OR Spectrum 28:375–398 (2006)
DOI 10.1007/s00291-005-0032-1

REGULAR ARTICLE

Zhiping Chen · Ruiyue Lin

Mutual fund performance evaluation using data envelopment analysis with new risk measures

Published online: 25 January 2006

Abstract The data envelopment analysis (DEA) technique has been found very useful for evaluating the mutual fund performance. This applied study extends previous results in two ways: to properly reflect the pervasive skewness and leptokurtosis in return distributions of actively managed funds, new risk measures value-at-risk (VaR) and conditional value-at-risk (CVaR) are introduced into inputs of the existing DEA models; to fairly evaluate the relative performance of the same fund during different time periods, we creatively treat the same fund during different periods as different decision making units. Except for confirming current empirical conclusions, detailed empirical analyses using data of the Chinese mutual fund market show that, VaR and CVaR, especially their combinations with traditional risk measures, are very helpful for comprehensively describing return distribution properties and fund characteristics such as the asset allocation structure, which, in turn, can better evaluate the overall performance of mutual funds. Treating the same fund during different time periods as different funds can not only show the specific performance variation, but reveal the reasons for that variation.

Keywords Data envelopment analysis · Performance evaluation · Mutual funds · Risk measures · Return

This research was partially supported by the National Natural Science Foundation of China (10571141), the second author's research was also supported by the Natural Science Foundation of Wenzhou University (2005L002). The authors are grateful to the anonymous referee, Professors Edwin Fischer and Hans–Otto Guenther (Editors) for their valuable comments and suggestions on a former version of this paper.

Z. Chen (✉)
Department of Scientific Computing and Applied Softwares, Faculty of Science,
Xi'an Jiaotong University, 710049 Xi'an, Shaanxi, People's Republic of China
E-mail: zchen@xjtu.edu.cn

R. Lin
The School of Mathematics & Information Science, Wenzhou University,
325035 Wenzhou, Zhejiang, People's Republic of China

1 Introduction

Due to its academic and practical importance, mutual fund performance assessment has been an important area of research in finance. The past four decades have witnessed the proliferation of managed funds, international diversifications to reduce market risks, and attractiveness to many investors around the globe. Hence, there is now a pressing need for a credible and robust measure for assessing and ranking the performance of managed funds. Organizations such as Morningstar Incorporated have developed their own fund performance measures due to the increasing demand in the financial service industry.

Three of the earliest performance measures, still in use today, are the Treynor index (Treynor 1965) of the excess return per unit of the systematic risk, the Sharpe index (Sharpe 1966) of the excess return per unit of the total risk, and the Jensen's α (Jensen 1968), which is defined as the difference between the actual portfolio return and the estimated benchmark return. Since these pioneer works, numerous studies have been concerned with measuring performance in two dimensions (i.e., risk and return) by relying on the Capital Asset Pricing Model (CAPM). The results of these studies depend, to a large extent, on the benchmark portfolio used, the measurement of the risk, and the main criticism made on the use of CAPM is the validity of its underlying assumptions.

To overcome drawbacks of using the variance of portfolio returns as a risk measure and to model non-normal distributions in portfolio returns, performance measures that incorporate higher moments (see, for example, Stephens and Proffitt 1991) or that are more concerned with the downside deviation (Sortino and Price 1994) have also been developed. For example, the reward-to-half-variance index is defined as the excess return per unit of the square-root of the lower semi-variance (Ang and Chua 1979). To capture nonlinearities in β resulting from market timing activities, Ferson and Schadt (1996) modify the classic CAPM performance evaluation techniques to account for time variation in risk premiums by using a conditional CAPM framework. By assuming that portfolio returns are a function of additional influences, multi-index models (Schneeweis and Spurgin 1998; and others) are also used to identify the factors that serve as proxies for the fund risk. Although these improvements can, in some degree, model the skewness in portfolio return distributions and the time-varying risk, it is difficult for them to describe the "fat tails" phenomenon in return distributions, which is now well recognized in the risk management literature. Meanwhile, with the liberalization of financial investments and the globalization of the world economy, financial markets are more volatile than before. Extreme returns (losses) are more and more concerned by fund managers from insurance and investment companies. However, the above mentioned risk measures are incapable of modelling extreme losses. In the practical implementation, almost all the above performance evaluation approaches are based on some kind of regression models, and assume that all portfolios obey the same underlying functional relationship among different measures. This functional form is usually unknown in reality and needs to be estimated. If the relationship is incorrectly expressed before the estimation, then the results are unreliable. Therefore, in order to properly measure the fund risk and thus to evaluate the fund's performance, other new, better modern risk measures should be adopted. On the other hand, the multitude of approaches also suggests that more than one measure of risk may be needed to accurately assess performance. Except for shortcomings in measuring the fund risk, the traditional numerical indexes do not take into account the transaction costs associated with the mutual fund investment. However, an investment in mutual fund portfolios requires the subscription and redemption costs. Funds usually charge loads or other fees to recover the costs of conducting financial transactions on behalf of the investor. Consequently, the inclusion of transaction costs in portfolio performance measurement is very important.

Recently, the data envelopment analysis (DEA) (Cooper et al. 2000) technique has been adopted for assessing mutual fund performance. Contrary to other performance measures, the DEA approach has the ability to incorporate many factors that are associated with the fund performance. Especially, this approach allows to define mutual fund performance indexes that can take into account different risk measures and the investment costs. Moreover, the DEA approach can naturally envisage other output indicators, in addition to the mean return considered by the traditional indexes. Murthi et al. (1997) first introduced a DEA based relative performance measure, the DEA portfolio efficiency index (DPEI), that does not require specification of a benchmark, but incorporates transaction costs. The $I_{\mathrm{DEA}-1}$ index proposed by Basso and Funari (2001) can be seen as a generalization of the DPEI that allows consideration of different risk measures. The $I_{\mathrm{DEA}-1}$ index is extended in the same paper to a two outputs DEA performance measure, $I_{\mathrm{DEA}-2}$, which includes among the outputs a stochastic dominance indicator that reflects both the investors' preference structure and the time occurrence of the returns. The most comprehensive index for mutual fund performance measure is the generalized DEA performance indicator $I_{\mathrm{DEA}-g}$ (Basso and Funari 2002). Considering that each traditional index may be applicable under some conditions and may help to shed light on a particular aspect of the link between risk and return, $I_{\mathrm{DEA}-g}$ is derived by augmenting output variables in $I_{\mathrm{DEA}-2}$ with a few traditional performance indexes. There are also some empirical focused papers which discuss the application of existing DEA models for mutual fund performance measure under different circumstances. Typical examples include those considering returns over different lengths of time periods (McMullen and Strong 1998), multiple time horizons (Morey and Morey 1999), and so on.

For fund performance measure, the advantages of DEA typically include the following. As a non-parametric analysis technique, DEA does not require any theoretical model as the measurement benchmark. Instead, DEA measures how well a fund performs relative to the best set of funds within the declared objective category. It can address the problem of endogeneity of transaction costs by considering the transaction costs such as expense ratios, loads as well as the return simultaneously in the analysis. The greatest advantage of using the DEA method over other approaches for measuring the fund performance is that DEA reveals the reason for a fund being inefficient and shows how to restore the fund to its optimum level of efficiency.

When choosing risk measures as the inputs, only traditional measures, that is, the standard deviation of fund returns (σ), the root of the lower semi-variance ($\sqrt{HV}$), or the β coefficient, were considered in the existing DEA models for performance measure. Unfortunately, it is now known that the distribution of many financial return series is very often asymmetric and skewed, with skewness, leptokurtosis ('fat tails'), or both, pervasive. None of the above measures is suitable for describing these distribution characteristics. Moreover, there is an increasing tendency nowadays for fund managers to introduce derivative products into their portfolios. For these and other actively managed funds, their time-varying risks and asymmetric return distributions with obvious leptokurtosis pose new challenges to the performance evaluation. The introduction of risk measures being capable of describing skewed return distributions and/or distributions with fat tails is becoming more and more urgent. Therefore, in order to properly measure the fund risk, and thus to better evaluate the mutual fund performance with DEA models, it is necessary to introduce recently developed new risk measures into the DEA model.

Recent developments in risk theory suggest that quantile-based measures are well-suited for computing risk, because it is more sensible for an investor to be concerned with the risk of the loss rather than the gain. This kind of measures is suitable for asymmetric return distributions with skewness and/or fat tails. One important example is the value-at-risk (VaR) (Morgan 1996). Given a pre-specified confidence level and a particular time horizon, a portfolio's VaR is the maximal loss one expects to suffer at that confidence level by holding that portfolio over that time horizon. Nevertheless, recent research has shown that VaR has undesirable properties such as lack of subadditivity, this results in that VaR is not a coherent risk measure (Artzner et al. 1999), coherency is now regarded as a basic property that any sound risk measure should possess. Recognizing the drawbacks of VaR to mainly stem from its inability to respond to the magnitude of the possible losses below the threshold it identifies, a viable risk measure, the conditional value-at-risk (CVaR), was introduced in Rockafellar and Uryasev (2000). CVaR is defined as the conditional expectation of losses exceeding VaR in a specified period at a given confidence level. Due to its nice mathematical properties and its ease to compute, CVaR as a coherent measure is now the most promising risk measure.

To date, most studies on ranking mutual funds using the DEA technique have been mainly on US managed funds. Because of the financial market liberalization and globalization, foreign investors' interest in the Chinese financial market is growing considerably. Therefore, ranking of Chinese mutual funds would be of international interest. Results of this kind of empirical research should be instructive for the performance evaluation of mutual funds in other emerging markets.

In view of the above limitations in traditional performance indexes and existing DEA models, the main purpose of this study is to improve the performance index $I_{\mathrm{DEA}-g}$ by introducing VaR and CVaR into it. With these new risk measures and all those input–output variables ever considered in current DEA models, our new DEA fund performance indexes can take into account all the different aspects relevant to the fund performance, and allow therefore to compute an indicator of the overall performance of the mutual fund investment. Furthermore, we investigate the sensitivity of the DEA relative efficiencies to various input–output variable combinations measured across two different time periods. This is achieved by carrying out a series of empirical researches with trading data from the Chinese closed-end fund market.

The remainder of the paper is organized as follows. Section 2 discusses typical DEA models and the computation of new risk measures, and presents our DEA model for the appraisal of mutual funds. To show the applicability and properties of the proposed DEA indexes, Section 3 analyzes the results of the empirical application of our new DEA model based on the Chinese fund data. Some concluding remarks are made in Section 4.

2 A DEA model with new risk measures

In recent years, DEA has enjoyed both rapid growth and widespread acceptance. A DEA model can be analyzed in two ways, an input orientation and an output orientation. An input orientation provides information as to how much proportional reduction of inputs is necessary while maintaining the current levels of outputs for an inefficient unit to become DEA-efficient, this is very important for the fund performance appraisal. The input orientation version of the DEA model is therefore used for our empirical study. This section is divided into three parts. First, typical DEA models and fund performance indexes are briefly explained; the efficient computation of VaR and CVaR is then discussed; finally presented is our new DEA model proposed for evaluating the mutual fund performance.

2.1 The typical DEA models

Suppose we have a set of n decision making units, $j=1, \cdots, n$. For each unit, there are t outputs, $r=1, \cdots, t$ and m inputs, $i=1, \cdots, m$. Let y_{rj} (x_{ij}) be the rth (ith) known output (input) of unit j. Define

$$h_j = \frac{\sum_{r=1}^{t} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}},$$

where $u_r \geq 0$, $v_i \geq 0$ are unknown variables. The DEA relative efficiency measure h_{j_0} for a target decision making unit j_0 can be determined by solving the following famous CCR model (Charnes et al. 1978)

$$\begin{aligned} \max\ h_{j_0} &= \frac{\sum_{r=1}^{t} u_r y_{rj_0}}{\sum_{i=1}^{m} v_i x_{ij_0}} \qquad (1)\\ \text{s.t.}\ \frac{\sum_{r=1}^{t} u_r y_{rj}}{\sum_{i=1}^{m} v_i x_{ij}} &\leq 1, \quad j=1,\cdots,n,\\ \frac{u_r}{\sum_{i=1}^{m} v_i x_{ij_0}} &\geq \varepsilon, \quad r=1,\cdots,t,\\ \frac{v_i}{\sum_{i=1}^{m} v_i x_{ij_0}} &\geq \varepsilon, \quad i=1,\cdots,m. \end{aligned}$$

where ε is a positive non-Archimedean infinitesimal smaller than any positive real number and is used to prevent the weights from being zero. The above fractional programming problem can be transformed into an equivalent linear program. The input-oriented BCC model (Banker et al. 1984) is the dual of this linear program together with a constraint capturing returns to scale characteristics, and can be described as

$$\min \theta - \varepsilon \sum_{r=1}^{t} s_r^+ - \varepsilon \sum_{i=1}^{m} s_r^- \tag{2}$$

$$\begin{aligned} \text{s.t. } & x_{ij_0}\theta - s_i^- - \sum_{j=1}^{n} x_{ij}\lambda_j = 0, i = 1, \cdots, m, \\ & -s_r^+ + \sum_{j=1}^{n} y_{rj}\lambda_j = y_{rj_0}, \quad r = 1, \cdots, t, \\ & \sum_{j=1}^{n} \lambda_j = 1, \lambda_j \geq 0, \quad j = 1, \cdots, n, \\ & s_i^- \geq 0, \quad i = 1, \cdots, m, \\ & s_r^+ \geq 0, \quad r = 1, \cdots, t, \end{aligned}$$

where ε is the same as that in Eq. (1), θ and $\lambda_j \geq 0$ are dual variables, s_i^- and s_r^+ are slack variables, $\sum_{j=1}^{n} \lambda_j = 1$ is the variable returns to scale constraint. The constraints in the model ensure that $\theta \leq 1$. Denote the optimal solution of problem (2) as (θ*, λ_j*, s_i^-*, s_r^+*). The unit j_0 is called weak DEA efficiency if θ* = 1, and is DEA efficiency if θ*=1 and $s_i^-* = 0, i = 1, \cdots, m, s_r^+* = 0, r = 1, \cdots, t$. Otherwise, it is labelled as inefficient when compared to the other units.

By treating each mutual fund as a decision making unit and assigning factors relevant with the fund performance as either inputs (such as risk, transaction costs) or outputs (return), the above CCR model, BCC model and other DEA models can then be used to evaluate the relative performance of mutual funds. This is exactly what have been done in many DEA-based fund performance appraisal papers. One requirement in usual DEA models is that x_{ij} and y_{rj} must be nonnegative. However, it is very likely that returns on some funds are negative. This problem can be solved by utilizing the translation invariance property of BCC models (Pastor 1996). That is, if x_{ij} and y_{rj} are displaced by $\beta_i \geq 0$ and $\pi_r \geq 0$, respectively, such that $\widehat{x}_{ij} = x_{ij} + \beta_i \geq 0$ and $\widehat{y}_{rj} = y_{rj} + \pi_r \geq 0, j = 1, \cdots, n$. This new set of data can then be used to measure the performance of mutual funds.

The first attempt to apply the DEA methodology to obtain a mutual fund efficiency indicator is the DPEI developed by Murthi et al. (1997). Concretely, the DPEI of the target fund j_0 is defined as the optimal value of the following DEA model

$$\max \text{DPEI}_{j_0} = \frac{R_{j_0}}{\sum_{i=1}^{I} w_i c_{ij_0} + \upsilon\sigma_{j_0}} \tag{3}$$

$$\begin{aligned} \text{s.t. } & \frac{R_j}{\sum_{i=1}^{I} w_i c_{ij} + \upsilon\sigma_j} \leq 1, \quad j = 1, \cdots, J, \\ & w_i \geq \varepsilon, \ \upsilon \geq \varepsilon, \end{aligned}$$

where J is the number of funds in the category, I is the number of different transaction costs, R_j is the return rate of the jth fund, σ_j is the standard deviation of the return for the jth fund, c_{ij} is the value of the ith transaction cost for the jth fund, ε is the same as that in problem (1), and weights w_i and υ are variables of the problem.

A second DEA indicator for the mutual fund performance is the $I_{\mathrm{DEA}-1}$ index proposed by Basso and Funari (2001). This indicator differs from the DPEI in two aspects: included in it are only the investment costs which directly weigh on the investors, i.e., subscription and redemption fees; except for σ_j, other usual risk measures such as $\sqrt{HV_j}$, β_j of the jth fund are also taken into account in this index. It is clear that a highly desirable property for a mutual fund is that it is not dominated by other funds, the eventual dominance relations among the analyzed mutual funds are useful to consider in the performance measurement. This factor can be added to the outputs in the $I_{\mathrm{DEA}-1}$ index as well as fund return, which results in a two outputs DEA portfolio performance measure $I_{\mathrm{DEA}-2}$. $I_{\mathrm{DEA}-2}$ is defined as the optimal value of the following problem:

$$\begin{aligned} \max\ I_{j_0,\mathrm{DEA}-2} &= \frac{u_1 R_{j_0} + u_2 d_{j_0}}{\sum_{i=1}^{h} v_{j_0} q_{ij_0} + \sum_{i=1}^{I} w_i c_{ij_0}} \\ \text{s.t.}\ & \frac{u_1 R_j + u_2 d_j}{\sum_{i=1}^{h} v_i q_{ij} + \sum_{i=1}^{I} w_i c_{ij}} \le 1, \ j = 1, \cdots, J, \\ & u_r \ge \varepsilon, \qquad r = 1, 2, \\ & v_i \ge \varepsilon, \qquad i = 1, \cdots, h, \\ & w_i \ge \varepsilon, \qquad i = 1, \cdots, I, \end{aligned} \tag{4}$$

where $q_{1j},\cdots,q_{hj}$ are h different risk measures considered for the jth fund, d_j is the stochastic dominance indicator for the jth fund, which can be determined by using approaches in Basso and Funari (2001, 2002). Other parameters and variables have the similar meanings as those in model (3). Furthermore, by augmenting the outputs in model (4) to include a few traditional performance indexes, the generalized DEA performance index $I_{\mathrm{DEA}-g}$ can then be derived from $I_{\mathrm{DEA}-2}$.

2.2 The calculation of VaR and CVaR

Due to its intuitive meaning and being capable of answering questions like: how much one can expect to loss in one day, week, ⋯, with a given probability? VaR quickly attracted interests from both academicians and practitioners since its appearance in 1996. Here we define VaR equivalently, in terms of returns, as the minimal portfolio return for a pre-specified confidence level $\alpha\times100\%$. Assume that the portfolio return at the end of the holding period is $\mathbf{R}$, which is a random variable with the distribution function being $F : F(u) = P\{\mathbf{R} \le u\}$. Then

$$\mathrm{VaR}(\alpha) = \min\{u : F(u) \ge 1-\alpha\} = \min\{u : P\{\mathbf{R} \le u\} \ge 1-\alpha\},$$

that is, VaR(α) is the $(1-\alpha)\times100\%$ percentile of the return distribution. There are mainly three approaches for computing VaR: the analytical method, the historical simulation method and the Monte Carlo simulation method, here the normal distribution is frequently adopted for describing random returns (Duffie and Pan 1997). Therefore, it is difficult for these methods to consider the skewness and/or kurtosis of the fund return distribution. Considering this, the stable distributions, which include leptokurtic and asymmetric distributions, are utilized in this paper to properly compute VaR and CVaR of each mutual fund.

A random variable **R** is stable distributed if it has a domain of attraction (Consiglio et al. 2002), the characteristic function which identifies a stable distribution is given by

$$\Phi_{\mathbf{R}}(t) = E(\exp(it\mathbf{R})) = \begin{cases} \exp\{-\gamma^{\tau}|t|^{\tau}(1 - i\eta \operatorname{sgn}(t)\tan(\pi\tau/2)) + i\delta t\} & \text{if } \tau \neq 1 \\ \exp\{-\gamma|t|(1 - i\eta\frac{2}{\pi}\operatorname{sgn}(t)\log(t)) + i\delta t\} & \text{if } \tau = 1. \end{cases}$$

Thus a stable distribution is identified by four parameters: the index of stability $\tau \in (0, 2]$ which is a coefficient of kurtosis, the skewness parameter $\eta \in [-1, 1]$, $\delta \in \Re$ and $\gamma \in \Re^{+}$, which are the location and the dispersion parameters, respectively. When $\tau = 2$ and $\eta = 0$, the stable distribution has a Gaussian density. stable distributions with $\tau < 2$ are leptokurtotic and present fat tails. $\eta > 0$ identifies distributions whose tails are more extended towards right, while $\eta < 0$ typically characterizes distributions whose tails are extended towards the negative values of the distribution.

For the practical implementation, each fund's VaR can be computed as follows.

Step 1. Basing on each fund's historical return data, the stable distribution parameters τ, η, γ and δ are estimated by maximizing the likelihood function

$$L(\tau, \eta, \gamma, \delta) = \prod_{R} f(R|\tau, \eta, \gamma, \delta),$$

where $f(R|\tau, \eta, \gamma, \delta)$ is the density function of a stable random variable (Kogon and Williams 1998). In our empirical experiment, the STABLE program[1] based on Nolan (1997, 1999) is used to estimate four parameters.

Step 2. With the estimated stable distribution, simulate a sufficiently large number, say N=10,000, of stable distributed random numbers as the possible future returns of the considered fund.

Step 3. The fund's VaR can be estimated by computing the empirical quantile of the simulated return distribution. That is, it is taken as the $(1 - \theta) \times N$ $-$th smallest value in the monotonically ordered random number series, here θ is the selected confidence level.

Nevertheless, VaR suffers from being unstable and difficult to work with numerically (Duffie and Pan 1997). To overcome drawbacks of VaR and to provide a proper risk measure, CVaR can be adopted. In terms of the portfolio return, CVaR can be defined as the conditional expectation of portfolio returns below the VaR return. Unlike VaR, CVaR possesses good continuity and can be easily computed (Rockafellar and Uryasev 2000). To ensure the consistence with the calculation of VaR, we also estimate CVaR by relying on the stable distribution. According to the above definition, each fund's CVaR can be directly computed by utilizing the above simulated data during the calculation of VaR. Concretely,

[1] See the web site http://www.cas.american.edu/jpnolan

Step 4. The CVaR value of the corresponding fund is taken as the mean of all those simulated returns which are not larger than the $(1-\theta)\times N-\text{th}$ smallest value in the ordered random number series.

2.3 The DEA model with new risk measures

By introducing new risk measures VaR and CVaR into its inputs, our new DEA model is an improvement on the DEA model corresponding to the performance index $I_{\text{DEA}-g}$. This improvement, as discussed in Section 1, is rather important for properly describing mutual funds' risks and thus evaluating their performances. Therefore, the new multiple input–multiple output DEA performance indicator for the target mutual fund $j_0(1\le j_0\le n)$ is determined as the optimal value of the following fractional program

$$
\begin{aligned}
&\max \frac{u_1 o_{j_0}+u_2 d_{j_0}+\sum_{r=1}^{p}\omega_r I_{rj_0}}{\sum_{i=1}^{h}\upsilon_i q_{ij_0}+\sum_{k=1}^{K}\omega_k c_{kj_0}} \\
&\text{s.t. } \frac{u_1 o_j+u_2 d_j+\sum_{r=1}^{p}\omega_r I_{rj}}{\sum_{i=1}^{h}\upsilon_i q_{ij}+\sum_{k=1}^{K}\omega_k c_{kj}}\le 1,\ j=1,\cdots,n, \\
&\quad u_r\ge\varepsilon, \qquad r=1,2, \\
&\quad \omega_r\ge\varepsilon, \qquad r=1,\cdots,p, \\
&\quad \upsilon_i\ge\varepsilon, \qquad i=1,\cdots,m, \\
&\quad \omega_k\ge\varepsilon, \qquad k=1,\cdots,K,
\end{aligned}
\tag{5}
$$

where, for each fund $j: 1\le j\le n$, the output variable o_j is either the expected return or the expected excess return, d_j is the stochastic dominance indicator, see Basso and Funari (2001, 2002) for its determination, I_{rj}, r=1, ···, p, are the traditional performance indexes, which might include the Treynor index, the Sharpe index, and the Jensen's α. The input indicators are composed of K investment costs c_{1j}, ···, c_{Kj}, which can be the sales charge, redemption fees, administrative expenses, advisory fees and other operational expenses, and h risk measures q_{1j}, ···, q_{hj}, which, except for traditional measures such as σ_j, $\sqrt{HV_j}$, β_j of the jth fund, include new types of risk measures like VaR and CVaR. The rest coefficients or variables have the same meanings as those in model (4). ε is a convenient small positive number that prevents the weights from being zero (Cooper et al. 2000).

For computational convenience in real applications, just like that from problem (1) to problem (2), the problem (5) can be converted into an input-oriented BCC-type linear program, the mutual fund's relative performance is then evaluated by solving this linear programming problem.

Although problem (5) has a similar form as the DEA model for determining $I_{\text{DEA}-g}$, this new model not only takes into account all those factors ever considered by traditional fund performance indexes and existing DEA performance indicators, but include, in its inputs, modern risk measures like VaR and CVaR. These new measures are very useful for describing the asymmetry, and/or the fat-tailedness of complexly composed or dynamically adjusted funds' return distributions, which make it possible for us to compute a reasonable and robust indicator of the overall performance of the mutual fund investment.

3 Empirical analysis and discussion of the results

The effect of the proposed new DEA models for the mutual fund performance measure will be empirically examined in this section, the input–output data used here comes from the China Security Investment Fund Research Database, which is compiled by the GTA Information Technology Company, Shenzhen, China. As an emerging security market, the Chinese closed-end fund market has only a very short history. The first closed-end fund in China was sold to public in April 1998. Nevertheless, the industry has grown steadily since then. By the end of 2002, there were more than 50 publicly traded closed-end funds with more than 100 billion RMB in floating volume. These funds have become a very important force in Chinese security markets. Comparatively, there are only a few open-ended mutual funds in China, it is difficult to collect the necessary trading data on them. We will thus only consider close-ended mutual funds in the following.

Since the Chinese security market has been a bear market in recent 3 years, the net asset value of most funds decreased significantly during this period. Due to this, our empirical analysis would be divided into two parts: for the first part, the relative efficiency of mutual funds in the period 11/1/2000 to 12/31/2002 is investigated; for the second part, the relative efficiency of mutual funds is examined for the period 9/1/1999 to 9/1/2000. In this way, we can not only compare the DEA performance appraisals for both the normal and the abnormal fund markets, but investigate how the whole market environment could affect the DEA performance evaluation results. We use daily data of chosen funds. For each fund, the daily return with dividend reinvestmented is used to calculate its expected return, the standard deviation and lower semi-variance of returns, the β coefficient and other indictors. Due to the limitation of data availability, only the average annually total cost for each fund could be obtained in the considered time periods, which is thus used as the single transaction cost factor. Except for this indicator, we have considered, among the inputs, various combinations of three classic risk measures (σ, the β coefficient, and $\sqrt{HV}$) and two new risk measures VaR and CVaR.

3.1 The performance appraisal for the period 11/1/2000 to 12/31/2002

There were only 33 publicly traded closed-end funds in the Chinese closed end fund market before October 2000. The relative performance of these funds is thus investigated here. First of all, to get an overview about the distribution properties of evaluated funds, we display in Table 1 parameter estimates of a stable density fitted on the return data of each fund. The STABLE program is used to estimate parameters τ, η, γ, δ. Note that all the funds have an index of stability τ that is lower than 2, while the skewness parameter η is obviously different from zero, with most being positive. Therefore, all these funds' return distributions show statistically significant leptokurtotic and skewness that often stretches the tails on the positive returns. According to our discussion in Section 1, the introduction of VaR and CVaR into the DEA model is very important for us to properly evaluate these funds' performance.

Table 2 reports, for each fund j (identified by its fund name), the average return $\overline{R}_j$, the average transaction cost c_j, the Jensen index α_j, β_j, VaR_j, CVaR_j, σ_j, and $\sqrt{HV_j} = \sqrt{\sum \left(\min\left(\mathbf{R}_j - \overline{R}_j, 0\right)\right)^2}$ of the fund's random return $\mathbf{R}_j$. Here $\beta_j = \mathrm{Cov}\left(\mathbf{R}_j, \mathbf{R}_m\right)/\mathrm{Var}(\mathbf{R}_m)$, $\mathrm{Cov}\left(\mathbf{R}_j, \mathbf{R}_m\right)$ is the co-variance between $\mathbf{R}_j$ and the daily market return with dividend reinvestmented $\mathbf{R}_m$, Var($\mathbf{R}_m$) is the standard deviation of $\mathbf{R}_m$; α_j is the intercept in the regression equation obtained by

Table 1 Estimated parameters of a stable density fitted on return data of each fund

Fund name	τ_j	η_j	γ_j	δ_j
Longyuan	1.7223	0.1604	0.8919	−0.1600
Hanbo	1.6374	0.0450	0.8628	−0.1597
Xing'an	1.7058	0.1197	0.8386	−0.1322
Handing	1.4658	0.0835	0.9532	−0.2008
Jinding	1.5886	0.0241	0.8168	−0.1394
Xingke	1.6768	0.0718	0.8197	−0.1308
Jinyuan	1.7006	0.0925	0.8719	−0.1563
Jinsheng	1.5726	−0.0355	0.7875	−0.0949
Yuze	1.5547	0.0598	0.7622	−0.1121
Tongzhi	1.6327	0.0515	0.7524	−0.1409
Yuhua	1.7210	0.1312	0.7599	−0.1125
Hanxing	1.7055	−0.0758	0.6779	−0.0534
Jingfu	1.7359	−0.0169	0.7360	−0.0539
Tianyuan	1.7049	−0.1063	0.7241	−0.0062
Pufeng	1.7419	0.0202	0.6945	−0.0513
Xinghe	1.8554	−0.4309	0.7111	0.0199
Yulong	1.7194	0.1573	0.7072	−0.0748
Anshun	1.8506	−0.2002	0.6962	0.0055
Hansheng	1.8524	−0.0649	0.7669	−0.0371
Jinghong	1.6942	0.1412	0.8197	−0.0939
Tongyi	1.8536	−0.4525	0.7102	0.0428
Taihe	1.7457	−0.0537	0.7965	0.0052
Jinxin	1.7598	−0.0113	0.7051	−0.0351
Tongsheng	1.7381	−0.0087	0.6738	−0.0235
Yuyuan	1.6868	0.0166	0.6765	−0.0256
Jingyang	1.7098	−0.1775	0.8394	0.0354
Jingbo	1.4715	0.0904	0.9137	−0.0887
Puhui	1.6901	−0.1122	0.6829	−0.0071
Yuyang	1.7920	0.0125	0.7547	−0.0616
Anxin	1.5857	0.2231	0.5944	−0.0785
Xinghua	1.6742	0.3235	0.6372	−0.0952
Jintai	1.7450	0.1871	0.6496	−0.0703
Kaiyuan	1.7901	0.1288	0.7391	−0.0595

regressing the excess daily return of fund j on the excess daily market return in the examined time period; VaR_j and CVaR_j, based on parameter estimates in Table 1, are computed by using the method presented in subsection 2.2, with θ=95%.

Note that all the mean returns in Table 2 are negative, which entails that the values of the Sharpe, Treynor and reward-to-half-variance indexes are negative and, above all, meaningless. The only traditional performance index that can be included in the outputs is then the Jensen index α_j, which is also uniformly negative. Therefore, in order to satisfy the nonnegativity requirement on output factors, we utilize the translation invariance property of BCC models. Concretely, $\overline{R}_j$ and α_j are displaced by 0.15 and 0.16, respectively, before we solve the corresponding input-oriented BCC-type linear program derived from problem (5).

Table 2 Values of different indicators for the chosen Chinese mutual funds

Fund name	β_j	σ_j	$\sqrt{\mathrm{HV}_j}$	VaR_j	CVaR_j	c_j	$\overline{R}_j$	α_j
Longyuan	0.92263	1.58962	1.03912	2.37625	3.35982	0.02015	−0.07911	−0.09545
Hanbo	0.97461	1.71368	1.08483	2.48364	3.53843	0.02292	−0.08083	−0.09809
Xing'an	0.90809	1.51792	1.00909	2.25550	3.25187	0.02256	−0.07473	−0.09080
Handing	0.90561	2.17011	1.50377	3.05269	5.32136	0.01955	−0.13866	−0.15470
Jinding	0.96657	1.61799	1.06117	2.43202	3.61732	0.02388	−0.09901	−0.11612
Xingke	0.87801	1.51362	1.01010	2.26856	3.34069	0.02349	−0.08692	−0.10247
Jinyuan	0.98204	1.63188	1.06088	2.38842	3.44946	0.02123	−0.08829	−0.10567
Jinsheng	0.95085	1.64114	1.10056	2.39539	3.88285	0.02211	−0.08796	−0.10479
Yuze	0.90984	1.57424	1.08326	2.26701	3.75914	0.02205	−0.08294	−0.09905
Tongzhi	0.85410	1.55378	1.02989	2.16809	3.49462	0.02188	−0.10216	−0.11728
Yuhua	0.77645	1.38116	0.98230	2.01804	2.99347	0.02325	−0.06894	−0.08270
Hanxing	0.67314	1.37539	0.97926	1.87791	3.16912	0.02206	−0.06723	−0.07916
Jingfu	0.79437	1.49198	1.05117	1.96572	3.31617	0.01928	−0.05554	−0.06961
Tianyuan	0.71081	1.56559	1.21967	1.97439	4.06003	0.02636	−0.05647	−0.06906
Pufeng	0.76494	1.60610	1.26122	1.83155	3.42702	0.02101	−0.05961	−0.07316
Xinghe	0.62591	1.49056	1.13955	1.83055	3.24425	0.02287	−0.04856	−0.06265
Yulong	0.68020	1.67387	1.33861	1.83564	3.65562	0.02475	−0.06947	−0.08153
Anshun	0.62505	1.53797	1.23705	1.75229	3.37826	0.02833	−0.05014	−0.06270
Hansheng	0.65970	1.62172	1.28575	1.93320	3.47862	0.02430	−0.06992	−0.08317
Jinghong	0.81686	1.76986	1.35789	2.16519	3.93195	0.02207	−0.07224	−0.08670
Tongyi	0.74914	1.87805	1.62934	1.81388	4.08869	0.02961	−0.07855	−0.09182
Taihe	0.83981	1.73931	1.35559	2.07277	3.94562	0.02565	−0.04123	−0.05610
Jinxin	0.71138	1.47387	1.10556	1.84072	3.31227	0.04144	−0.05194	−0.06454
Tongsheng	0.71514	1.47534	1.14939	1.76759	3.34758	0.02467	−0.05391	−0.06658
Yuyuan	0.67737	1.66829	1.33571	1.81364	3.58946	0.02645	−0.04968	−0.06329
Jingyang	0.76111	1.74127	1.37221	2.28783	4.18645	0.02366	−0.05248	−0.06596
Jingbo	0.94881	2.05256	1.45820	2.79759	5.20241	0.02346	−0.04618	−0.06297
Puhui	0.74486	1.82532	1.55034	1.88639	3.80947	0.02599	−0.07022	−0.08341
Yuyang	0.62670	1.84009	1.53423	1.95409	3.73115	0.02685	−0.08525	−0.09784
Anxin	0.62690	1.91548	1.65091	1.60122	4.21836	0.03279	−0.08358	−0.09618
Xinghua	0.57584	1.63162	1.31365	1.61729	3.35793	0.02758	−0.04739	−0.06037
Jintai	0.64579	1.69125	1.40521	1.66042	3.43329	0.02606	−0.06584	−0.07729
Kaiyuan	0.72887	1.92907	1.64527	1.86856	3.37594	0.02762	−0.07722	−0.08719

To investigate the sensitivity of the DEA relative efficiency to various input factor-output factor combinations, we consider 22 combinations of input–output indicators in this analysis, see Table 3 for details. Our emphasis is put on different combinations of five risk measures, the reasons are: as quantile-based risk measures, VaR and CVaR only concern the lower tail of the fund's return distribution. This might not be enough for describing the fund's full domain return information. In this regard, VaR or CVaR should be combined with other risk measures so that different risk characteristics of the fund return can be simultaneously modelled. Meanwhile, one of the DEA model's advantages lies in its flexibility in including multiple input and multiple output factors for the relative efficiency evaluation. In view of these points, it might be better to add, except for VaR or CVaR, usual risk measures in the inputs of the DEA model. The dual simplex algorithm in Matlab 7.0 is used to solve each derived BCC-type linear program. To avoid the floating overflow caused by extremely large numbers in the input and output matrices, we pre-process the raw input-matrix and output-matrix by standardizing their elements. Specifically, elements in each row of these two matrices are divided by the maximal element of that row. The relative rankings of 22 DEA runs are given in Table 4.[2]

First, we examine the influence of α and β on the DEA relative efficiency. From the relevant relative rankings (under runs 1–1 to 1–16) in Table 4, it can be seen that, whether including α or not does not significantly affect the evaluation result, except for *Jinxin*. For this fund, if σ_j or $\sqrt{HV_j}$ is chosen as the risk measure, the inclusion of α considerably affects the relative ranking, and causes it to change from DEA inefficiency to DEA efficiency. The reason for this variation is due to *Jinxin*'s larger α and smaller value of the corresponding risk measure, when compared with other funds. For both cases with and without the β coefficient, the correlation coefficients of relative rankings for each pair of DEA runs considering α and not considering α, respectively (for example, runs 1–1 and 1–9), are all greater than 0.988. Therefore, for Chinese closed-end funds, it is not important to consider the Jensen index in DEA models. One explanation for this is: classical performance measures depend on the CAPM, which does not hold for the Chinese security markets (Ma et al. 2000). Unlike α, whether to include β or not does impact the performance appraisal. The most significant change occurs for *Yuyang*, *Anxin*, *Xinghua* and *Jintai*, whose relative rankings increase considerably with the inclusion of β. This phenomenon is caused by these four funds' very small β coefficients, which are the smallest ones among all funds but those DEA efficient funds. For both cases with and without α, the correlation coefficients of relative rankings for each pair of DEA runs including β and not including β, respectively (for example, runs 1–1 and 1–13), are all smaller than 0.897. Consequently, we might conclude that, whether to include the β coefficient or not would affect the performance evaluation of Chinese closed-end funds.

Second, we investigate the DEA relative efficiency when only one of $\sigma_{j,}\sqrt{HV_j}$, VaR_j, or CVaR_j is chosen as the risk measure (i.e., runs 1–5 to 1–8). It can be seen from Table 4 that, for the first 11 funds, $R_{1-6,j} \geq R_{1-5,j}$, $1 \leq j \leq 11$, while $R_{1-6,j} \leq R_{1-5,j}$, $12 \leq j \leq 33$, for the rest 22 funds except for *Jinghong*. The reasons for the

[2] Due to the space limitation, we do not report here corresponding efficiency scores, which can be provided upon request.

Table 3 Indicators considered in different DEA runs

DEA run	β_j	σ_j	$\sqrt{HV_j}$	VaR_j	$CVaR_j$	c_j	$\overline{R}_j$	α_j
Run 1–1	√	√				√	√	√
Run 1–2	√		√			√	√	√
Run 1–3	√			√		√	√	√
Run 1–4	√				√	√	√	√
Run 1–5		√				√	√	
Run 1–6			√			√	√	
Run 1–7				√		√	√	
Run 1–8					√	√	√	
Run 1–9	√	√				√	√	
Run 1–10	√		√			√	√	
Run 1–11	√			√		√	√	
Run 1–12	√				√	√	√	
Run 1–13		√				√	√	√
Run 1–14			√			√	√	√
Run 1–15				√		√	√	√
Run 1–16					√	√	√	√
Run 1–17		√		√		√	√	
Run 1–18		√			√	√	√	
Run 1–19	√	√		√		√	√	
Run 1–20	√	√			√	√	√	
Run 1–21	√	√	√	√		√	√	
Run 1–22	√	√	√		√	√	√	

above phenomenon are: $\sqrt{HV_j}$ only measures those "negative returns" below the mean return, but σ_j measures both positive and negative deviations of the return from its expected value; all the first 11 funds but one have positive skewness parameters, while the skewness parameters of rest 22 funds are mostly negative. Intuitively, funds with a positive skewness should perform better than funds with a negative skewness. Therefore, $\sqrt{HV_j}$ is better than σ_j in terms of reflecting the fund's return distribution characteristics. By comparing results obtained under DEA runs adopting VaR or CVaR as the risk measure with those obtained under DEA runs using σ_j or $\sqrt{HV_j}$ as the risk measure, it can be seen that the relative rankings of many funds, such as *Hanxing*, *Yuhua*, *Pufeng*, *Anxin* and *Xinghua*, change greatly. Especially, the DEA efficient (inefficient) fund *Hanxing* (*Yuhua*) under DEA runs with σ_j or $\sqrt{HV_j}$ as the risk measure becomes DEA inefficient (efficient) under the DEA run using CVaR as the risk measure. We first analyze *Hanxing*, *Yuhua*, *Taihe*, *Xinghe*, *Jingfu*, the last three funds are DEA efficient under all four DEA runs. The expected return of *Taihe* is the highest one among all 33 funds. The DEA efficiency of *Taihe* thus comes from its high return rate. Although *Yuhua*'s σ_j and $\sqrt{HV_j}$ are slightly bigger than corresponding values of *Hanxing*, its $CVaR_j$ is the smallest one among these five funds. This makes *Yuhua* become DEA efficient under the DEA run adopting CVaR as the risk measure. *Hanxing*'s σ_j and $\sqrt{HV_j}$ are the smallest ones among corresponding values of five funds, the DEA efficiency of *Hanxing* under DEA runs with σ_j or $\sqrt{HV_j}$ as the risk measure

Table 4 Relative rankings in different DEA runs

Fund name	Run 1–1	Run 1–2	Run 1–3	Run 1–4	Run 1–5	Run 1–6	Run 1–7	Run 1–8	Run 1–9	Run 1–10	Run 1–11	Run 1–12	Run 1–13	Run 1–14	Run 1–15	Run 1–16	Run 1–17	Run 1–18	Run 1–19	Run 1–20	Run 1–21	Run 1–22
Longyuan	13	9	13	9	11	7	12	9	13	8	13	9	12	8	12	9	15	11	16	12	14	9
Hanbo	31	23	32	27	26	17	32	18	31	23	32	27	26	18	32	18	33	20	33	29	27	24
Xing'an	17	11	30	13	13	10	30	11	17	11	30	13	13	10	30	11	19	13	20	14	16	13
Handing	11	10	10	8	9	9	9	7	11	10	10	8	9	9	9	7	14	10	14	11	15	12
Jinding	29	24	33	29	22	18	33	24	29	24	33	29	23	19	33	25	31	26	32	31	28	25
Xingke	21	15	31	20	17	12	31	16	21	15	31	20	17	12	31	16	22	16	23	20	18	16
Jinyuan	25	16	21	18	19	13	18	15	25	16	21	18	19	13	18	15	25	15	27	18	19	17
Jinsheng	28	22	29	31	20	16	29	27	28	22	29	31	20	16	29	27	28	22	31	30	25	23
Y uze	22	21	28	28	18	14	27	23	22	21	28	28	18	15	28	24	23	19	24	24	24	21
Tongzhi	18	13	25	21	14	11	22	17	18	13	25	21	14	11	22	17	20	17	21	21	17	14
Y uhua	8	8	18	1	7	6	20	1	8	7	18	1	7	7	20	1	12	1	13	1	12	1
Hanxing	1	1	9	1	1	1	11	8	1	1	9	1	1	1	11	8	1	1	1	1	1	1
Jingfu	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Tianyuan	19	30	26	30	16	24	25	31	20	32	26	30	16	24	25	31	21	18	22	23	26	27
Pu f eng	12	14	1	12	15	19	1	13	12	14	1	11	15	21	1	14	1	14	1	13	1	15
Xinghe	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Yulong	20	25	12	22	27	27	13	28	19	25	12	22	27	27	13	28	16	29	15	22	20	26
Anshun	10	12	17	10	10	20	15	12	10	12	17	10	10	20	15	12	13	12	12	10	13	11
Hansheng	15	19	16	17	25	26	16	20	15	18	16	17	25	26	16	20	18	23	19	17	23	20
Jinghong	24	27	20	24	24	22	23	25	24	27	20	24	24	23	23	26	29	28	26	26	30	29
Tongyi	33	33	24	33	32	32	21	32	33	33	24	33	32	32	21	32	27	32	30	33	33	33
Taihe	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Jinxin	1	1	27	11	6	8	26	10	7	9	27	12	1	1	27	10	11	8	11	8	1	1
Tongsheng	9	18	11	14	8	15	10	14	9	19	11	16	8	14	10	13	1	9	1	9	1	10
Yuyuan	26	29	15	25	21	25	14	19	26	29	15	25	21	25	14	19	17	21	18	27	22	30

Table 4 (continued)

Fund name	Run 1–1	Run 1–2	Run 1–3	Run 1–4	Run 1–5	Run 1–6	Run 1–7	Run 1–8	Run 1–9	Run 1–10	Run 1–11	Run 1–12	Run 1–13	Run 1–14	Run 1–15	Run 1–16	Run 1–17	Run 1–18	Run 1–19	Run 1–20	Run 1–21	Run 1–22
Jingyang	27	28	22	26	23	23	28	26	27	28	22	26	22	22	26	22	32	27	28	28	31	31
Jingbo	7	7	8	7	5	5	8	5	6	6	8	7	6	6	8	6	10	6	10	7	11	8
Puhui	30	31	19	32	29	29	17	30	30	31	19	32	29	29	17	30	24	31	25	32	29	32
Yuyang	14	17	14	15	30	30	24	29	14	17	14	14	30	30	24	29	30	30	17	15	21	18
Anxin	23	26	1	23	33	33	1	33	23	26	1	23	33	33	1	33	1	33	1	25	1	28
Xinghua	1	1	1	1	12	21	1	6	1	1	1	1	11	17	1	1	1	7	1	1	1	1
Jintai	16	20	1	16	28	28	1	22	16	20	1	15	28	28	1	21	1	25	1	16	1	19
Kaiyuan	32	32	23	19	31	31	19	21	32	30	23	19	31	31	19	23	26	24	29	19	32	22

is due to its low risk value. Similarly, the reason for *Pufeng, Anxin* and *Xinghua* to change from the DEA inefficiency under other three DEA runs to the DEA efficiency under the DEA run using VaR as the risk measure is due to their very small VaR values, when compared with values of other risk measures.

The Chinese closed-end funds mainly consist of stocks, corporate bonds, government bonds and currency accounts, among which stocks should be the riskiest kind of securities. For each fund, the ratio of its included stocks' total market value to the sum of capital values of corporate bonds, government bonds and currency accounts included in it can be used as an indicator of the fund's riskiness. The higher the ratio, the riskier the fund in terms of its investment structure. The corresponding ratios of *Taihe*, *Hanxing*, *Xinghe*, *Jingfu* and *Yuhua* are 2.1925, 1.9475, 1.9051, 1.8863 and 1.3742, respectively. The smallest ratio of *Yuhua* explains why it becomes DEA efficient under the DEA run with CVaR used as the risk measure. Meanwhile, *Hanxing*'s ratio is the second largest one among these five funds. Therefore, the inclusion of CVaR can indeed better reflect the fund's potential risk, and thus appraise its relative efficiency. Moreover, the skewness parameters of funds *Yuhua*, *Pufeng*, *Anxin* and *Xinghua* are all significantly positive, while the skewness parameter of *Hanxing* is negative. Finally, it can be seen from Table 4 that, the relative rankings of most funds gotten under DEA runs with VaR or CVaR used as the risk measure are near to each other. This is because both CVaR and VaR measure the loss risk by considering the left tail distribution of the fund return. In a word, the above discussion shows the significant advantage of VaR and CVaR in evaluating mutual funds' relative efficiency, when compared with σ_j and $\sqrt{HV_j}$.

Thirdly, we analyze results obtained under DEA runs which combine VaR or CVaR with other risk measures (runs 1–17 to 1–22). Compared with those gotten under DEA runs just using VaR or CVaR as the risk measure, it is easy to see that relative rankings of many funds do change to some extent. In particular, *Xinghua* is always DEA efficient after the inclusion of the β coefficient; *Hanxing* changes from DEA inefficiency when CVaR is used as the single risk measure to DEA efficiency when CVaR is combined with other risk measures; *Jinxin* becomes DEA efficient after the inclusion of $\sqrt{HV_j}$. To further compare results obtained under DEA runs with different risk measure combinations, we computed the correlation coefficients among corresponding relative rankings, which are shown in Tables 5 and 6, respectively. It can be seen from Tables 5 and 6 that most of the correlation coefficients are not high, except for a few large values which are due to one of the following reasons: the large correlation coefficient (0.848) between β and VaR (runs 1–7 and 1–11, 1–17 and 1–19); the large correlation coefficient (0.826) between σ_j and CVaR (runs 1–8 and 1–18, 1–12 and 1–20); the increasing number

Table 5 Correlation coefficients among relative rankings obtained under DEA runs 1–7, 1–11, 1–17, 1–19 and 1–21

	Run 1–11	Run 1–17	Run 1–19	Run 1–21
Run 1–7	0.97022	0.86120	0.83289	0.73248
Run 1–11		0.82766	0.85194	0.72986
Run 1–17			0.96949	0.83808
Run 1–19				0.94234

Table 6 Correlation coefficients among relative rankings obtained under DEA runs 1–8, 1–12, 1–18, 1–20 and 1–22

	Run 1–12	Run 1–18	Run 1–20	Run 1–22
Run 1–8	0.88652	0.94438	0.82191	0.87936
Run 1–12		0.83149	0.97361	0.89185
Run 1–18			0.85637	0.88191
Run 1–20				0.94930

of common risk measures used in relevant DEA runs (runs 1–21 and 1–19, 1–20 and 1–22). Consequently, the combination of VaR or CVaR with traditional risk measures does affect the performance appraisal. The DEA efficiency evaluation with VaR or CVaR can be further improved if it is combined with one or more suitably chosen traditional risk measures.

The above empirical results show the necessity and practical value for introducing new risk measures into the DEA model for mutual fund performance evaluation. Considering that the Chinese security markets have been a bear market during the examined time period, we will further examine these models by using another set of data in the next subsection.

3.2 The performance appraisal for the period 9/1/1999 to 9/1/2000

To find out the effect of our new DEA models on the mutual fund performance evaluation under the normal market environment, the relative efficiency of Chinese mutual funds for the period 9/1/1999 to 9/1/2000 is examined here. Evaluated are all 14 closed-end funds which were sold to public before August 1999. Just as that for the previous group of funds, parameter estimates[3] of a stable density fitted on daily returns of each of 14 funds show that the return distribution of each fund exhibits statistically significant skewness and kurtosis. VaR and CVaR should then be very suitable for measuring these funds' risk.

The input indicators are the same as those considered in the previous subsection. To ensure the evaluation effect, the number of decision making units in a DEA model should be at least three times of the sum of the number of input indicators and the number of output indicators (Cooper et al. 2000). Due to this requirement and the small number of funds, the expected return is chosen as the unique output indicator in the following. Values of inputs and outputs are computed in the same way as that in the last subsection and are reported in Table 7.

Different from what in Table 2, it can be seen from Table 7 that most of expected returns are now positive, which thus represents the normal situation about Chinese mutual funds. Since there are still four negative expected returns, the average returns are all displaced by 0.03 before we solve the corresponding input-oriented BCC-type linear program, derived from problem (5), by the dual simplex algorithm in Matlab 7.0. Eight different combinations of input indicators, denoted as DEA runs 2–1 to 2–8 and shown in Table 8, will be examined. The relative rankings of these DEA runs are reported in Table 9. To easily compare

[3] Again, due to the space limitation, we do not present concrete estimates here, which can be provided upon request.

Table 7 Values of different indicators for the chosen Chinese mutual funds

Fund name	β_j	σ_j	$\sqrt{HV_j}$	VaR_j	CVaR_j	c_j	$\overline{R}_j$
Pufeng	0.40656	1.05935	0.65106	1.29678	1.77961	0.01809	0.00165
Xinghe	0.37343	1.00413	0.59708	1.33944	1.62170	0.01934	−0.01026
Yulong	0.64496	1.49662	0.88220	1.93941	2.55882	0.02536	0.03634
Anshun	0.56082	1.38569	0.82821	1.81193	2.40201	0.02399	0.02027
Hansheng	0.70869	1.68066	1.14719	1.89649	3.42960	0.03032	−0.01058
Jinghong	0.69344	1.75178	1.24080	1.98329	3.49551	0.02947	−0.01741
Tongyi	0.64096	1.66652	1.17617	1.80521	3.31715	0.03141	0.03749
Taihe	0.57271	1.29788	0.74682	1.63644	2.13242	0.02662	0.02914
Puhui	0.62318	1.87726	1.40749	1.88502	3.89802	0.03258	0.00798
Yuyang	0.71022	2.18790	1.69501	2.00819	4.58211	0.03856	0.01046
Anxin	0.80085	2.09666	1.51444	2.21869	4.21344	0.04099	0.01164
Xinghua	0.69960	2.23873	1.78272	1.81485	4.44250	0.03474	−0.01219
Jintai	0.63458	1.70302	1.23185	1.80440	3.40903	0.03187	0.01139
Kaiyuan	0.72392	1.94762	1.44598	1.89733	3.99072	0.03765	0.02272

performances of 14 funds in two periods adopted here and that in the last subsection, the relative performance of these funds, under the same eight DEA runs, in the period 11/1/2000 to 12/31/2002 is reexamined by using their input and output values in Table 2, these DEA runs are denoted as 3–1 to 3–8 for ease of distinction. The corresponding appraisal results are given in Table 10.

Compared with DEA runs using σ_j or $\sqrt{HV_j}$ as the risk measure (runs 2–1, 2–2), the relative rankings obtained under the DEA run using CVaR as the risk measure (run 2–4) do not alter obviously; nevertheless, the corresponding results gotten under the DEA run using VaR as the risk measure (run 2–3) do change significantly for nearly half of 14 funds, especially for *Xinghe, Jinghong, Puhui* and *Xinghua*. While being DEA efficient under three DEA runs with σ_j, $\sqrt{HV_j}$, or CVaR as the risk measure, respectively, *Xinghe* becomes inefficient under the current DEA run. The relative ranking of *Puhui* changes from 11 under those three DEA runs to nine under the present DEA run. With regard to the corresponding risk value of efficient funds *Pufeng*, *Yulong*, *Tongyi* and *Taihe*, *Xinghe*'s σ_j, $\sqrt{HV_j}$, or CVaR is always the smallest one, but its VaR value is significantly larger than that of *Pufeng*. This is exactly the reason that makes *Xinghe* inefficient under the DEA

Table 8 Indicators considered in different DEA runs

DEA run	β_j	σ_j	$\sqrt{HV_j}$	VaR_j	CVaR_j	c_j	$\overline{R}_j$
Run 2–1 3–1		√				√	√
Run 2–2 3–2			√			√	√
Run 2–3 3–3				√		√	√
Run 2–4 3–4					√	√	√
Run 2–5 3–5		√		√		√	√
Run 2–6 3–6		√			√	√	√
Run 2–7 3–7	√			√		√	√
Run 2–8 3–8	√				√	√	√

Table 9 Relative rankings in different DEA runs

Fund name	Run 2–1	Run 2–2	Run 2–3	Run 2–4	Run 2–5	Run 2–6	Run 2–7	Run 2–8
Pufeng	1	1	1	1	1	1	1	1
Xinghe	1	1	5	1	1	1	1	1
Yulong	1	1	1	1	1	1	1	1
Anshun	6	6	6	6	6	6	6	6
Hansheng	9	10	12	10	12	9	12	12
Jinghong	10	8	13	8	13	10	13	11
Tongyi	1	1	1	1	1	1	1	1
Taihe	1	1	1	1	1	1	1	1
Puhui	11	11	9	11	9	11	9	9
Yuyang	12	13	11	13	11	13	11	10
Anxin	13	14	14	14	14	12	14	13
Xinghua	14	12	10	12	10	14	10	14
Jintai	7	7	8	7	8	7	8	8
Kaiyuan	8	9	7	9	7	8	7	7

run using VaR as the risk measure. For *Xinghe* (*Pufeng*), the ratio of its included stocks' market value to the sum of capital values of corporate bonds, government bonds and currency accounts included in it is 2.71 (2.56). In this sense, *Xinghe* is riskier than *Pufeng*, just as their VaR values show. The performance deterioration of *Jinghong* under the DEA run with VaR as the single risk measure is due to its large VaR value, which is the third largest one among those of 14 funds, and its average σ_j, $\sqrt{HV_j}$, and CVaR values. Just the opposite, the performance improvement of *Puhui* and *Xinghua* is due to their large values of σ_j, $\sqrt{HV_j}$, and CVaR and average values of VaR.

To more comprehensively examine the influence of different risk measures, VaR or CVaR is also combined with traditional risk measures. Since the correlation

Table 10 Relative rankings in different DEA runs

Fund name	Run 3–1	Run 3–2	Run 3–3	Run 3–4	Run 3–5	Run 3–6	Run 3–7	Run 3–8
Pufeng	1	1	1	1	1	1	1	1
Xinghe	1	1	1	1	1	1	1	1
Yulong	8	8	8	10	8	10	7	11
Anshun	4	5	9	6	1	5	11	5
Hansheng	7	6	10	9	10	9	10	10
Jinghong	5	4	7	7	9	7	9	9
Tongyi	13	13	13	13	13	13	14	14
Taihe	1	1	1	1	1	1	1	1
Puhui	10	10	11	12	11	12	12	13
Yuyang	11	11	14	11	14	11	8	7
Anxin	14	14	1	14	1	14	1	12
Xinghua	6	7	1	4	1	4	1	1
Jintai	9	9	1	8	1	8	1	8
Kaiyuan	12	12	12	5	12	6	13	6

coefficient between σ_j and $\sqrt{HV_j}$ is very high (0.969), VaR (CVaR) is only combined with either σ_j or β (runs 2–5 to 2–8). The obtained results in Table 9 can be analyzed by using the similar method as that for corresponding DEA runs in Table 3, which is thus omitted due to the space limitation. What's more important is that, just like that in the previous subsection, the results here demonstrate the high desirability and importance about the introduction of VaR and CVaR into the DEA performance evaluation models for measuring the overall performance of mutual funds under the normal market circumstance.

Finally, we investigate the impact of different reference data sets and fund collections on the relative performance evaluation of mutual funds. This is accomplished by comparing those corresponding relative rankings in Tables 4, 9 and 10 which are obtained under the same DEA model. The relative rankings of many funds change significantly because of different data sets. For instance, from mostly the average performance in Table 4, *Pufeng* becomes DEA efficient in Tables 9 and 10; although its relative rankings are higher than those of *Pufeng* under most DEA runs in Table 4, *Anshun* is always DEA inefficient except for one under corresponding DEA runs in Tables 9 and 10. Compared next are results obtained with DEA models which simultaneously consider multiple risk measures (runs 2–5 to 2–8 versus runs 3–5 to 3–8). As expected, some funds' relative rankings increase while other funds' corresponding rankings decrease. Two special funds are *Pufeng* and *Taihe*, which are always DEA efficient in two time periods. In contrast, *Yulong* and *Tongyi* change from DEA efficiency in the period 9/1/1999–9/1/2000 to DEA inefficiency in the period 11/1/2000–12/31/2002. Although these two funds' transaction costs in the period 11/1/2000–12/31/2002 are a bit smaller than those in the period 9/1/1999–9/1/2000, their risks increase considerably and expected returns decrease obviously. Therefore, the performances of *Tongyi* and *Yulong* degraded indeed after November 2000. The above analysis shows that the variation of reference sets do affect the performance appraisal of mutual funds.

As a technique for evaluating the relative performance, the DEA efficiency score of each fund is relative to other funds in the selected reference set. Even for the same 14 funds, their inputs and outputs in periods 11/1/2000–12/31/2002 and 9/1/1999–9/1/2000, respectively, would form two different reference sets. Therefore, if a fund's efficiency score in the period 11/1/2000–12/31/2002 is larger than that in the period 9/1/1999–9/1/2000, we could not conclude that this fund's performance is improved. One nature question then is: how can the performances of the same fund in different time periods be fairly compared? To this end, we creatively synthesize two different sample sets of 14 funds in two periods as one new reference set with 28 decision making units. To distinguish the same fund in two periods, the fund name attached by 1 (2) is used to indicate that it corresponds to the time period 9/1/1999–9/1/2000 (11/1/2000–12/31/2002). For instance, *Pufeng*1 corresponds to the fund *Pufeng* considered in 9/1/1999–9/1/2000. Four DEA runs with the same input–output indicator combinations as those in DEA runs 2–5 to 2–8 are examined, which are denoted as DEA runs 4–1 to 4 – 4, respectively. Because of negative expected returns, this indicator is displaced by 0.15 before the derived input-oriented BCC-type linear program is solved. The obtained relative rankings are given in Table 11.

Obviously, funds *Pufeng*1, *Xinghe*1, *Yulong*1, *Tongyi*1 and *Taihe*1 are DEA effcient under all four DEA runs in both Table 9 and Table 11; while being DEA

Table 11 Relative rankings in different DEA runs

Fund name	Run 4–1	Run 4–2	Run 4–3	Run 4–4	Fund name	Run 4–1	Run 4–2	Run 4–3	Run 4–4
*Pufeng*1	1	1	1	1	*Pufeng*2	7	7	7	7
*Xinghe*1	1	1	1	1	*Xinghe*2	12	9	12	9
*Yulong*1	1	1	1	1	*Yulong*2	17	11	17	13
*Anshun*1	6	6	6	6	*Anshun*2	16	17	16	22
*Hansheng*1	25	21	25	25	*Hansheng*2	15	10	15	10
*Jinghong*1	27	22	27	23	*Jinghong*2	9	8	9	8
*Tongyi*1	1	1	1	1	*Tongyi*2	19	23	19	24
*Taihe*1	1	1	1	1	*Taihe*2	21	12	21	15
*Puhui*1	18	24	18	14	*Puhui*2	23	13	23	16
*Yuyang*1	22	27	22	21	*Yuyang*2	26	15	26	19
*Anxin*1	28	26	28	27	*Anxin*2	10	25	10	26
*Xinghua*1	20	28	20	28	*Xinghua*2	11	18	11	18
*Jintai*1	13	16	13	12	*Jintai*2	14	14	14	17
*Kaiyuan*1	8	20	8	11	*Kaiyuan*2	24	19	24	20

efficient under corresponding DEA runs in Table 10, *Pufeng*2, *Xinghe*2 and *Taihe*2 become DEA inefficient in Table 11; the relative rankings of *Hansheng*2, *Jinghong*2, and *Xinghua*2 (*Yulong*2 and *Tongyi*2) are much better (worse) than the corresponding rankings of *Hansheng*1, *Jinghong*1, and *Xinghua*1 (*Yulong*1 and *Tongyi*1), respectively. By using the same analysis method as above, changes in these funds' relative rankings can be explained by variations in their input and output indicators, especially their expected daily returns and values of VaR and CVaR in two time periods. This is omitted due to the limitation of space.

Results in Table 11 further demonstrate the importance of the reference set selection. If we want to objectively evaluate a fund's performance variation during a long time period, the approach introduced above should be a good choice. That is, the whole time period is first divided into several subperiods; each fund is treated as different funds in these subperiods; the input and output indicator values during these subperiods are then combined as a whole reference set and the DEA relative performance appraisal is carried out; the performance variation of each fund can finally be investigated by comparing its relative rankings during different subperiods.

4 Conclusion

The DEA technique often imposes the assumption that all relevant indicators are included in the model. When applied for the mutual fund performance evaluation, this is good and easy to accomplish for taking account of different investment costs, but not the case for describing the fund risk. The risk measures ever considered in DEA models are only traditional measures, which are often highly correlated (especially for σ and $\sqrt{HV}$). Simultaneously including them could cause some degree of redundancy and is thus unnecessary. On the other hand, these measures are incapable of properly reflecting the asymmetry and fat-tailedness of the fund return distribution, which pervasively exist in nowadays actively managed funds. For these reasons, two modern risk measures VaR and CVaR have been introduced into inputs of the DEA model in this paper to measure the mutual fund's risk and to evaluate its performance. Empirical results show that these two measures can indeed better describe return distribution properties and thus risk of mutual funds. Meanwhile, there is no consensus among researchers and managers as to which input and output indicators should be included in a DEA model unambiguously. The sensitivity of DEA efficiencies to various risk measure combinations, especially the combination of VaR and/or CVaR with traditional measures, has thus been empirically investigated in detail. We found that, the proper combination of VaR (CVaR) with other risk measures can more comprehensively reflect mutual funds' risk properties and thus better measure the overall performance of mutual funds.

Different from existing researches, we treat the same fund during different time periods (or different subperiods over a long time period) as ”different” funds, each one corresponds to a specific period (subperiod), and then evaluate their relative performances. This innovation in the field of fund efficiency evaluation is very useful for fairly examining performance variations of a mutual fund.

The above observations not only can provide guidance to analysts in selecting the appropriate input and output indicators when using the DEA technique in mutual fund appraisal or to managers in identifying the source of ineffciencies, but are useful to investors in selecting suitable mutual funds.

Due to the data limitation about the Chinese mutual funds, only different combinations of risk measures have been considered in this paper. It is worthwhile for us to investigate the impact of various combinations of different risk measures and investment costs, as well as different outputs, on the DEA relative performance appraisal by utilizing trading data from other fund markets, which is left for future research.

References

Ang JS, Chua JH (1979) Composite measures for the evaluation of investment performance. J Financ Quant Anal 14:361–384

Artzner P, Delbaen F, Eber JM, Heath D (1999) Coherent measures of risk. Math Financ 9: 203–228

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale efficiencies in data envelopment analysis. Manage Sci 30:1078–1092

Basso A, Funari S (2001) A data envelopment analysis approach to measure the mutual fund performance. Eur J Oper Res 135:477–492

Basso A, Funari S (2002) A generalized performance attribution technique for mutual funds. Working Paper n.01.08, GRETA

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2:429–444

Consiglio A, Massabo I, Ortobelli S (2002) Non-Gaussian distribution for VaR calculation: an assessment for the Italian market. Working paper, University of Calabria

Cooper WW, Seiford LM, Tone K (2000) Data envelopment analysis: a comprehensive text with models, applications, references and DEA solver software. Kluwer, Boston

Duffie D, Pan J (1997) An overview of value at risk. J Deriv 4:7–49

Ferson WE, Schadt RW (1996) Measuring fund strategy and performance in changing economic conditions. J Finance 51:425–461

Jensen MC (1968) The performance of mutual funds in the period 1945–1964. J Finance 23: 389–416
Kogon SM, Williams DB (1998) Characteristic function based estimation of stable parameters. In: Adler R, Feldman R, Taqqu M (eds) A practical guide to heavy tailed data. Birkhauser, Boston
Ma SC, Song LM, Wang W (2000) Empirical study of factors affecting stock prices in the Shanghai stock market. Stat Res 8:24–28 (in Chinese)
McMullen PR, Strong RA (1998) Selection of mutual fund using data envelopment analysis. J Bus Econ Stud 4:1–12
Morey MR, Morey RC (1999) Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. Omega 27:241–258
Morgan JP (1996) RiskMetrics. Technical document, 4th edn. Morgan, New York
Murthi BPS, Choi YK, Desai P (1997) Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach. Eur J Oper Res 98:408–418
Nolan JP (1997) Numerical computation of stable densities and distribution functions. Commun Stat-Stoch Models n. 13
Nolan JP (1999) An algorithm for evaluating stable densities in Zolotarev's parameterization. Math Comput Model n. 29
Pastor J (1996) Translation invariance in data envelopment analysis: a generalization. Ann Oper Res 66:93–102
Rockafellar RT, Uryasev S (2000) Optimization of conditional value-at-risk. J Risk 2:21–42
Schneeweis T, Spurgin R (1998) Multifactor analysis of hedge funds, managed futures, and mutual fund return and risk characteristics. J Altern Invest 1–24 Fall
Sharpe WF (1966) Mutual fund performance. J Bus 34:119–138
Sortino FA, Price LN (1994) Performance measurement in a downside risk framework. J Invest 59–64 Fall
Stephens A, Proffitt D (1991) Performance measurement when return distributions are nonsymmetric. Q J Bus Econ 23–39 Autumn
Treynor JL (1965) How to rate management of investment funds. Harvard Bus Rev 43:63–75