Domestic Open-End Equity Mutual Fund Performance Evaluation Using Extended TOPSIS Method with Different Distance Approaches

Ching-Hui Chang  
Department of Applied Statistics and Information Science,  
Ming Chuan University, Taoyuan, Taiwan, R.O.C.

Jyh-Jiuan Lin*  
Department of Statistics,  
Tamkang University, Tamsui, Taipei, Taiwan, R.O.C.

Miao-Chen Chiang  
Department of Statistics,  
Tamkang University, Tamsui, Taipei, Taiwan, R.O.C.

Wen-Rong Ho  
Department of Banking and Finance,  
Kainan University, Taoyuan, Taiwan, R.O.C.

ABSTRACT

The most commonly used measures for evaluating the competing mutual funds are “Treynor Ratio”, “Sharpe Ratio” and “Jensen’s Alpha”. One also uses another measure called the “Information Ratio”. However, it is not clear which measure is the most robust. The purpose of our study is to evaluate the performance of mutual funds under the broad framework of Multi-Attribute Decision Analysis approach where each criterion can be taken into consideration in making a final ranking of the mutual funds. In this paper we adopt the concepts of “Ideal” and “Anti-Ideal” solutions as suggested by Hwang and Yoon (1981), and study the extended TOPSIS (Technique for Order Preference by Similarity to the Ideal Solution) method using two different “distance” ideas, namely – “Minkowski’s $L_1$ metric” and “Mahalanobis” distances. The broad framework with the two aforementioned distances is then applied to evaluate the performance of 82 Taiwanese mutual funds for consecutive 34 months. The empirical results show that using TOPSIS methods incorporated Minkowski’s distance measure to evaluate the mutual funds’ performance perform well.

Keywords: multiple criteria analysis, mutual fund, performance evaluation, TOPSIS

1. INTRODUCTION

In the current financial scenario investment through mutual funds has emerged as one of the most popular financial strategies. Mutual funds tend to attract investors due to the following advantages: (i) affordability, (ii) diversification, and (iii) professional management. Thus, a great deal of efforts has been made and many criteria have been developed on evaluating the mutual funds’ performance. However, it is not clear which criterion is the most robust when they don’t agree in ranking the funds. The purpose of this study is to evaluate the performance of mutual funds under the broad framework of Multi-Attribute Decision Analysis (MADA) approach where all criteria can be taken into consideration in making a final ranking of the mutual funds. The rest of the paper is organized as follows.

We review the four popular criteria measures (or “attributes”) used to evaluate the performance of the mutual funds briefly in section 1.1 and the MADA approach in section 1.2. The research method, which is essentially extended TOPSIS, used in this paper is introduced in section 2. Finally an application with one real life data is discussed in section 3.

1.1 Performance Evaluation Measures of Mutual Funds

The three most commonly used measures of risk adjusted performance are “Treynor Ratio” ($TR$, 1965), “Sharpe Ratio” ($SR$, 1966) and “Jensen’s Alpha” ($JA$, 1968). Besides, another measure of risk adjusted relative return is the “Information Ratio” ($IR$).

Some of these ratios have been applied to evaluate the performance by many researchers including Pendaraki et al. (2005) and Lin et al. (2007). Pendaraki et al. (2005) employed one commonly used classification techniques, the UTADIS (UTilités Additives DIscriminantes) MCDA method, to develop mutual fund’s performance models supporting the selection of a small set of mutual funds which will compose the final portfolios. Then a
goal programming model is adopted to determine the proportion of the selected mutual funds in the final portfolios. Lin et al. (2007) who studied 82 domestic open-end equity mutual funds in Taiwan as same as in this paper. However, it focuses on the choice of the objective weights only under the Euclidean distance. This paper extends it taking not only the objective weights but also the distance measurement into consideration under the framework of MADA approach which is introduced as follows.

1.2 Multi-Attribute Decision Analysis

A decision making problem is always complicated especially if there are more than one criterion under consideration. One can imagine the task of ranking several job applicants when they can be evaluated in terms of multiple criteria or attributes like - technical skill, interpersonal communication skill, leadership skill, etc. One method to handle such a decision making problem is the Data Envelopment Analysis (DEA) which has attracted the attention of many researchers including Charnes et al. (1978) who have successfully used DEA to evaluate banks’ efficiency. However, apart from some technical difficulties, DEA is more concerned about identifying the alternatives which are not “efficient” in some sense.

Alternatively, Multi-Criteria Decision Analysis (MCDA) is another approach to tackle the above-mentioned decision making problems. Basically, Multi-Criteria Decision Problems, MCDP, has two branches - Multi-Attribute Decision Problems (MADP) and Multi-Objective Decision Problems (MODP). Our interest lies in MADP which can evaluate several competing decision rules under multiple (and often conflicting) criteria. The Technique for Order Preference by Similarity to the Ideal Solution Method (TOPSIS) is one of the many MADA methods. For a good review on this see Wang (2005). It has been pointed out by Deng et al. (2000) that (i) the concept of TOPSIS is rational, (ii) the computation involved is easy, and (iii) it allows objective weights to be incorporated into the comparison process. TOPSIS method adopts the concepts of “ideal” and “anti-ideal” solutions as suggested by Hwang and Yoon (1981) and computes the weighted distances to measure the relative distances away from the ideal and anti-ideal solutions for each alternative (i.e., decision rule). The idea of TOPSIS is shown in Figure 1. Not only the best alternative should be as close as possible to the ideal solution but also it should be as far as possible away from the anti-ideal solution.

Basically, TOPSIS is composed of two components: weights and distances. The most commonly used weights are - mean weight, standard deviation weight and the well-known Shannon and Weaver’s (1947) entropy weight. In addition, Sinha (2003) proposed coefficient-variation weight. Among the most commonly used distances one can use “Euclidean distance” and Sinha’s (2003) “City block distance”. Lin et al. (2005) applied the aforementioned weight methods under Euclidean and City block distances to the academic performance problems. Wu (2004) applied the TOPSIS method to study the Taiwanese mutual fund selection. However, the weight method used in his research is through factor analysis and the distance measurement is Minkowski’s $L_2$ metric with $\lambda = 2$ (Euclidean). Lin et al. (2007) also studies the performance of uni-criterion and multi-criteria indices using the extended TOPSIS method under the Euclidean distance. Therefore, the question regarding the different distance measurements effect to the performance evaluation remains open.

The purpose of this paper is not only to compare the competing mutual funds, but also to see how the TOPSIS method affects the performance evaluation of the mutual funds by using different “distance” ideas under a specific weight method.

2. THE EXTENDED TOPSIS METHOD WITH DIFFERENT DISTANCE APPROACHES

2.1 Data Preparation

Suppose a decision matrix $X$ with $n$ alternatives and $m$ criteria is expressed as

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}, \quad (1)$$

where $x_{ij}$, the $(i, j)^{th}$ element of the matrix, denotes the score of the $i^{th}$ alternative (candidate) with respect to the $j^{th}$ criterion. Since the rate of the return could take negative values, additional caution with location shift has been taken in the data process step to suit the methodology later. Therefore, all elements of the matrix $X$ as in (1) will always take the nonnegative values. Now the vector $x_j = (x_{1j}, x_{2j}, \ldots, x_{nj})$, for $j = 1, 2, \ldots, m$ will be used as the scores for all $n$ alternatives with respect to criterion $j$ for the rest of the paper. Since the larger the original observation, the larger the transformed new score and the better the performance of the mutual fund, we define the Ideal Row (IDR) and the Anti-ideal Row (ANIDR) as in (2) and (3).

$$IDR = (\max x_{j1}, \max x_{j2}, \cdots, \max x_{jm})$$

$$= (a_1, a_2, \cdots, a_m) = a’ \quad (2)$$

$$ANIDR = (\min x_{j1}, \min x_{j2}, \cdots, \min x_{jm})$$

$$= (\bar{a_1}, \bar{a_2}, \cdots, \bar{a_m}) = \bar{a}’ \quad (3)$$

Figure 1. The idea of TOPSIS method

\[ \text{Figure 1. The idea of TOPSIS method} \]
\[ ANIDR = (\min_i x_{i1}, \min_i x_{i2}, \ldots, \min_i x_{in}) \]
\[ = (b_1, b_2, \ldots, b_m) = \mathbf{b'} \]

2.2 Objective Weights

Four objective weights (say, \( w_i \)) are used in the paper. They are introduced briefly in the following.

(a) Mean Weight (MW) Method: Mean weight method assigns equal weight to each criterion. It reflects a neutral attitude of the decision maker and guarantees the objectivity of performance evaluation process.

(b) Entropy Weight (EW) Method: Shannon and Weavers' (1947) entropy is a measure of uncertainty in information and the entropy concept is commonly used in MCDP. Its value reflects the relative importance of its corresponding criterion in terms of the amount of the information it contains and it indicates the inherent contrast intensity of the corresponding criteria.

(c) Coefficient-Variation Weight (CV) Method: Sinha (2003) has suggested that the sample coefficient of variation can be used as the objective weight in multi-criteria decision problems. However, the sum of the total coefficient of variation is not 1. To make the interpretation of the weights easier, the sum of \( CV \) weights is normalized to unity through 
\[ \tilde{c}_j = \frac{c_j}{\sum_{j=1}^{m} c_j} \]
where \( c_j \) is the sample coefficient of variation under criterion \( j \).

(d) CRITIC (CR) Method: Diakoulaki et al. (1995) proposed the CRITIC method, which aims at the determination of the objective weights that incorporate both contrast intensity and conflict.

2.3 Weighted Distances

Two distance ideas used in this paper are introduced briefly in the following.

(a) Minkowski's \( L_\lambda \) Metric Weighted Distance: The Minkowski's \( L_\lambda \) metric weighted distances from the Ideal Row and the Anti-ideal Row are defined as in (4) and (5) respectively.

\[
WD_i(IDR) = \left[ \sum_{j=1}^{n} \left( \frac{1}{w_j} \left( x_j - a_j \right)^\lambda \right)^{\frac{1}{\lambda}} \right]^{\frac{1}{\lambda}}; \quad (4)
\]
\[
WD_i(ANIDR) = \left[ \sum_{j=1}^{n} \left( \frac{1}{w_j} \left( x_j - b_j \right)^\lambda \right)^{\frac{1}{\lambda}} \right]^{\frac{1}{\lambda}}; \quad (5)
\]
where \( i = 1, 2, \ldots, n \). Note that Minkowski distance is equal to the City block (or Manhattan) distance when \( \lambda = 1 \), the Euclidean distance when \( \lambda = 2 \), and the Tchebycheff distance when \( \lambda = \infty \). Yu (1973) suggests \( \lambda = 1 \) under “majority” rule and \( \lambda = \infty \) under the “Minimum Individual Regret of Opponent” rule. It was also pointed out by Lai et al. (1994) that the credibility of the Minkowski’s \( L_\lambda \) distance function decreases as parameter \( \lambda \) increases. Therefore, the cases studied in this paper only focus on those two most commonly used distance measurements with \( \lambda = 1 \) (City block), \( \lambda = 2 \) (Euclidean) and Minkowski’s \( L_\lambda \) distance with small \( \lambda = 1.5 \) (say); those cases are referred as \( CB \), \( EU \) and \( MI \), respectively. Besides, the concept among these three measurements is demonstrated in Figure 2.

(b) Mahalanobis Weighted Distance (MA): Mahalanobis distance measure offers an option to be able to take the variation as well as the correlations between the variables (attributes or criteria here) into consideration. For example, as shown in Figure 3, the Euclidean distance between cases A and B is as the same as that between cases C and D when the dimension is equal to 2. However, if the relationship between two variables is positively correlated, then the Mahalanobis distance between A and B is greater than that between C and D. On the contrary, if the relationship between two variables is negatively correlated, then the Mahalanobis distance between A and B is less than that between C and D. The variance-covariance matrix of those criteria is equivalent to the role of objective weights, and is called Pooled Within-Group Covariance Matrix. It can be seen that the Mahalanobis distance is equal to the Euclidean distance when the standardized criteria are independent. It provides the strong motive for us to study the affect of correlation on the performance evaluation of the mutual funds by using TOPSIS method through the Mahalanobis distance measurement in this paper. More details regarding the Mahalanobis distance can be seen in Chow (2002).

The Mahalanobis weighted distances from the Ideal Row and the Anti-ideal Row are defined as in (6) and (7) respectively.

\[
WD_\lambda(i, IDR) = [(x_i - \mathbf{a'})\Sigma^{-1}(x_i - \mathbf{a})]^{1/2}; \quad (6)
\]
\[
WD_\lambda(i, ANIDR) = [(x_i - \mathbf{b'})\Sigma^{-1}(x_i - \mathbf{b})]^{1/2}, \quad (7)
\]
where \( i = 1, 2, \ldots, n \) and \( \Sigma \) is the variance-covariance matrix of \( X \).

2.4 Evaluation Indices

The overall evaluation index for alternative \( i \) under each distance is defined as

\[
I_{WD_\lambda, i} = \frac{WD_\lambda(i, ANIDR)}{WD_\lambda(i, IDR) + WD_\lambda(i, ANIDR)}, \quad (8)
\]
where \( k = 1, 2 \) and \( i = 1, 2, \ldots, n \). The larger the index
value, the better the performance of the alternative.

The broad framework with the four aforementioned distances is then applied to evaluate and rank the performance of 82 Taiwanese mutual funds as discussed in section 3.

3. EMPIRICAL STUDY

This article studies 82 domestic market-oriented mutual funds which are classified as open-end equity mutual funds by the Securities Investment Trust & Consulting Association of the R.O.C. over the period form September, 2002 to June, 2005. Four yearly based measures (TR, SR, JA and IR) and the actual yearly rate of return for each mutual fund, AR, are collected from the website (http://140.112.111.12) of Mutual Fund Performance Evaluation, Taiwan (which is co-chaired by Professor Tsun-Siou Lee and Professor Shean-Bii Chiu, Department of Finance, National Taiwan University). The rankings of AR are used to be the benchmark (since AR is the “true” performance of the mutual fund) to assess the TOPSIS method studied in this paper.

3.1 Evaluation of Indices

So far we have discussed thirteen extended TOPSIS indices based on different distance choices (EU, CR and MI) under different objective weights (MW, EW, CV and CR) and MA itself to evaluate the performance of the mutual funds. To be able to see which index is the most robust, we will investigate those 13 indices in two perspectives: (i) Spearman’s Rank Correlation Coefficient, rs, a popular nonparametric measure of association based on ranks, and (ii) Proportion of the rankings using an evaluating index matched with the rankings of actual return (AR) based on the true yearly performances (henceforth called, PRMAR). It is a number between 0 and 1 that indicates the consistency of the evaluation indices with the true performance (AR). On the other hand, one may also look at the ability of identifying the best (or better) alternative(s) using an index for those periods. To do so, one could use the proportion of the rankings using an evaluating index matched with the ranking(s) of the top ones, PRMAR(K) (K could be any number interested which indicates the top rankings). Therefore, the higher the PRMAR(K), the better ability of identifying the Top 1 to Top K mutual funds, and hence the better the evaluating index.

3.2 Numerical Results

(a) Objective Weights Assigned to Each Criterion:

The averages and the standard deviations of all weights of each criterion among four objective weights over 34 periods are shown in Table 1. MW method assigns four criteria equal weight (0.2500, 0.2500, 0.2500 and 0.2500) as it should be. For EW method, averagely, it assigns objective weights 0.2384, 0.2156, 0.2855 and 0.2605 to TR, SR, JA and IR respectively. For CV method, averagely, it assigns objective weights 0.2503, 0.2315, 0.2667 and 0.2515 to TR, SR, JA and IR respectively. Not only that, all the standard deviations of the objective weights are not greater than 0.1206. That means we may conclude that both methods are putting about the same weights on each criterion averagely. However, the CR method tends to give IR larger objective weight. Averagely, it assigns 0.2300, 0.1640, 0.2249 and 0.3811 to TR, SR, JA and IR. Later, one will see that CR method under EU has maximum r_s counts and it may be due to it put more weights on IR. IR is a more informative index among the four uni-criterion indices in terms of Spearman’s rank correlation coefficient r_s as mentioned in Lin et al. (2007).

(b) Spearman’s Rank Correlation Coefficients, r_s:

The averages and the standard deviations of the Spearman’s rank correlation coefficients, r_s, between the overall index under each evaluating method and AR over 34 periods are provided in Table 2. Basically, one can see that most of the average r_s for each case are all above up to 0.96 except MA case. That means those approaches adopted all perform well in terms of r_s. Therefore, it is plausible to use one of those indices except MA case to evaluate mutual funds. Note that the off-diagonal elements of the variance-covariance matrix Σ used in the MA case are just the regular covariances and it plays the role of weight. On the other hand, all the weight methods adopted are especially and dedicatedly designed to reflect the information carried by the criteria. Therefore, it could...
be one of the reasons why MA case does not perform well as other indices do.

For convenience of interpreting the numerical results, we use the new notation as defined in Table 3. The “P” stands for index and the subscripts of it stands for the corresponding weight method and distance measure. Mahalanobis is a weight method as well as a distance measure, so \( I_{M}\) is used to refer to the Mahalanobis case. One could see from Table 3 that \( I_{MC}, I_{EC}, I_{CE}\) and \( I_{EC}\) perform best under \( MW, EW, CV\) and \( CR\) weight methods respectively. Besides, \( I_{MC}\) has overall maximum \( r\) counts (8) among 13 indices during 34 periods. Thus, without any weight methods preference, \( I_{MC}\) is recommended if someone is looking for an index which gives high correlation with the rankings of actual return, \( AR\).

(c) Proportion of Rankings Matched with Rankings of \( AR\): For proportion of rankings matched with rankings of \( AR\), the proportion of rankings of the overall, Top 1, Top 1-3, Top 1-5, and Top 1-10 matched with \( AR\) are listed in Table 3. One can obtain the following results.

(i) For overall, the \( MI\) distance measure performs best under the \( MW \) and \( EW\) weight methods; so does the \( EU\) distance measure under the \( CV\) and \( CR\) weight methods on the other hand.

(ii) For Top 1, the \( MI\) distance measure performs best under \( MW, EW\) weight methods; the \( EU\) and \( MI\) distance measures are both tied under the \( EW\) weight method; the \( CB\) distance measure performs best under the \( CR\) weight method.

(iii) For Top 1-3, the \( CB\) distance measure performs best under the \( MW\) and \( CR\) weight methods; the \( EU, CB\) and \( MI\) distance measures are all tied under the \( EW\) weight method; the \( MI\) distance measure performs better under the \( CV\) weight method.

(iv) For Top 1-5, the \( CB\) distance measure performs best under the \( MW, EW\) and \( CR\) weight methods; the \( CB\) and \( MI\) distance measures are tied under the \( CV\) weight method.

(v) For Top 1-10, the \( MI\) distance measure performs best under the \( MW, EW\) and \( CV\) weight methods; the \( CB\) distance measure performs best under the \( CR\) weight method.

<table>
<thead>
<tr>
<th>Weight</th>
<th>( MW) Mean St.D.</th>
<th>( EW) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td>TR</td>
<td>SR</td>
</tr>
<tr>
<td>Mean</td>
<td>0.2500</td>
<td>0.2500</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight</th>
<th>( CV) Mean St.D.</th>
<th>( CR) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td>TR</td>
<td>SR</td>
</tr>
<tr>
<td>Mean</td>
<td>0.2503</td>
<td>0.2315</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0498</td>
<td>0.0372</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight</th>
<th>( EU) Mean St.D.</th>
<th>( CB) Mean St.D.</th>
<th>( MI) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>( I_{ME}) 4 (14)</td>
<td>( I_{MC}) 1 (16)</td>
<td>( I_{MM}) 1 (5)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.9682</td>
<td>0.9692</td>
<td>0.9684</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0305</td>
<td>0.0277</td>
<td>0.0297</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight</th>
<th>( CV) Mean St.D.</th>
<th>( CR) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td>( I_{ME}) 3 (16)</td>
<td>( I_{MC}) 3 (15)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.9667</td>
<td>0.9671</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0332</td>
<td>0.0317</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight</th>
<th>( EU) Mean St.D.</th>
<th>( CB) Mean St.D.</th>
<th>( MI) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance</td>
<td>( I_{ME}) 3 (17)</td>
<td>( I_{MC}) 1 (13)</td>
<td>( I_{MM}) 0 (5)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.5294</td>
<td>0.3294</td>
<td>0.3829</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0588</td>
<td>0.0392</td>
<td>0.0765</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Weight</th>
<th>( CR) Mean St.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criteria</td>
<td>( I_{ME}) 5 (18)</td>
</tr>
<tr>
<td>Mean</td>
<td>0.4118</td>
</tr>
<tr>
<td>St.D.</td>
<td>0.0470</td>
</tr>
</tbody>
</table>

Table 1. Weights of Each Criterion under Four Objective Weight Methods

Table 2. Spearman’s Rank Correlation Coefficients \( r_s\)

Table 3. Methods Evaluation for Fixed Weight Methods for 34 Periods

Maximum * tied † the occurrence of maximum counts for both indices when there is a tie.
Overall speaking, $I_{EM}$ has the highest PRMAR(82) (0.1614). It means $I_{EM}$ performs best to describe the performance of overall mutual funds in terms of PRMAR(82). In addition, $I_{MIN}$, $I_{ES}$, $I_{EM}$ and $I_{CM}$ are all tied and good at “picking” the best (Top 1) mutual funds in terms of PRMAR(1). Besides, $I_{MC}$ and $I_{CM}$ are both tied and outperform the rest of 11 indices if someone is interested in selecting Top 1-3 mutual funds. Similarly one could easily find a suitable index to evaluate the mutual funds in terms of PRMAR($K$) according their own interest. Not surprisingly, it has seen that the PRMAR($K$) is decreasing in $K$ since it is more challenging to “predict” the Top 1-10 than the Top 1.

4. CONCLUSION

The inferences drawn here are concluded as follows:
(a) the index using CRITIC weight method under the Euclidean distance measure is recommended if the correlation between the rankings is concerned.
(b) all the multi-criteria indices (the TOPSIS method using four different distance measures, namely – “Euclidean distance”, “City block distance”, “Minkowski distance” and “Mahalanobis distance”) except Mahalanobis case are pretty much consistent in their evaluations, and they produce high Spearman’s rank correlation with the rankings of the rate of actual return. Therefore, using extended TOPSIS incorporated the Euclidean, City-block or Minkowski ($L_\lambda$ with $\lambda = 1.5$) distance to evaluate the performance of the mutual funds all perform well.

ACKNOWLEDGEMENT

The second author’s research has been supported by a research grant from the National Science Council (NSC 95-2118-M-032-007).

REFERENCES