Assessing Equity Mutual Funds' Performance Using A Multicriteria Methodology: A Comparative Analysis

KONSTANTINA PENDARAKI,[*]
MICHAEL DOUMPOS,[*]
CONSTANTIN ZOPOUNIDIS[*,a]

[*]Department of Production Engineering and Management,
Technical University of Crete, Greece

Abstract

The evaluation of the performance of mutual funds (MFs) has been a very interesting research topic not only for researchers, but also for managers of financial, banking and investment institutions. This paper explores the performance of a non-parametric approach in developing MF performance evaluation models. The proposed approach is based on the UTADIS multicriteria decision aid method. The data set consists of daily data for 33 Greek domestic equity MFs, and is used to estimate the performance of the method in classifying the funds into two groups. A cross-validation procedure is employed to evaluate the predictive performance of the models and a comparison with linear discriminant analysis is also performed. The results indicate the superiority of the UTADIS method as opposed to the traditional discrimination technique.

1. Introduction

Within the E.U at present 26,512 MFs are operating, with total assets amounting to €3,503 bn (data as of 31/12/2001; Association of Greek Institutional Investors). As in Europe, the collective investments industry in Greece is growing rapidly. According to recent data of the Association of Greek Institutional Investors, there are currently 27 MF management companies in Greece managing 266 MFs, with assets amounting to €23.86 bn (data as of 29/03/2002). A decade earlier, there were only 7 MF management companies managing 7 MFs with assets amounting to €431.4 million. The American Investment Company Institute numbers more than 8,200 MFs, while the listed companies on the Stock Exchanges of NYSE and NASDAQ at the

---

[a] *Corresponding author:* Financial Engineering Laboratory, University Campus, 73100 Chania, Greece, e-mail: kostas@dpem.tuc.gr

end of 1999 were about 7,800.

This situation highlights the rapid growth of the MF market worldwide. Thus, it is very difficult for investors to choose funds according to their decision policy, the risk levels they are willing to take, and their profitability goals. Moreover, the selection of successful MFs is not only a difficult process but also an uncertain one. Today, in the USA numerous business magazines, private firms, and financial institutions specialize in giving regular rankings and ratings for MFs. Representative examples are the cases of Morningstar, Moody's, Standard & Poor's, etc..

In Greece, there are no such institutions providing assessments of MF performance to Greek investors. The adoption of the evaluation systems of the foreign markets in the Greek case is not feasible, due to the major differences between the Greek market and the other developed markets. Furthermore, according to Sharpe (1998), such measures, like Morningstar's, are only appropriate for single fund investments. In a multifund portfolio such measures are not appropriate to describe the desirability of a fund, since in this case the only relevant issue is the fund's contribution to the total portfolio risk.

The analysis of the nature and definition of risk in portfolio selection and management shows that the risk is multidimensional and is affected by a wide range of financial and stock market data, qualitative criteria and macroeconomic factors. Many of the models used in the past are based on unidimensional approaches that are not appropriate to the multidimensional nature of risk (Colson and Zeney, 1979; Hurson and Zopounidis, 1995).

The empirical literature on the evaluation of the performance of MF portfolios includes the Treynor index (1965), the Sharpe's index (1966), the Jensen's performance index (1968), the Treynor-Mazuy model (1966), the Henriksson-Metron model (1981), the CAPM, and several optimization models. Even though the performance measures proposed in past studies have been widely used in the assessment of portfolio performance, researchers have noted several restrictions in their application, such as the use of a proxy variable of the theoretical market portfolio that can be criticized as inadequate, the evaluation of the performance of an investment manager for long rather than short periods, the acceptance of the assumption of borrowing and lending at the same interest rate, the validity of the Capital Asset Pricing Model, the persistence of the performance of investment managers over time, etc. Moreover, most of the aforementioned studies concerning MF performance evaluation use linear models as primary tools and they, possibly, do not capture the complexity presented in the data.

This paper proposes an alternative approach for the development of MF performance appraisal models. The proposed methodology integrates statistical analysis tech-

niques for the selection of appropriate evaluation criteria with a multicriteria decision aid (MCDA) classification approach, namely the UTADIS method (Doumpos and Zopounidis, 2002). The proposed multicriteria approach enables the consideration of all possible factors related to MF performance in a global non-linear evaluation index. The development of a MF evaluation model through the UTADIS method is performed through non-parametric estimation techniques based on linear programming. This approach provides increased flexibility during model development, since the decision maker can incorporate his/her preferences as constraints in the developed model. Thus, the developed model takes into account the decision makers' (MF manager, investor) subjective judgment and policy. This is the first time such an approach has been employed in MF evaluation. The output of this approach involves the development of a model that supports the evaluation of MF performance and the identification of the funds with positive future prospects. Such a model is of major help to MF managers and investors. MF managers may use such a model to develop and implement appropriate MF management strategies that improve future MF performance. In addition, such a model can be used by MF managers and investors as a tool that supports the selection of MFs that are appropriate investment vehicles in the medium and long term. The results obtained from the application of this methodology to a sample of Greek MFs in the three-year period 1999-2001, during which the Greek market experienced many difficulties, show that it can be an efficient tool for MF managers and investors in evaluating the future prospects of MFs using well-known risk, return and MF management criteria.

The rest of the paper is organized as follows. Section 2 reviews the past research on MF appraisal. Section 3 outlines the main features of the proposed multicriteria methodology. Section 4 is devoted to the application of the proposed methodology, with a description of the data set used and the variable employed, while section 5 discusses the empirical results obtained. Finally, section 6 concludes the paper and summarizes the main findings of this research.

2. Review of past empirical studies

According to prior research, investors pay great attention to the selection of the MFs that best suit their own financial situation (Morey and Morey, 1999). Thus, it is obvious that MF classification helps investors to choose funds according to their decision policy, the risk levels they are willing to take, and their profitability needs. In the international literature, there is a wide variety of studies regarding the development of different models for the evaluation of MF performance, often leading to controversial results and findings.

Several of the existing studies showed that MF managers can outperform the mar-

ket (Henriksson, 1984; Grinblatt and Titman, 1989; Gallo and Swanson, 1996), while other studies reached the opposite result (Elton et al., 1993, 1996). Hendricks et al. (1993), Goetzmann and Ibbotson (1994), and Brown and Goetzmann (1995), in their research works found evidence of performance persistence. On the other hand, Malkiel (1995), Kahn and Rudd (1995), and Carhart (1997), showed slight or no evidence of performance persistence.

Several studies have examined the market-timing and stock selection abilities of MF managers. The academic literature on this area among others concerns the existence or lack of superior performance of investment managers or even the correlation between selectivity and timing abilities (Lehmann and Modest, 1987; Cumby and Clen, 1990; Koh et al., 1993; Sorros, 2001).

Several studies have reported that MF returns are negatively related to fund expense ratios showing that funds that performed well had lower expenses ratios than those that performed poorly (Jensen, 1968; Elton et al., 1993, Malkiel, 1995; Carhart, 1997). Furthermore, another important aspect of MF evaluation literature focuses on the way that MF expenses are related to their management (Gruber, 1996) and the investors' profits (Grossman and Stiglitz, 1980; Daniel et al., 1997).

Most of the previous studies utilize linear models on the evaluation of MF performance as the primary tools and because of that they may not be able to capture complexity in the data. However, recently alternative non-parametric and non-linear approaches have been investigated, with promising results. Some examples include artificial neural network techniques (Indro et al., 1999), data envelopment analysis (Morey and Morey, 1999; Basso and Funari, 2001) and MCDA methodologies (Zopounidis and Pendaraki, 2002).

The increasing importance of the MF classification problem has received great attention from both the practitioners and the researchers. MFs are usually classified based on their stated objectives and their attributes (characteristics, investment style, and risk/return measures). For instance, MFs are classified according to their investment purpose in growth funds, income funds etc. The international literature ignored for many years the topic of MF classification. Apart from the first work by LeClair (1974), three recent works have also focused on this research topic (DiBartolomeo and Witkowski, 1997; Brown and Goetzmann, 1997; Kim et al., 2000).

In the same way, the present study explores the performance of a non-parametric approach based on the UTADIS method, in developing MF performance models through a classification approach.

3. The UTADIS multicriteria decision aid method

The method used to develop a MF performance classification framework in this study is the UTADIS MCDA method (Doumpos and Zopounidis, 2002). The objective of

the UTADIS method is to develop an additive utility model for the classification of a set of alternatives in predefined homogeneous classes with the minimum classification error. In the considered case, the alternatives correspond to the MFs, whereas the classification involves two groups, i.e., the high performance funds and the low performance ones.

The method operates on the basis of a non-parametric ordinal regression-based framework that is similar to the one commonly used in traditional statistical and econometric classification techniques (e.g., discriminant analysis, logit, probit, etc.). Initially, using a training sample the classification model is developed. If the classification accuracy of the model in the training sample is satisfactory, then it can be applied to any other sample for extrapolating purposes. The model development process is briefly outlined below (a detailed description can be found in Zopounidis and Doumpos, 1999).

Let the training sample consist of n MFs (alternatives) $\alpha_1, \alpha_2, \ldots, \alpha_n$ described over a set of m evaluation criteria (variables) $g_1, g_2, \ldots, g_m$. The funds under consideration are classified into q ordered classes $C_1, C_2, \ldots, C_q$ (C_k is preferred to C_{k+1}, k=1, 2, …, q-1). The additive utility model, which is developed through the UTADIS method, has the following form:

$$U(a_j) = \sum_{i=1}^{m} p_i u_i'(g_{ij}) = \sum_{i=1}^{m} u_i(g_{ij}) \in [0, \ 1] \tag{1}$$

where $U(\alpha_j)$ is the global utility of a fund α_j, p_i is the weight (significance) of criterion g_i (the criteria weights sum up to one) and $u_i(g_{ij})$ is the marginal utility of the fund on the evaluation criterion g_i (ranging between 0 and p_i). The global utility provides an overall evaluation of the performance of the fund, whereas the marginal utilities provide the partial evaluations of the fund's performance on each individual evaluation criterion.

Given a fund's overall performance (global utility) its classification in the predefined classes can be performed through the introduction of a set of utility thresholds (cut-off points) $u_1, u_2, \ldots, u_{q-1}$ in the utility scale (each threshold u_k distinguishes the classes C_k and C_{k+1}). The comparison of the global utility of a fund α_j with the utility thresholds leads to the classification of the fund as follows:

$$\left. \begin{array}{ll} U(a_j) \geq u_1 & \Rightarrow a_j \in C_1 \\ u_2 \leq U(a_j) < u_1 & \Rightarrow a_j \in C_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots & \\ U(a_j) < u_{q-1} & \Rightarrow a_j \in C_q \end{array} \right\} \tag{2}$$

Estimations of the global utility model (additive utility function) and utility thresholds are accomplished through the solution of a mathematical programming problem of the following form (Doumpos and Zopounidis, 2002):

$$\text{Minimize} \quad F = \sum_{j=1}^{n}(\sigma_j^+ + \sigma_j^-) \tag{3}$$

Subject to:

$$\sum_{i=1}^{m} u_i(g_{ij}) - u_1 + \sigma_j^+ \geq 0 \qquad \forall a_j \in C_1 \tag{4}$$

$$\left. \begin{array}{l} \displaystyle\sum_{i=1}^{m} u_i(g_{ij}) - u_{k-1} - \sigma_j^- \leq -\delta \\[2em] \displaystyle\sum_{i=1}^{m} u_i(g_{ij}) - u_k + \sigma_j^+ \geq 0 \end{array} \right\} \quad \forall a_j \in C_k \tag{5}$$

$$\sum_{i=1}^{m} u_i(g_{ij}) - u_{k-1} - \sigma_j^- \leq -\delta \qquad \forall a_j \in C_q \tag{6}$$

$$u_i(g_{i*}) = 0$$
$$\sum_{i=1}^{m} u_i(g_i^*) = 1 \tag{7}$$
$$u_i(g_{ij}) \geq u_i(g_{it}), \quad \forall g_{ij} \geq g_{it}$$

$$u_{k-1} - u_k \geq s, \qquad k = 2, 3, ..., q-1 \tag{8}$$

$$u_i(g_{ij}), \ \sigma_j^+, \ \sigma_j^- \geq 0, \ \forall a_j, g_i \tag{9}$$

Constraints (4)-(6) are derived from the classification rules (2) with σ_j^+, σ_j^- ($\sigma_j^+ \cdot \sigma_j^- = 0$) representing the classification errors for fund α_j (δ is a small positive user-defined constant). Constraints (7) define the marginal utilities as increasing functions of the criteria scale with the additional conditions that an ideal(anti-ideal) MF (a fund with ideal(anti-ideal) performance on all criteria) should have a global utility of one(zero).

Finally, constraint (8) defines the minimum difference between two consecutive utility thresholds u_k and u_{k-1} (s is a positive user-defined constant such that $s>\delta$).

The UTADIS method has been applied to several fields of financial management including bankruptcy prediction, credit risk assessment, country risk evaluation, credit cards assessment, portfolio selection and management (Doumpos and Zopounidis, 1998; 2002).

4. Application to mutual funds' performance assessment

4.1 Data Set Description

4.1.1 Sample

The sample used in this study is provided from Alpha Trust S.A. (a leading Greek MF management company) and consists of daily data of all domestic equity MFs over the period 1999-2001 (752 observations daily). The selected time period is chosen solely on the basis of data availability (data prior to 1999 were not available). At the end of 2001, the sample consisted of 72 domestic equity MFs. Nevertheless, for many of these MFs there were not complete data for the whole three-year period of the analysis, since many MFs started their operation in the last two years. Therefore, to have a consistent sample with data for the whole period the analysis is restricted to a smaller sample consisting of 33 MFs which were in operation during the entire period under scrutiny.

Some additional data required for the analysis, such as the return of the market portfolio and the return of a risk free asset, were obtained from the Athens Stock Exchange (ASE) and the Bank of Greece, respectively. In particular, the ASE General Index (ASE-GI) is used as a proxy for the market portfolio, while the three-month Treasury bill considered as the risk free asset.

The starting year 1999 of the period under scrutiny has been characterized as the year of equity MFs. During the whole year, equity MFs experienced high returns following the rapid increase in stock prices on the ASE. On the other hand, the main feature of the subsequent two years (2000-2001) was the considerable decline in stock prices and the negative returns of all MFs. Furthermore, during the last couple of years a considerable degree of homogeneity is observed in the changes in stock prices on the Greek market. Overall, during the period of the analysis the Greek market demonstrated an abnormal pattern of behaviour, similar in some respects to that of other developed markets. This peculiarity of the sample period could make it difficult to derive reliable conclusions on the behaviour of the MFs. Nevertheless, that is not the aim of this study. Instead, its aim is to investigate whether it is possible to develop models that could predict future MF performance. If this objective is achieved using

the selected sample period then it is natural to expect that even better results could be obtained in more "normal" market conditions.

Table 1 presents some descriptive statistics regarding the return and the percentage change of net asset value (NAV) of the MFs in the sample. In particular, in the period (1999-2001) the return of the equity MFs ranged between -29.48% to 64.80%. The mean return of the 33 equity MFs in the three years period is 9.05% with a standard deviation of 19.47%. The standard deviation of the daily returns of the examined MFs ranged between 45.23% to 63.29%. The NAV percentage change ranged between -22.72% to 2840.32%. The mean NAV percentage change of the 33 equity MFs in the three year period is 492.33% with a standard deviation of 721.82%. The variation in these percentages among different MFs is due to the excessive growth of some funds due to inflows by investors into these mutual funds.

Table 1. Descriptive statistics of the sample data

Descriptive Statistics	Return	NAV change (in %)
Mean	9.047	492.338
Standard Deviation	19.776	721.822
Kurtosis	0.717	3.487
Skewness	0.479	2.042
Minimum	-29.487	-22.719
Maximum	64.803	2840.322

As has already been mentioned, the classification of MFs in homogeneous groups is accomplished through the development of a global evaluation index, in the form of an additive utility function (UTADIS method). The realization of this scheme depends heavily on the specification of the groups for the classification of the MFs. The actual decision maker (MF manager, investor) can play an active role in this stage of the analysis through the definition of the MF groups according to his/her investment policy.

MFs are usually classified according to their stated objectives and their attributes (characteristics, investment style, and risk/return measures). In this study after discussion with Greek MF managers, it was decided to employ an approach involving the

classification of the MFs according to their performance in relation to the ASE General Index (ASE-GI), which is used as a proxy of the market. This approach leads to the classification of the MFs as opposed to the ASE-GI used as the reference/benchmark point. The classification of the MFs is determined on the basis of their return R as opposed to the return of the market (R_M) as follows:

− Group 1: High performance funds with $R > R_M(1+k)$, and

− Group 2: Low performance funds with $R < R_M(1+k)$.

Both the MFs return R as well as the market return R_M (return of ASE-GI) are considered for the first semester in 2002 (the most recent data available during the time this study was conducted). Bearing in mind that the data used in the analysis cover the period 1999-2001, it is clear that this classification of the MFs is based on their future returns over a subsequent time period, thus providing a basis for relating their past performance characteristics to their future prospects as investment opportunities.

The parameter k is set equal to 20% to distinguish the MFs with considerably higher returns than the market from the rest of the funds. In particular, the MFs of the first group are the ones with the best perspectives and constitute good investment opportunities compared to the rest of MFs. On the other hand, the MFs of the second group are the ones with lower performance than the ones of the first group. Given the small sample size the mark up point k had to be chosen so that enough instances from each group of MFs were available. The choice of a mark up point higher than the selected value of 20% led to the identification of a very small number of high performance funds, whereas the choice of a lower mark up point led to a very large number of high performance funds. The selected mark up point gave a reasonable distinction of the funds in the sample as high and low performance funds.

4.1.2 Evaluation criteria

The criteria that are used to evaluate MF performance in the three years of the analysis are: (1) the return in the 3-year period, (2) the mean return, (3) the standard deviation of the returns, (4) the coefficient of variation of the returns, (5) the NAV percentage change in the 3-year period, (6) the geometric mean of excess return over benchmark, (7) the value at risk (VaR), (8) the Sharpe index, (9) the Modigliani measure, (10) the information ratio, (11) the beta coefficient β, (12) the Treynor index, (13) the Jensen's alpha α coefficient, (14) the Treynor & Mazuy's α coefficient, (15) the Treynor & Mazuy's γ coefficient, (16) the Henriksson & Merton's α coefficient, (17) the Henriksson & Merton's γ coefficient, and (18) the Treynor & Black appraisal ratio. All these variables refer to different performance and risk measures and are briefly described below.

The arithmetic return on an MF investment includes both income and capital gains or losses. It is calculated as the percentage change in the net asset value of an MF over a given period taking also into account the dividends paid. The arithmetic returns in this study are calculated for the period 1999-2001 used in the analysis.

The standard deviation is the most commonly used measure of variability. For an MF the standard deviation σ is used to measure the variability of daily returns, thus measuring the total risk of the fund. The standard deviation of daily returns (752 daily returns) is transformed in this analysis to refer to the three year time period using the simple formula $\sigma\sqrt{752}$.

An alternative measure of risk refers to the coefficient of variation. The coefficient of variation measures the risk per unit of return achieved, and takes positive or negative values and values higher or lower than unity. The utility of this coefficient refers to the comparison of total risk among MFs.

The return of MFs and other risky investments is often considered in relation to a risk-free asset. In the case of MFs the measure used to investigate this issue is the geometric mean of excess return over the return R_f of a risk free asset. The excess return of a fund is considered as the difference between the fund's return and the risk-free return. The geometric mean of a fund's excess return over a benchmark shows how well the manager of a fund was able to pick stocks. In this analysis the 3-month Treasury bill rate is used as a proxy for R_f.

The beta (β) coefficient is a measure of a fund's risk in relation to the market risk. It represents the systematic risk and the CAPM implies that it is a crucial factor in pricing risky assets. For the calculation of beta coefficient the following regression is used: $R = \alpha + \beta R_M + \varepsilon$, where α is a coefficient measuring the return of a fund when the market is constant, and ε is an error term that represents the impact of non-systematic factors that are independent from the market fluctuations.

Another well-known measure of risk is Value at Risk (VaR). VaR measures the maximum losses that an investor can have in a certain time period for a given confidence level. The calculation of VaR is based on the variance-covariance (VC) approach (Jorion, 2000). Apart from the computational efficiency of the VC approach, another major advantage is that it enables the easy computation of VaR for each time period using widely available daily data. Given the daily standard deviation σ and the expected daily return $\overline{R}$ of an MF, its VaR for a time period of N days for a confidence level α is calculated as $VaR = N\overline{R} + Z_\alpha \sigma \sqrt{N}$, where the term Z_α is obtained from the standard normal cumulative distribution as the probability $P(Z \leq Z_\alpha) = 1 - \alpha$. In this study VaR is calculated for the three year period of the analysis with α=97.5% ($Z_\alpha = -1.96$).

The traditional total performance measures, Sharpe index (1966), and Treynor index (1965) are used to measure the expected return of a fund per unit of risk. These

measures are defined as the following ratios $(R - R_f)/\sigma$ (Sharpe's index) and $(R - R_f)/\beta$ (Treynor's index). The evaluation of MFs with these two indices shows that an MF with higher performance per unit of risk is the best-managed fund, while an MF with lower performance per unit of risk is the worst managed fund.

Modigliani and Modigliani (1997) proposed an alternative measure of risk-adjusted performance that an average investor can easily understand. Given the standard deviation σ_I of an MF's excess return over the index, the Modigliani measure is defined as the ratio $R\sigma_I/\sigma$. The fund with the highest Modigliani measure presents the highest return for any level of risk. According to this approach each portfolio is adjusted to the level of risk of the market benchmark, thereby matching the portfolio's risk to that of the market. The Modigliani measures the performance of this risk-matched portfolio.

Another performance measure that is derived from comparing a fund to its benchmark is the information ratio calculated as the ratio $(R - R_M)/\sigma'$, where σ' is the standard deviation of the MF's excess return over the market portfolio.

The Jensen alpha (1968) measure is the intercept in a regression of the time series of fund excess returns against the time series of excess returns on the benchmark. Both the Treynor index and the Jensen alpha assume that investors are well diversified and, therefore, they are only taking into account systematic risk when evaluating fund performance. The Jensen alpha (α) measure is given by the regression model $R - R_f = \alpha + \beta(R_M - R_f) + \varepsilon$. The coefficient α will be positive if the manager has any forecasting ability and zero if he has no forecasting ability.

The Treynor and Mazuy (1966) model measures both market timing and security selection abilities of fund managers. Treynor and Mazuy add the quadratic term $\gamma(R_M - R_f)^2$ to Jensen's equation to test for market timing skills. The market timing and the security selection performance of MFs are also examined through the Henriksson and Merton model (1981): $R - R_f = \alpha + \beta(R_M - R_f) + \gamma Z_M + \varepsilon$, where $Z_M = \max(0, R_M - R_f)$. In both the Treynor-Mazuy and the Henriksson-Merton models, the parameters α and γ provide estimates on the performance of the MF managers. In particular, α shows the stock selection ability of the manager, while the parameter γ shows his market-timing ability. Positive values for α and γ show that the MF manager has forecasting abilities, negative values indicate forecasting inability, and values close to zero show no ability at all.

Finally, another measure regarding the MF managers' forecasting abilities is the Treynor and Black (1973) appraisal ratio, defined as the ratio α/s, where α is the Jensen alpha coefficient, and s is the standard deviation of the error term in the regression used to obtain the α coefficient. Higher (lower) values of this measure show higher (lower) forecasting ability of the portfolio manager.

4.1.3 Statistical analysis

The incorporation in the analysis of all the above evaluation criteria would result in the development of an unrealistic MF assessment model with limited practical value, whose implementation would require a considerable amount of information and data. To overcome this problem, factor analysis is performed to select the most relevant criteria, which best describe the performance of the MFs in the sample. Of course, it could be possible to override factor analysis if an MF expert was available to determine the most significant indicators.

In this case study, factor analysis is performed using all the available data (18 criteria) for the three-year period. The application of factor analysis led to the development of four factors that account for 88,5% of the total variance in the data. The loadings of the evaluation criteria on the extracted factors are illustrated in Table 2 (only the loadings higher than 0.8, in absolute terms, are presented). The results show that the first factor considers criteria that are related to the return of the MFs, the second factor is mainly characterized by the forecasting ability of the MF managers measured through the Jensen's α coefficient, while the third factor represents the market-timing abilities of the MF managers measured through the Treynor-Mazuy and Henriksson-Merton γ coefficients. Finally, the fourth factor is related to the systematic risk of the MFs (beta coefficient).

From the criteria with loadings higher than 0.8 (in absolute terms) in each factor the most significant one is selected for the further analysis and the development of the MF assessment models. This leads to the selection of four criteria overall, one from each factor. The significance of the criteria to be selected is considered through an ANOVA test. The results of the ANOVA test for all the 18 evaluation criteria are summarized in Table 3. Thus, on the basis of the factor analysis results and the statistical significance of the considered criteria, the following four evaluation criteria are finally selected: (a) the return in the 3-year period, (b) the beta coefficient, (c) the Henriksson & Merton's γ coefficient, and (d) the Treynor & Black appraisal ratio. The selected criteria consider all aspects of MFs performance including: (1) a return criterion measuring the expected outcome of the investment (return over the three year period), (2) a risk criterion measuring the uncertainty about the outcome of the investment (beta coefficient), and (3) two management evaluation criteria measuring the efficiency of the MF managers in terms of their forecasting, market-timing and security selection abilities (Henriksson & Merton's γ coefficient, Treynor & Black appraisal ratio). With regard to the statistical significance of the four selected criteria, it should be noted that the return in the 3-year period and the Treynor & Black appraisal ratio are statistically significant at the 5% level, while the beta coefficient and the Henriksson & Merton's γ coefficient are statistically significant at the 10% level (cf. Table 3).

Table 2. Factor analysis results (factor loadings)

Criteria	1st factor	2nd factor	3d factor	4th factor
Return	0.939			
Mean return	0.941			
Geometric mean of excess return	0.941			
VaR	0.889			
Sharpe index	0.901			
Modigliani measure	0.901			
Information ratio	0.929			
Beta coefficient				0.970
Treynor index	0.906			
Jensen's alpha		0.940		
Treynor & Mazuy γ			0.937	
Herniksson & Merton γ			0.956	

Rotation method: Varimax with Kaiser Normalization (all entries correspond to factor loadings higher than 0.8 in absolute terms)

Table 3. One-Way ANOVA results

	Group means			
Criteria	Group 1	Group 2	F	p
Return	17.060	2.371	5.091	0.031
Mean return	28.787	16.257	4.251	0.048
Standard deviation	54.305	54.537	0.031	0.861
Coefficient of variation	1.738	19.699	1.054	0.312
% Change of NAV	671.915	342.692	1.741	0.197
Geometric mean of excess return	19.273	5.216	4.429	0.044
VaR	77.651	90.635	4.281	0.047
Sharpe index	-1.049	-1.278	3.725	0.063
Modigliani measure	-60.314	-73.485	3.725	0.063
Information ratio	1.197	0.530	3.871	0.058
Beta coefficient	0.877	0.911	3.102	0.088
Treynor index	-64.460	-76.540	2.737	0.108
Jensen's alpha	-0.003	-0.017	2.262	0.143
Treynor & Mazuy α	0.018	-0.003	3.848	0.059
Treynor & Mazuy γ	-0.006	-0.003	3.120	0.087
Herniksson & Merton α	0.034	-0.002	5.598	0.024
Herniksson & Merton γ	-0.066	-0.035	3.537	0.069
Treynor & Black appraisal ratio	-0.006	-0.039	4.438	0.043

Note: Group 1: high performance funds, group 2: low performance funds

5. Presentation of results

In order to investigate the performance of the UTADIS method and compare it with the linear discriminant analysis (LDA), several validation tests are conducted using a leave-one-out cross-validation approach (LOO-CV; Stone, 1974). Given the limited data availability it was not possible to collect a secondary holdout data sample to perform an ex-ante assessment of the proposed methodological framework. Despite this limitation, the use of LOO-CV enables the unbiased assessment of the proposed UTADIS approach in distinguishing between low and high performance funds in the considered period. Further analysis considering additional future data would provide more insight into the true performance of the method, but still, the LOO-CV analysis provides a useful basis for deriving some initial results and conclusions. Finally, it should be noticed that the general Greek market conditions for the period after the first semester of 2002, which is employed for the classification of the MFs in this study, up to the first months of 2003 have not shown any noticeable changes and, consequently, it is expected that the results of the present analysis would not be significantly different if more recent data were employed.

Bearing in mind the above remark, in the LOO-CV analysis conducted in this study 33 replications are performed. In each replication t (t=1, 2, …, 33) the MF α_t is excluded from the sample and a classification model is developed on the basis of another 32 MFs (training sample). This classification model is then tested on MF α_t (validation sample). After all the replications are performed the expected performance of the UTADIS method and LDA is estimated as the average accuracy for the MFs in the validation samples.

On the basis of the above methodology, Table 4 summarizes some statistics on the significance of each criterion in the discrimination between high and low performance MFs according to the models developed through the UTADIS method. The results clearly indicate that the Treynor & Black appraisal ratio is the dominant factor in the discrimination among high and low performance funds, with an average weight of 57.70%. Thus, the classification models developed through the UTADIS method assigned the higher weight in a management evaluation criterion. This finding indicates that MF managers that have been successful in the past in terms of their market forecasting abilities, are expected to be able to follow the same successful management strategy in the future, thereby ensuring higher MF performance. Another criterion that is found significant from the results of the UTADIS method is the beta coefficient with an average weight of 29.87%. Therefore, it is concluded that the driving factors of MF performance during this difficult period for the Greek market were the funds' risk in relation to the market risk and the forecasting ability of managers; both have a very important role in the evaluation of the MFs performance. It is important to notice,

despite the peculiarities of the sample period with regard to the behaviour of the Greek market, that the above findings are consistent with the results of previous studies, such as those of Elton et al. (1996), Ferson and Schadt (1996), Carhart (1997), Daniel, et al. (1997).

Table 4. Statistics on the weights of the evaluation criteria according to the UTADIS method (leave-one-out cross validation results)

Criteria	Average weight	St. error
Annual return	12.37%	5.03%
Beta coefficient	29.87%	7.53%
Henriksson & Merton's γ coefficient	0.06%	0.01%
Treynor & Black Appraisal ratio	57.70%	8.00%

Table 5 summarizes the average LOO-CV classification results for the UTADIS method. For comparative purposes the results of LDA are also reported. The entries "High performance-High performance" and "Low performance-Low performance" represent average classification accuracy for each of the two groups, while all the other entries of the table correspond to average classification errors. The final column of the table presents the average overall accuracy of each method and the corresponding standard error (in parenthesis).

The results obtained indicate that UTADIS outperforms LDA both in the training and the validation samples. In particular, in the training sample, the overall classification accuracy of the UTADIS method is 80.52% while accuracy of LDA is 77.98%. Of course, higher model fit in the training sample does not ensure higher generalizing ability, which is the ultimate objective in decision models, developed through regression-based techniques. In that respect, the results on the validation tests are of particular interest towards the evaluation of the predictability of UTADIS. The comparison of the methods according to the validation sample results indicates that in terms of the overall classification accuracy, UTADIS performs better than LDA. In particular, in the validation sample, the overall classification accuracy of the UTADIS method is 78.33% while for LDA it is 69.44%. Moreover, the average classification errors in the UTADIS method are lower than the ones in the LDA method for both the "low performance" group and the "high performance" group. Misclassifications involving

Table 5. Summary of average classification results (leave-one-out cross validation)

		High Performance	Low Performance	Overall accuracy
		Training		
UTADIS	High Performance	75.50%	24.50%	80.52%
	Low Performance	14.46%	85.54%	(0.45)
LDA	High Performance	72.48%	27.52%	77.98%
	Low Performance	16.52%	83.48%	(0.36)
		Validation		
UTADIS	High Performance	73.33%	26.67%	78.33%
	Low Performance	16.67%	83.33%	(7.12)
LDA	High Performance	66.67%	33.33%	69.44%
	Low Performance	27.78%	72.22%	(8.00)

Note: Parentheses indicate the standard error of overall accuracy

funds of the "low performance" group may result in capital losses for the investor, whereas misclassifications involving funds of the "high performance" group may lead to opportunity costs. In both cases, the UTADIS models perform better than LDA.

Although these results involve a specific time period characterized by several peculiarities regarding the operation of the Greek market, they illustrate the potential that the proposed methodological framework provides in assessing MFs performance. In particular, the results show that it is possible to develop models to identify the future performance of MFs in terms of their past risk/return characteristics as well as on the basis of their management efficiency. This is an important finding for both the MF managers and the investors. The former may use such models as a support tool in the design and implementation of the MF management strategy in order to improve the future performance of their MFs. On the other hand, investors may use such performance appraisal models to select appropriate MFs for medium-long term

investments. Of course, such models should not be static. Instead, they should be periodically updated when market conditions change.

6. Concluding remarks and future perspectives

The performance of MFs has become an increasingly important issue among portfolio managers and investors. The aim of this study is to propose a methodological framework for predicting future MF performance based upon characteristice MF relative returns and risks.

In order to achieve this goal a sample of 33 Greek domestic equity funds was employed during a period which was quite difficult for the Greek market. After a thorough statistical analysis of the sample (factor analysis and ANOVA test) a limited set of MF performance measures was identified and employed to develop appropriate MF performance models. For the development of these models the UTADIS classification method was employed. The results obtained from the analysis show that this methodological framework enables MF managers and investors to develop efficient models for assessing MF performance on the basis of available data. In this study the assessment of MF performance was realized in terms of the future MF returns compared to ASE General Index used as a benchmark in order to distinguish between MF with high and low performance. Of course, other classification and performance evaluation schemes are also applicable.

The development of models such as the ones considered in this study is of major support to MF managers and investors. The main use of these models is to support the selection of appropriate MFs for investing over a medium-long term period. Furthermore, such models can be employed by MF managers to monitor the performance of their MFs and to design appropriate strategies that ensure high future performance.

A further investigation of the proposed methodology to consider an extensive data period would provide a useful insight into the efficiency of the methodology itself, as well as into the characteristics of the Greek MF market. Comparisons with other markets are also worth investigating to analyze the similarities and differences between the Greek case and international experience. Finally, further examination of the proposed methodological framework in other performance assessment problems and comparative studies among other methods to identify their relative strengths and weakness is also an interesting research topic.

References

Basso, A. and Funary, S. 2001, "A data envelopment analysis approach to measure the mutual fund performance", *European Journal of Operational Research*, 135, 477-492.

Brown, S.J., Goetzmann, W.N., 1995, "Performance persistence" *The Journal of Finance*, 50, 679-698.

Carhart, M., 1997, "On persistence in mutual fund performance", *The Journal of Finance*, LII, 57-82

Colson, G. and Zeleny, M., 1979, Uncertain prospects ranking and portfolio analysis under the condition of partial information, in Verlag Anton Hain (eds.), *Mathematical Systems in Economics*, 44, Maisenheim.

Cumby R.E. and Glen, J.D., 1990, "Evaluating the performance of international mutual funds", *The Journal of Finance*, XLV, 497-521.

Daniel, K., Grinblatt, M., Titman, S. and Wermers, R., 1997, "Measuring mutual fund performance with characteristic-based benchmarks", *Journal of Finance*, 52, 1035-1058.

DiBartolomeo, D. and Witwowski, E., 1997, "Mutual fund misclassification: Evidence based on style analysis", *Financial Analysts Journal*, 53, 32-43

Doumpos, M. and Zopounidis, C., 1998, "The use of the preference disaggregation analysis in the assessment of financial risks", *Fuzzy Economic Review*, 3, 39-57.

Doumpos, M. and Zopounidis, C., 2002, *Multicriteria Decision Aid Classification Methods*, Dordrecht: Kluwer Academic Publishers,.

Elton, E.J., Gruber, M.J., Das, S., Hlavka, M., 1993. "Efficiency with costly information: A reinterpretation of evidence from managed portfolios", *The Review of Financial Studies*, 6, 1-22.

Elton, E.J., Gruber, M.J. and Blake, C.R., 1996, "The persistence of risk-adjusted mutual fund performance", *Journal of Business*, 69, 133-157.

Ferson, W. and Schadt, R., 1996, "Measuring fund strategy and performance in changing economic conditions", *Journal of Finance*, 51, 425-461.

Gallo, J.G., Swanson, P.E., 1996, "Comparative measures of performance for U.S.-based international equity mutual funds", *Journal of Banking and Finance*, 20, 1635-1650.

Goetzmann, W.N., Ibbotson, R., 1994, "Do winners repeat? Patterns in mutual fund return behavior", *The Journal of Portfolio Management*, 20, 9-18.

Grinblatt, M. and Titman, S., 1989, "Portfolio performance evaluation: Old issues and new insights", *Review of Financial Studies*, 2, 393-421.

Grinblatt, M. and Titman, S., 1994, "A study of monthly fund returns and performance evaluation techniques", *Journal of Financial and Quantitative Analysis*, 29, 419-443.

Grossman, S.J. and Stiglitz, J., 1980, "On the impossibility of informationally efficient markets", *The American Economic Review*, 70, 398-408.

Gruber, M. J.,1996, "Another puzzle: The growth in actively managed mutual funds", *Journal of Finance*, 51, 783-810.

Hendricks, D., Patel, J., Zeckhauser, R., 1993, "Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988", *The Journal of Finance*, 48, 93-130.

Henriksson, R., 1984, "Market timing and mutual fund performance: An empirical investigation", *Journal of Business*, 57, 73-96.

Henriksson, R. and Merton, R., 1981, "On market timing and investment performance", *Journal of Business*, 54, 513-534.

Hurson, Ch. and Zopounidis, C., 1995, "On the use of Multicriteria decision aid methods for portfolio selection", *Journal of Euro-Asian Management*, 1/2, 69-94.

Hurson, Ch. and Zopounidis, C., 1997, Gestion de portfeuille et Analyse Multicritère, *Economica*, Paris.

Indro, D.C., Jiang, C.X., Patuwo, B.E. and Zhang, G.P., 1999, "Predicting mutual fund performance using artificial neural networks", *Omega*, 27, 373-380.

Jensen, C.M., 1968, "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, 23, 389-416.

Jorion, P., 2000, *Value at Risk: The New Benchmark for Managing Financial Risk* (2nd edition), New York: McGraw-Hill

Kahn, R.N., Rudd, A., 1995, "Does historical performance predict future performance", *Financial Analysts Journal*, 43-53.

Kim, M., Shukl, R. and Tomas, M. 2000, "Mutual fund objective misclassification", *Journal of Economics and Business*, 52, 309-323.

Koh, F., Phoon, K., Tan, C., 1993, "Market timing abilities of fund managers: Parametric and non-parametric tests", *Journal of Business Finance & Accounting*, 20, 155-166.

LeClair, R. T. 1974, "Discriminant analysis and the classification of mutual funds", *Journal of Economics and Business*, 26, 220-224.

Lehmann, B.N. and Modest, D.M., 1987, "Mutual Fund Performance Evaluation: A Comparison of Benchmarks and Benchmark Comparisons", *The Journal of Finance*, XLII, 233-265.

Malkiel, B.G., 1995, "Returns from investing in equity mutual funds 1971-1991", *Journal of Finance*, 50, 549-572.

Modigliani, F. and Modigliani L., 1997, "Risk-adjusted performance", *Journal of Portfolio Management*", 23, 45-54.

Morey M.R. and Morey, R.C., 1999, "Mutual fund performance appraisals: A multi-horizon perspective with endogenous benchmarking", *Omega, Int. J. Mgmt Sci.*, 27, 241-258.

Sharpe, W.F., 1966, "Mutual Fund Performance", *Journal of Business*, 39, 119-138.

Sharpe, W.F., 1998, "Morningstar's risk adjusted ratings", *Financial Analysts Journal*, (July/August), 21-23.

Sorros, J. 2001, "Equity mutual fund managers performance in Greece", *Managerial Finance*, 27, 68-74.

Stone, M., 1974, "Cross-validation choice and assessment of statistical predictions", *Journal of the Royal Statistical Society*, 36, 111-147.

Treynor, J.L., 1965, "How to rate management of investment funds", *Harvard Business Review*, 43, 63-75.

Treynor, J.L. and Mazuy, K.K., 1966, "Can mutual funds outguess the market", *Harvard Business Review*, (July/August), 131-136.

Treynor, J. and Black, F., 1973, "How to use security analysis to improve portfolio selection", *Journal of Business*, 46, 66-68.

Zopounidis, C. and Doumpos, M., 1999, " A multicriteria decision aid methodology for sorting decision problems: The case of financial distress", *Computational Economics*, 14, 197-218.

Zopounidis, C. and Pendaraki, K., 2002, "An integrated approach on the evaluation of equity mutual funds' performance", *European Journal of Business and Economic Management*, (in press).