitz, H.M., 1959. Portfolio Selection. Wiley, New York. Reprinted and expanded by Blackwell, Cambridge, MA, 1991.

Roll, R., 1992. A mean/variance analysis of tracking error. The Journal of Portfolio Management 18 (Summer), 13–22.

Rudolf, M., 1994. Algorithms for Portfolio Optimization and Portfolio Optimization. Haupt, Bern, Stuttgart, Wien.

Sharpe, W.F., 1971. A linear programming approximation for the general portfolio analysis problem. Journal of Financial and Quantitative Analysis 6 (December), 1263–1275.

Speranza, M.G., 1993. Linear programming models for portfolio optimization. Finance 14, 107–123.