

Performance measurement of hedge funds using data envelopment analysis

Martin Eling

Published online: 15 December 2006
© Swiss Society for Financial Market Research 2006

Abstract Data envelopment analysis (DEA) is a nonparametric method from the area of operations research that measures the relationship of produced outputs to assigned inputs and determines an efficiency score. This efficiency score can be interpreted as a performance measure in investment analysis. Recent literature contains intensive discussion of using DEA to measure the performance of hedge funds, as this approach yields some advantages compared to classic performance measures. This paper extends the current discussion in three aspects. First, we present different DEA models and analyze their suitability for hedge fund performance measurement. Second, we systematize possible inputs and outputs for DEA and again examine their suitability for hedge fund performance measurement. Third, two rules are developed to select inputs and outputs in DEA of hedge funds. Using this framework, we find a completely new ranking of hedge funds compared to classic performance measures and compared to previously proposed DEA applications. Thus, we propose that classic performance measures should be supplemented with DEA based on the suggested rules to fully capture hedge fund risk and return characteristics.

Keywords Data envelopment analysis · Performance measurement · Hedge funds

JEL Classification G10 · G11 · G23

M. Eling (✉)
University of St. Gallen, Institute of Insurance Economics,
Kirchlistrasse 2, 9010 St. Gallen, Switzerland
e-mail: martin.eling@unisg.ch

1 Introduction

As an alternative to traditional investments in stocks or bonds, hedge funds have become very popular and have been the subject of controversial discussion in science as well as in practice in the last years (Favre and Galéano 2001; Banz and de Planta 2002). Hedge fund performance is often evaluated by classic performance measures such as the Sharpe ratio, under which hedge funds appear to be very attractive investments. However, recent research has pointed out some special qualities of hedge fund returns, thus making the suitability of classic measures to assess hedge fund performance less certain (Malkiel and Saha 2005; Eling 2006). In particular, when hedge fund returns are compared to those of traditional investments, they tend to exhibit stronger deviations from normally distributed returns (Borner 2004; Kassberger and Kiesel 2006). Consideration of this problem has led to the development of numerous new performance measures including Omega, the Sortino ratio, the Calmar ratio, and the modified Sharpe ratio. An overview is provided in Lhabitant (2004). These new measures differ from the classic Sharpe ratio in that the standard deviation is replaced by alternative risk measures, such as lower partial moments, drawdown, or value at risk.

Both classic and new approaches to hedge fund performance measurement, however, exhibit one fundamental drawback. As two-dimensional measures, they allow the integration of only one aspect of risk and one aspect of return in performance measurement. The Sharpe ratio, for example, takes into consideration the standard deviation as risk, but has no measure that focuses on the risk of loss, e.g., the lower partial moments. Thus, the use of the Sharpe ratio ignores the tail risk of hedge funds, which is due to the fact that many hedge funds pursue dynamic trading strategies and exhibit option-like returns with unstable exposures to market factors in time (Gregoriou et al. 2005; Fung and Hsieh 1997). Therefore, hedge fund performance is often reported in a multitude of separate measures, as two-dimensional performance measures cannot account for the full complexity of hedge fund returns (Wilkins and Zhu 2005, p. 161; for a typical fund fact sheet see <http://www.hedgefundweb.com> or <http://www.iasg.com>).

In this paper, data envelopment analysis (DEA) is presented as an alternative method for hedge fund performance measurement that does not exhibit the drawback discussed above. DEA is a nonparametric approach that uses linear programming to measure the relationship of produced goods and services (outputs) to assigned resources (inputs). In the hedge fund context, different risk measures might be regarded as inputs and the fund return or related measures as outputs. As an optimization result, DEA determines an efficiency score, which can be interpreted as a performance measure. Originally, DEA was developed to measure efficiency in the public sector, e.g., educational institutions (Banker et al. 1984) or hospitals (Sherman 1984). However, nowadays it has been used in a wide variety of settings, including banks (Sherman and Gold 1985), insurance companies (Fecher and Pestieau 1993), trading companies (Mahajan 1991), and manufacturing companies (Parkan 1991).

Since the end of 1990s, DEA has also been used to measure the performance of financial investments, particularly of mutual funds. The first application of DEA to mutual funds was performed by [Murthi et al. \(1997\)](#), who examined the efficiency of 2,083 mutual funds in 1993. Their motivation to use DEA was to overcome a number of shortcomings of classic two-dimensional performance measures. DEA is very flexible and allows more than one factor as input and output. Thus, compared to classic two-dimensional performance measures, DEA offers a multi-dimensional performance analysis. Compared to other extensions of the classic two-dimensional framework such as higher-moment extensions of the capital asset pricing model presented by [Hwang and Satchell \(1999\)](#) or the generalized version of the capital asset pricing model proposed by [Leland \(1999\)](#), DEA offers further advantages. It is a nonparametric analysis that does not require any theoretical model as a benchmark such as the capital asset pricing model or the arbitrage pricing theory. Instead, DEA measures how well a fund performs relative to the best funds. Furthermore, it can address the problem of endogeneity of transaction costs in the analysis by simultaneously considering expense ratios, turnover, and loads, as well as returns. [Basso and Funari \(2001\)](#) measured the efficiency of 47 mutual funds between 1997 and 1999. Their contribution was to develop a generalized DEA-based performance measure that can integrate both classic performance measures (as Sharpe, Treynor, and Jensen) and the approach of [Murthi et al. \(1997\)](#). Additional applications of DEA for measuring mutual fund performance are [McMullen and Strong \(1998\)](#), [Bowlin \(1998\)](#), [Morey and Morey \(1999\)](#), and [Choi and Murthi \(2001\)](#).

To date, there have been to our knowledge four applications of DEA to hedge funds. [Gregoriou \(2003\)](#), who examined 168 funds of hedge funds for the period from 1997 to 2001, focused on different downside risk measures in order to capture the tail risk in asymmetrically distributed hedge fund returns. The study by [Gregoriou \(2003\)](#) was extended by [Gregoriou et al. \(2005\)](#) in that an additional 446 single hedge funds were examined, using the same inputs and outputs, as well as the same investigation period. [Wilkins and Zhu \(2005\)](#) considered 271 single hedge funds in the June 2001 to May 2002 period. They introduced the returns to scale estimation in DEA of hedge funds. Therefore, DEA not only measures the fund efficiency but also allows subdividing these funds depending on increasing, constant, or decreasing returns to scale. [Nguyen-Thi-Thanh \(2006\)](#), who analyzed 38 hedge funds in the period 2000–2004, discussed the use of DEA in hedge fund portfolio selection when investors face multi-dimensional objectives and incorporated these by adding several optimization constraints. However, in the literature to date, there are only a few DEA applications with various inputs and outputs. What is missing is a common framework of DEA models and discussion of their use in hedge fund performance measurement. In particular, there is no rule or standard for the selection of inputs and outputs when using DEA to evaluate hedge funds.

This study extends previous work on DEA-based hedge fund performance measurement in three aspects. First, three different DEA models and their suitability for hedge fund performance measurement are analyzed. We consider

those three models that have been used to date for performance measurement of mutual funds, which are the DEA models presented by [Charnes et al. \(1978\)](#), [Banker et al. \(1984\)](#), and [Andersen and Petersen \(1993\)](#). Second, 14 inputs and 8 outputs for DEA are systematized and, again, their suitability for hedge fund performance measurement is examined. Third, two rules that can be used to select inputs and outputs in hedge fund performance measurement are developed. The first rule is based on Spearman's rank correlation and the second on principal component analysis. All aspects are first discussed at the theoretical level and then implemented in an empirical investigation.

An important result of the empirical study is that we find completely new rankings of hedge funds compared to classic performance measures and previously proposed DEA applications when using the selection rules.¹ Thus, we are convinced that DEA performance measurement based on these selection rules is able to capture additional information contained in hedge fund returns and leads to an improvement in hedge fund performance measurement. For these purposes, we propose that classic hedge fund performance measurement should be supplemented with DEA based on the suggested rules to fully capture hedge fund risk and return characteristics.

The remainder of the paper is organized as follows. Section 2 contains a discussion of the application of DEA at the theoretical level. The theoretical results are applied in an empirical investigation in Section 3. Section 4 concludes.

2 Theoretical investigation

2.1 DEA models

2.1.1 Mathematical formulation and interpretation of DEA models

In the literature, several DEA models are proposed that differ, e.g., in regard to the underlying returns to scale assumption or concerning the range of the efficiency score. An overview on different models is provided in [Cooper et al. \(2006\)](#). As mentioned, three approaches have been used to date for performance measurement of mutual funds: [Charnes et al. \(1978\)](#), [Banker et al. \(1984\)](#), and [Andersen and Petersen \(1993\)](#). The main features of these models and the differences between them are explored in Table 1.

DEA was introduced by [Charnes et al. \(1978\)](#). DEA builds upon the method suggested by [Farell \(1957\)](#) for computation of the technical efficiency. [Farell \(1957\)](#) developed the concept of "best practice frontier," which represents the technical frontier achievable, i.e., it shows for each input level the maximum attainable output and for each output level the minimum necessary input,

¹ This result is of special relevance because [Eling and Schuhmacher \(2006\)](#) found that the two-dimensional measures recently proposed in the literature (e.g., Omega) do not change rankings of hedge funds too much.

Table 1 DEA models

Model	Charnes et al. (1978)	Banker et al. (1984)	Andersen and Petersen (1993)
Assumption concerning returns to scale	Constant	Variable	Variable
Range of efficiency score	Between 0 and 1	Between 0 and 1	Starts with 0, but larger than 1 is possible

This table shows two key characteristics (the assumption concerning returns to scale and the range of the efficiency score) for those DEA models that have been used to date for performance measurement of mutual funds

respectively. The efficiency of a fund can then be determined by the relative distance between the actually observed output and this efficient frontier. Thus, a fund is classified as inefficient if its output (e.g., return) and input (e.g., risk) are below the best practice frontier.

Charnes et al. (1978) presented a solution algorithm for the problem posed by Farell (1957) based on linear programming. By assuming constant returns to scale, possible economies of scale are ignored in the model. The starting point of the model is the formula for calculating the efficiency score e_i of fund i ($i = 1, \dots, I$). This formula gives the relationship of the weighted sum of the outputs to the weighted sum of the inputs:

$$e_i = \frac{\sum_{r=1}^R u_r y_{ri}}{\sum_{j=1}^J v_j x_{ji}}, \tag{1}$$

with $r = 1, \dots, R$ as outputs, $j = 1, \dots, J$ as inputs, y_{ri} as amount of output r of fund i , x_{ji} as amount of input j of fund i , u_r as output-weighting factor, and v_j as input-weighting factor. In the optimization, weighting factors are selected so that the efficiency score is maximized, subject to technical constraints. The efficiency measure is standardized between 0 and 1, i.e., the most efficient fund receives the value 1; completely inefficient funds have the value 0. Thus, the following nonlinear optimization problem results:

$$\begin{aligned} \text{Maximize}_{u_r, v_j} e_i &= \frac{\sum_{r=1}^R u_r y_{ri}}{\sum_{j=1}^J v_j x_{ji}}, \quad \text{subject to } \frac{\sum_{r=1}^R u_r y_{ri}}{\sum_{j=1}^J v_j x_{ji}} \\ &\sum_{j=1}^J v_j x_{ji} \leq 1, \quad u_r \geq 0, \quad v_j \geq 0. \end{aligned} \tag{2}$$

By setting the denominator of the objective function to 1, this nonlinear optimization problem can be transformed into an equivalent linear optimization

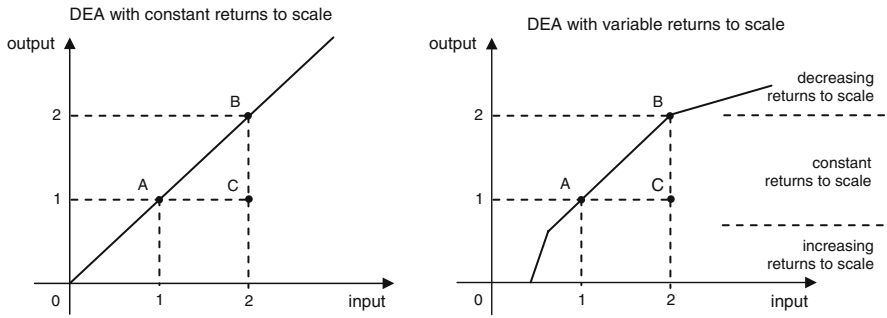


Fig. 1 DEA with constant and variable returns to scale. This figure shows the DEA efficient frontier with constant returns to scale (*left*) and with variable returns to scale (*right*) in the case where one input is used for the production of one output

problem:

$$\begin{aligned}
 &\text{Maximize}_{u_r} \sum_{r=1}^R u_r y_{ri}, \quad \text{subject to } \sum_{j=1}^J v_j x_{ji} = 1, \\
 &\sum_{r=1}^R u_r y_{ri} - \sum_{j=1}^J v_j x_{ji} \leq 0, \quad u_r \geq 0, v_j \geq 0.
 \end{aligned} \tag{3}$$

The determined efficiency score can be interpreted as savings potential. For example, a value of 0.8 means that with efficient production the same output quantity could be achieved with a 20% smaller input. The weighting factors u_r and v_j , which are the shadow prices in the optimization, also provide powerful economic insights. The ratio of the shadow prices corresponds to (1) the marginal rate of substitution (if the shadow prices of two inputs are compared), (2) the marginal productivity (if the shadow prices of one input and one output are compared), and (3) the marginal rate of transformation (if the shadow prices of two outputs are compared).² Further economic insights are provided by the slack variables in the optimization, as they indicate the extent to which each input can be reduced to achieve an efficiency score of one (Choi and Murthi 2001; Kuosmanen et al. 2006). Thus, DEA not only measures efficiency, but can provide guidance as to how to improve the efficiency of inefficient funds.

The left part of Fig. 1 clarifies the principle of DEA with constant returns to scale in the case where one input is used for the production of one output.

The straight line ascending from the origin in the left part of Fig. 1 is called the efficient frontier. The points A, B, and C represent observed input–output combinations of different funds. Funds A and B exhibit the best input–output combinations and thus determine the efficient frontier. Fund C does not work

² See Steinmann (2002) and Murthi et al. (1997). Take the left part of Fig. 1 as an example: The slope of the straight line ascending from the origin (the efficient frontier) represents the marginal productivity and can be calculated as u_r divided by v_j .

efficiently; the output it obtains could be achieved with a lower input fund A employs, and the inputs of fund C are sufficient to achieve the output of fund B. The case of fund C is a good illustration of the efficiency score calculation. The efficiency score results from the relationship of the inputs necessary for efficient production (which is 1 in this case) to the actually used inputs (2 in this case). In this case the efficiency score is 0.5.

If size differences have an influence on fund efficiency, assuming constant returns to scale would lead to a blending of scale efficiency due to size differences and technical efficiency. By employing an additional variable c_i , the [Banker et al. \(1984\)](#) approach introduces variable returns to scale:

$$\begin{aligned} \text{Maximize } & \sum_{r=1}^R u_r y_{ri} + c_i, \quad \text{subject to } \sum_{j=1}^J v_j x_{ji} = 1, \quad \sum_{r=1}^R u_r y_{ri} \\ & - \sum_{j=1}^J v_j x_{ji} - c_i \leq 0, \quad u_r \geq 0, v_j \geq 0. \end{aligned} \tag{4}$$

The addition of variable returns to scale is the only difference between this model and that of [Charnes et al. \(1978\)](#). The right part of Fig. 1 shows the principle of DEA with variable returns to scale, again using one input and one output factor. Assuming variable returns to scale, the efficient frontier ascending from the origin in the figure runs piecewise linear. Three ranges can be distinguished within the efficient frontier—increasing ($c_i < 0$), constant ($c_i = 0$), and decreasing ($c_i > 0$) returns to scale. With increasing (decreasing) returns to scale, an increase of input leads to a disproportionately high (low) increase of output. With constant returns to scale, the increase of an input corresponds to the increase of an output.

The approach of super efficiency presented by [Andersen and Petersen \(1993\)](#) is similar to the [Banker et al. \(1984\)](#) approach. However, in the optimization, the fund under consideration is excluded from the calculation of the constraints. Therefore, efficient funds can attain an efficiency score larger than 1. This procedure makes a ranking of efficient funds possible. For inefficient funds, the efficiency score measured by [Andersen and Petersen \(1993\)](#) model corresponds to the value calculated by [Banker et al. \(1984\)](#) model.

2.1.2 Application of DEA models to hedge funds

Regarding the application of DEA for hedge fund performance measurement, we first discuss the use of constant and variable returns to scale. Especially with hedge funds, the optimal fund size is an important problem against the background of different scale effects moving in opposite directions.

On the one hand, hedge funds need a certain minimum capital to implement their strategies. For example, with the fixed-income arbitrage strategy very small market inefficiencies can be profitably exploited only if a large amount of capital is available. Moreover, fixed-cost degression results in larger funds

working more efficiently. These two effects indicate that hedge funds involve scale advantages and increasing returns to scale.

On the other hand, hedge funds that are too large experience scale disadvantages. First, large hedge funds can create the risk that the funds will affect market prices by their own trading actions. This is particularly problematic with illiquid securities because it is possible that closing the position will not be feasible at fair market prices. Second, finding profitable investment opportunities becomes more difficult with increasing fund size, a situation particularly true for funds specialized in a very close market segment. Thus, hedge funds are also subject to scale disadvantages and decreasing returns to scale.

It is unknown how the scale advantages and disadvantages work in sum. The above discussion implies, however, that small funds may obtain increasing returns to scale at first, but after attaining a certain size become subjected to decreasing returns to scale. Therefore, assuming variable returns to scale seems appropriate in applying DEA to hedge funds, which is in accordance with empirical evidence³ and the procedures employed by [Gregoriou \(2003\)](#); [Wilkens and Zhu \(2005\)](#), and [Gregoriou et al. \(2005\)](#). Nevertheless, DEA applications to traditional mutual funds usually use constant returns to scale ([Murthi et al. 1997](#), p. 411; [Basso and Funari 2001](#), p. 481), even though most of the scale effects discussed with respect to hedge funds also affect traditional mutual funds.⁴ In order to evaluate the impact of model selection, we include applications with constant and variable returns to scale in our empirical analysis.

2.2 Systematization of inputs and outputs

The selection of input and output factors is of high importance when using DEA. A basic result of the capital market theory is that there exists a functional relationship between risk and return of an investment—higher risk taking is rewarded with a higher return. Thus, in principle, risk measures are thought of as inputs, while return measures are interpreted as outputs. However, [Nguyen-Thi-Thanh \(2006\)](#) argues that some investors might be more concerned with central tendencies (mean, standard deviation), while others may care more about extreme values (skewness, kurtosis). Moreover, [McMullen and Strong \(1998\)](#) show that investors are concerned about risk and return over various time-horizons, as this provides more information than the performance of a single period only. [Basso and Funari \(2001\)](#) suggest that some investors may

³ [Ammann and Moerth \(2005\)](#) found a negative relationship between fund sizes and returns, but they also show that very small funds underperform on average. The underperformance of very small funds is explained by their higher total expense ratios. A negative relationship between standard deviations and fund sizes is also observed. Thus, we conclude that both risk and return are subject to variable returns to scale with hedge funds. Attributing the underperformance of very small funds to higher total expense ratios indicates that fund costs are also subject to variable returns to scale, which is in line with our discussion. Therefore, we conclude that costs are also subject to variable returns to scale with hedge funds.

⁴ For example, [Ferris and Chance \(1987\)](#) find a negative relationship between the expense ratio and the size of mutual funds.

Table 2 Inputs and outputs considered in the literature

Category	Gregoriou (2003), Gregoriou et al. (2005)	Wilkins and Zhu (2005)	Nguyen-Thi-Thanh (2006)
Inputs	Lower partial moments of order 1	Standard deviation	Standard deviation
	Lower partial moments of order 2	Lower partial moments of order 0	Excess kurtosis
	Lower partial moments of order 3		
Outputs	Higher partial moments of order 1	Average return	Average return
	Higher partial moments of order 2	Skewness	Skewness
	Higher partial moments of order 3	Minimum return	

This table shows the inputs and outputs used in the four papers that applied DEA to measure the performance of hedge funds. Gregoriou (2003) and Gregoriou et al. (2005) choose lower and higher partial moments; Wilkins and Zhu (2005) and Nguyen-Thi-Thanh (2006) also include the average return, the standard deviation of returns, and the higher moments of the return distribution in the DEA of hedge funds

also include ethical criteria in decision making, and Murthi et al. (1997) report that investors are interested in transaction costs. In the hedge fund context, it might also be important to consider hedge-fund-specific characteristics such as the minimum investment or the lock-up period (Nguyen-Thi-Thanh 2006). Therefore, the selection of input and output variables is a complex task as the decision-maker is confronted with a large number of different attributes that might be of differing importance. However, by the choice of inputs and outputs and by setting additional constraints in the optimization, DEA is able to include all the investor's financial, risk-aversion, diversification, and investment horizon constraints. Consequently, DEA can be a tailor-made decision-making tool that exactly suits the investor's preferences (Nguyen-Thi-Thanh 2006).

In this section possible inputs and outputs for DEA are presented, systematized, and discussed with respect to their suitability for hedge fund performance measurement. Table 2 shows the inputs and outputs suggested in Gregoriou (2003), Gregoriou et al. (2005), Wilkins and Zhu (2005), and Nguyen-Thi-Thanh (2006).

To capture the tail characteristics of hedge fund returns, Gregoriou (2003) and Gregoriou et al. (2005) use the lower partial moments (LPM) of order 1–3 as inputs and the higher partial moments (HPM) of order 1–3 as outputs. Wilkins and Zhu (2005) choose some measures often employed in hedge fund analysis. They consider the standard deviation and the portion of negative returns (which is the lower partial moment of order 0) as inputs and the average return, the skewness, and the minimum return as outputs. Because investors might have positive preferences for odd moments and negative preferences for even moments Nguyen-Thi-Thanh (2006) chooses standard deviation and excess kurtosis as inputs and average return and skewness as outputs. It

Table 3 Possible inputs and outputs

Inputs			
risk measures	standard deviation (SD)	} total risk } downside risk } systematic risk } unsystematic risk	} included in the empirical part } not included in the empirical part
	lower partial moments of order 0 (LPM ₀)		
	lower partial moments of order 1 (LPM ₁)		
	lower partial moments of order 2 (LPM ₂)		
	lower partial moments of order 3 (LPM ₃)		
	maximum drawdown (MD)		
	average drawdown (AD)		
	standard deviation of drawdown (SDD)		
	value at risk (VAR)		
	conditional value at risk (CVAR)		
	modified value at risk (MVAR)		
	beta factor		
	residual volatility		
	tracking error		
excess kurtosis			
loads, fees			
trading activity			
higher moments			
cost measures			
outputs			
return measures	arithmetic excess return (AER)	} included in the empirical part } not included in the empirical part	} included in the empirical part } not included in the empirical part
	geometric excess return (GER)		
	higher partial moment of order 0 (HPM ₀)		
	higher partial moment of order 1 (HPM ₁)		
	higher partial moment of order 2 (HPM ₂)		
	higher partial moment of order 3 (HPM ₃)		
higher moments	Skewness		→ the empirical part

This table shows a selection of the most common risk, return, and cost measures that might be regarded as inputs and outputs in DEA. Inputs include risk measures, cost measures, and higher moments. Outputs include return measures and higher moments. Eleven inputs and six outputs will be used in the empirical section of this paper. The formulas for calculating the measures included in the empirical investigation can be found in Appendix A

seems fair to conclude from the literature that there is no agreement as to which measures should be used as inputs and outputs for a DEA-based performance measurement of hedge funds.⁵

Table 3 sets out a selection of the most common risk, return, and cost measures that might be regarded as inputs and outputs in DEA. The measures presented do not necessarily comprise a complete list of all possible risk, return, and cost measures, as individual investors might consider other measures as relevant.

There is a wide variety of classic and newer risk measures that can be included as inputs. The standard deviation measures the total risk of an investment, which gives both the positive and negative deviations of the returns from the expected value. However, this is not the general understanding of risk. In contrast, LPMs only consider negative deviation of returns from a minimal acceptable return, a situation that most investors view with disfavor. Thus, LPMs might seem a more appropriate measure of risk. The choice of order *n* determines the extent to which the deviation from the minimal acceptable return is weighted.

The drawdown of a fund is the loss incurred over a given investment period. The three drawdown-based risk measures listed in Table 3 are defined as the

⁵ The difficulty in selecting inputs and outputs applies not only to hedge funds, but more so to banks. For a discussion of inputs and outputs of banks see Bessler and Norsworthy (2005).

largest possible loss (the so-called maximum drawdown), an average of the largest drawdowns, or a type of standard deviation of the largest drawdowns.

Value at risk describes the possible loss of an investment, which is not exceeded with a given probability of $1 - \alpha$ in a certain period. To take into account the distribution of returns below the value at risk, the literature frequently considers expected loss under the condition that the value at risk is exceeded (Artzner et al. 1999). This leads to the conditional value at risk. To include skewness and excess kurtosis in computing value at risk, the Cornish–Fisher expansion can be used, which leads to the modified value at risk (Gregoriou and Gueyie 2003).

Measures for systematic (beta factor) and unsystematic risk (residual volatility and tracking error) could be integrated into DEA. But these measures are less suitable for performance measurement of hedge funds in our context because they require the definition of a benchmark. Usually, hedge funds attempt to achieve positive returns independent of any market index and therefore the definition of a standardized benchmark is neither possible nor meaningful. Therefore, these risk measures are excluded from the empirical investigation.⁶

The cost measures that are used as inputs in conducting DEA for mutual funds are likewise less suitable for performance measurement of hedge funds, as front-end loads, administration, and management fees are already contained in the fund returns (all hedge fund returns are net of fees). It is important to note that one could argue that even if returns are net of all fees, fee consideration can still provide additional information about the fund manager's performance because there is a difference between a fund that has a gross return of 10% and a net return of 5%, and a fund that has a gross return of 20% and a net return of 5% (Choi and Murthi 2001). But even if information on returns before fees was available, a comparison of hedge funds on the basis of cost measures would not be meaningful due to the complexity and heterogeneity of the fee structure (Le Moigne and Savaria 2006). An example is the individual high watermark. Thus, standardized comparisons as used in evaluating mutual funds, e.g., with front-end loads, are not possible for hedge funds. The same is true for trading-activity-based measures, also included in mutual fund DEA. Because hedge fund managers do not publish their trading activity, there is no data available to make such a comparison between different hedge funds.

⁶ From a theoretical point of view, the beta factor is not suited if the performance of hedge funds is evaluated in isolation but it is the right measure if the investor only wants to add a small percentage of hedge funds to his existing investment portfolio. See Scholz and Wilkens (2003) for a related discussion. Therefore, it would be necessary to measure the systematic risk contribution of the fund under evaluation in relation to the investor's portfolio. As this goes far beyond the scope of this contribution, we restrict ourselves to the evaluation of hedge funds in isolation and exclude this portfolio theoretical investigation. However, in Appendix B we show that our main results are robust for an evaluation from the portfolio perspective. For this purpose, we measure the performance of a portfolio of traditional investments and a small portion of hedge funds (1, 5, and 10%). Detailed results are available upon request.

Compared to the input side, the number of possible return measures on the output side is small. In most of the literature, only excess return is used as an output, making it necessary merely to decide whether the average return should be computed arithmetically or geometrically. The arithmetic case assumes withdrawal of gains; the geometric case assumes reinvestment of gains. [Gregoriou \(2003\)](#) and [Gregoriou et al. \(2005\)](#) also use a counterpart of the LPMs as return measure, that is, the higher partial moments (HPM).

As mentioned, [Wilkins and Zhu \(2005\)](#) and [Nguyen-Thi-Thanh \(2006\)](#) use skewness and kurtosis in DEA as inputs and outputs. Similar to hedge fund costs, we will not integrate skewness or kurtosis in DEA because these characteristics are already displayed by many risk measures on the input side. Skewness and kurtosis, for example, explicitly enter the computation of the modified value at risk so that additional consideration of the higher moments is not necessary in our framework.

2.3 Input and output selection rules

The question now arise which and how many of these input and output measures should be used for a DEA of hedge funds. Counterintuitively, using too many inputs and outputs will be less helpful because when the number of inputs and outputs increases, more decision-making units tend to get an efficiency score of 1 as they become too specialized to be evaluated with respect to other units.⁷ A rule of thumb is that there should be a minimum of three funds per input and output in implementing a DEA model ([Bowlin 1998](#), p. 18). Thus, for practical reasons, there needs to be some limit on the number of inputs and outputs.

As mentioned, to date there has been no agreement or rule for the selection of inputs and outputs. From an economic point of view, the input and output selection should be based on the investor's preferences because different investors might consider different measures as relevant. The correct approach would thus be to evaluate the distribution function of the returns of a particular hedge fund by using the investor's utility function.⁸ The problem with this procedure is that hedge fund performance is often reported in a multitude of separate measures, as single measures cannot account for the full complexity of hedge fund returns. The investor is confronted with a large number of possible measures, any or all of which might be of particular relevance.

⁷ Problems related to discrimination and model misspecification arise when there are a relatively large number of variables compared to the number of funds, which may cause the majority of funds to be defined as efficient. Thus, the inclusion of more inputs and outputs in the model can artificially inflate the efficiency scores as the addition of each variable provides an additional dimension in which the model will seek similar comparison peers. See [Pedraja-Chaparro et al. \(1999\)](#) and [Adler and Golany \(2002\)](#).

⁸ There are different ways to integrate investor's preferences into DEA, see [Allen et al. \(1997\)](#) and [Halme et al. \(1999\)](#). A simple example would be a scoring approach similar to the proposed selection rule based on principal components; the only difference is that the input and output weights are not determined by an optimization but are estimated by the investor.

We recommend tackling this problem by using DEA, which is capable of integrating the complexity of hedge fund returns into performance measurement. For this purpose, those inputs and outputs should be selected that differ as much as possible from each other in order to reach the highest possible explanatory power in performance measurement.⁹ We propose two selection rules that meet this criterion. The first is based on Spearman's rank correlation; the second on principal component analysis.¹⁰

2.3.1 Selection rule based on Spearman's rank correlation

The selection rule based on Spearman (1904) consists of selecting those inputs and outputs that are correlated as little as possible. Here the selection of inputs and outputs is a three-step procedure: in the first step, all risk and return measures are computed for all funds. In the second step, the measurement values are ranked so as to determine an order of the funds. In the third step, this ranking is used to determine the rank correlation between the measures; those measures with the smallest rank correlations are selected as inputs and outputs. The selection rule based on Spearman's rank correlation coefficient (r_{sp}) is stated formally for the inputs as

$$\arg \min_{1 \leq j_1, j_2 \leq J} r_{sp} j_1, j_2 = 1 - \left(6 \sum_{i=1}^I d_i^2 \right) / (I^3 - I), \quad (5)$$

and for the outputs as

$$\arg \min_{1 \leq r_1, r_2 \leq R} r_{sp} r_1, r_2 = 1 - \left(6 \sum_{i=1}^I d_i^2 \right) / (I^3 - I), \quad (6)$$

with d_i as differences in ranking places for fund i . Under this selection rule, two inputs and two outputs are selected, which are those with the smallest rank correlation. (For more details on rank correlation analysis see Kendall and Gibbons 1990).

⁹ If the investor already knows which measures are relevant and if he knows the relative importance of each of these measures, there is no further selection rule necessary. Thus, the selection rules presented in the following are helpful (1) for those investors who cannot decide between the large number of possible measures, (2) for those investors who have chosen too many inputs and outputs, and (3) for those investors who cannot identify the relative importance of the inputs and outputs chosen.

¹⁰ The main motivation for the proposed selection rules is that they are helpful in achieving a high-quality DEA model. The DEA model could be misspecified if relevant factors are omitted or if irrelevant factors are included (Smith 1997). For example, an input that has a correlation of one to another input might be considered irrelevant for the DEA analysis, as it cannot provide any new information (Pedraja-Chaparro et al. 1999). The presented selection rules avoid these problems by distinguishing relevant from irrelevant factors and thus improve the model quality.

2.3.2 Selection rule based on principal component analysis

Principal component analysis is a variable reduction procedure. It extracts the most important information by building a number of artificial variables that account for most of the variance in the observed variables. These artificial variables are called principal components. According to the second selection rule, the principal components should be used in DEA to integrate the variability of all risk and return measures (Adler and Golany 2002). Technically, a principal component is a linear combination of optimally-weighted observed variables. Starting with the observed variables X_1 to X_N , principal component analysis provides the following linear system of equations:

$$\begin{aligned} C_A &= \gamma_{1A} \cdot X_1 + \gamma_{2A} \cdot X_2 \cdots + \gamma_{NA} \cdot X_N, \\ C_B &= \gamma_{1B} \cdot X_1 + \gamma_{2B} \cdot X_2 \cdots + \gamma_{NB} \cdot X_N, \cdots \end{aligned} \quad (7)$$

The first principal component C_A is the result of a linear combination of the observed variables X_1 to X_N with the factor loadings γ_{1A} to γ_{NA} . In the optimization, the factor loadings are weighted so as to maximize the portion of variance explained by the principal components. The eigenvector with the largest eigenvalue determines the first principal component, while the eigenvector with the second largest eigenvalue determines the second principal component (for more details on principal component analysis see Jackson 2003; Jolliffe 2002). The practical implementation of both selection rules is shown in the following section.

3 Empirical investigation

3.1 Data and procedure

In the empirical investigation we consider monthly returns for 30 hedge funds between January 1996 and December 2005. The hedge funds were selected from the Center for International Securities and Derivatives Markets (CISDM) hedge fund database.¹¹ To obtain a representative sample of the entire hedge fund universe, we randomly choose two funds from each of the 15 strategy groups included in the database.¹² The performance of these 30 funds is evaluated by 10 DEA applications. The results of the applications are compared

¹¹ The CISDM hedge fund database has been subject of many academic studies. For the properties of this database, see, e.g., (Edwards and Caglayan 2001; Kouwenberg 2003; Capocci and Hübner 2004).

¹² Altogether we analyzed 450 hedge funds. We repeated the investigation 15 times with 30 funds from each of the strategy groups and found that our main results are robust with regard to these variations. We also found robust results regarding different time-horizons. The results of all robustness tests are presented in Appendix B. Detailed results are available upon request.

Table 4 Overview of the applications

Application	input	Output	returns to scale	underlying modeling approach
1	standard deviation	excess return	constant	Charnes/Cooper/Rhodes (1978)
2	standard deviation	excess return	variable	Banker/Charnes/Cooper (1984)
3	LPM ₁ , LPM ₂ , LPM ₃	HPM ₁ , HPM ₂ , HPM ₃	variable	Banker/Charnes/Cooper (1984)
4	standard deviation, LPM ₀	average return, skewness, minimum return	variable	Banker/Charnes/Cooper (1984)
5	} selection rule based on	selection rule based on	constant	Charnes/Cooper/Rhodes (1978)
6		Spearman's rank correlation	variable	Banker/Charnes/Cooper (1984)
7	} selection rule based on	selection rule based on	variable	Andersen/Petersen (1993)
8		principal component analysis	constant	Charnes/Cooper/Rhodes (1978)
9		principal component analysis	variable	Banker/Charnes/Cooper (1984)
10	} selection rule based on	selection rule based on	variable	Banker/Charnes/Cooper (1984)
		principal component analysis	variable	Andersen/Petersen (1993)

This table shows the inputs, outputs, returns to scale assumption, and underlying modeling approach for the ten DEA applications considered in the empirical investigation. Applications 1–4 are already covered in the literature. Applications 5–10 extend the literature by choosing inputs and outputs following the selection rules presented in Sect. 2.3. The selection of inputs and outputs is based on Spearman's rank correlation for Applications 5–7 and on principal component analysis for Applications 8–10. The abbreviations are explained in Table 3

both among themselves and with classic performance measures. Table 4 gives an overview of the ten applications.

Application 1 clarifies the connection between classic performance measurement and DEA. Its input is standard deviation and its output is excess return. Assuming constant returns to scale, Application 1 results in the same ranking and evaluation of the investments as is made by the Sharpe ratio (SR_i), calculated as the arithmetic excess return divided by the standard deviation. In Application 2, the assumption of constant returns to scale is replaced by variable returns to scale. Applications 3–10 use a multi-dimensional performance measurement instead of a two-dimensional approach. Application 3 is the one presented by Gregoriou (2003) and Gregoriou et al. (2005). Application 4 was analyzed by Wilkens and Zhu (2005).

Applications 5–10 extend the literature by choosing inputs and outputs following the selection rules presented in Sect. 2.3. Applications 5–7 use the selection rule based on Spearman's rank correlation, but use various underlying DEA models: Application 5 is based on Charnes et al. (1978), Application 6 on Banker et al. (1984), and Application 7 on Andersen and Petersen (1993). Applications 8–10 use the selection rule based on principal component analysis. Again, these applications differ with respect to the underlying DEA model.

Various comparisons between the applications are possible and meaningful. The results of the classic Sharpe ratio can be compared with the results of the DEA analysis. Applications 1 and 2 allow a direct comparison of the Charnes et al. (1978) and Banker et al. (1984) models. This clarifies the influence of variable economies of scale on fund efficiency. Moreover, applications already proposed in the literature (Applications 3 and 4) can be contrasted with applications using the new selection rules (Applications 5–10). Section 3.2 presents the selection of inputs and outputs. Then, in Sect. 3.3, DEA performance measurement is performed. To solve the DEA optimization problem, the efficiency measurement system (EMS) software presented by Scheel (2000) is employed.

3.2 Selection of inputs and outputs

3.2.1 Selection rule based on Spearman's rank correlation

Following the selection rule based on Spearman's rank correlation, we first calculate all measures for the 30 funds. A constant risk-free interest rate of $r_f = 0.35\%$ per month was used. This corresponds to the interest on 10-year US treasury bonds as of 30 December 2005 (4.28% per annum). For the LPM-based measures, the minimal acceptable return corresponds to the risk-free monthly interest rate ($\tau = 0.35\%$). For the average drawdown and the standard deviation of drawdown, the five largest drawdowns are considered ($K = 5$). The value-at-risk-based performance measures were calculated using a significance level of $\alpha = 0.05$. The results for the inputs are presented in Table 5 (panel A). In the parentheses, the measurement values are ranked in order to determine the rank correlation in the next step.

The rankings presented in panel A of Table 5 illustrate the complexity of hedge fund performance measurement. For example, fund 27 has the highest standard deviation, but it ranks at tenth place with the modified value at risk. Fund 24 is in second place with the standard deviation, but at 21st place with the modified value at risk. These results confirm the significance of a multi-dimensional approach when assessing hedge fund performance and thus emphasize the benefit of DEA in performance measurement.

Using the ranking shown in Table 5, the rank correlation between the measures is derived, as set out in panel B of Table 5. The measures that exhibit the smallest rank correlations are selected as inputs and outputs. We find the minimum rank correlation comparing conditional value at risk and lower partial moment of order 0. The rank correlation here is 0.23.

The measurement values and rankings for the outputs are presented in Table 6. The minimum rank correlation is found with the higher partial moment of order 3 and higher partial moment of order 0. The rank correlation here is -0.33 .

Thus, in Applications 5–7 the conditional value at risk and the lower partial moment of order 0 will be used as inputs and the higher partial moment of order 3 and of order 0 will be used as outputs.

3.2.2 Selection rule based on principal component analysis

Principal components are calculated using SPSS. The results obtained for the inputs and outputs are shown in Table 7.

On the input side, the lower partial moment of order 2 has the highest impact on component A (factor loading 0.97), while the lower partial moment of order 0 is the most important factor for component B (factor loading 0.86). Components A and B together account for 88.20% of the variance in the input data. On the output side, the higher partial moment of order 2 (factor loading 0.95, component A) and the higher partial moment of order 0 (factor loading

Table 5 Measurement value and ranking of inputs

Panel A: measurement value and ranking (in parentheses) of inputs

Input fund	SD	LPM ₀	LPM ₁	LPM ₂	LPM ₃	MD	AD	SDD	VAR	CVAR	MVAR
1	1.50 (26)	28.33 (29)	0.37 (23)	0.01 (21)	0.00 (20)	8.52(23)	5.51 (20)	12.95 (21)	1.66 (26)	2.98 (24)	1.94 (20)
2	1.29 (27)	34.17 (22)	0.35 (24)	0.01 (24)	0.00 (25)	9.34(22)	4.90 (21)	13.52 (20)	1.49 (27)	2.27 (27)	1.58 (22)
3	1.63 (23)	34.78 (21)	0.34 (25)	0.00 (27)	0.00 (27)	6.16(26)	2.73 (27)	7.62 (27)	1.78 (24)	2.25 (28)	1.24 (25)
4	0.98 (29)	46.27 (5)	0.32 (27)	0.00 (29)	0.00 (29)	3.23(29)	1.88 (28)	4.63 (29)	1.09 (29)	1.44 (30)	1.05 (26)
5	3.89 (10)	36.21 (19)	0.94 (10)	0.07 (8)	0.01 (7)	22.85(11)	12.25 (10)	30.55 (11)	5.03 (10)	8.90 (10)	6.09 (7)
6	4.71 (8)	37.38 (17)	1.30 (8)	0.11 (6)	0.02 (2)	46.84(2)	14.20 (6)	48.75 (3)	6.91 (8)	13.14 (5)	7.79 (5)
7	6.92 (3)	51.28 (3)	1.53 (6)	0.15 (2)	0.03 (1)	48.71 (1)	16.83 (4)	52.48 (2)	10.64 (2)	15.94 (2)	11.43 (1)
8	5.68 (5)	41.18 (15)	1.70 (4)	0.13 (5)	0.02 (4)	32.10(6)	18.13 (3)	47.62 (4)	7.94 (5)	11.64 (6)	8.32 (3)
9	2.86 (13)	33.33 (23)	0.72 (14)	0.02 (15)	0.00 (15)	18.83(14)	7.91 (15)	21.76 (16)	3.48 (15)	4.35 (18)	3.46 (14)
10	5.22 (7)	40.83 (16)	1.73 (3)	0.13 (4)	0.01 (6)	38.77(4)	12.79 (8)	40.84 (6)	7.88 (6)	11.46 (7)	8.13 (4)
11	6.83 (4)	42.37 (13)	1.93 (1)	0.15 (1)	0.02 (3)	19.93(13)	18.18 (2)	40.77 (7)	9.48 (4)	13.20 (4)	9.43 (2)
12	2.81 (14)	33.33 (23)	0.79 (12)	0.04 (13)	0.00 (13)	18.37(15)	8.59 (14)	24.64 (14)	3.64 (14)	6.38 (14)	4.23 (12)
13	1.71 (22)	44.17 (8)	0.52 (18)	0.01 (20)	0.00 (21)	5.99(27)	4.65 (22)	10.67 (24)	2.20 (22)	3.14 (22)	2.05 (19)
14	2.96 (12)	52.50 (2)	1.20 (9)	0.05 (10)	0.00 (12)	43.14(3)	12.36 (9)	44.42 (5)	4.66 (11)	6.76 (13)	4.88 (9)
15	0.99 (28)	35.83 (20)	0.27 (29)	0.01 (25)	0.00 (23)	11.79(20)	3.67 (23)	12.28 (23)	1.16 (28)	3.01 (23)	1.52 (23)
16	2.36 (20)	41.67 (14)	0.65 (15)	0.02 (17)	0.00 (17)	12.02(19)	6.84 (16)	16.39 (18)	3.17 (19)	4.88 (16)	2.81 (17)
17	2.56 (15)	33.33 (23)	0.41 (21)	0.04 (12)	0.01 (8)	24.97(9)	5.52 (19)	25.06 (13)	3.77 (13)	21.53 (1)	3.48 (13)
18	1.61 (24)	30.67 (27)	0.28 (28)	0.00 (28)	0.00 (28)	4.80(28)	1.83 (29)	5.72 (28)	1.70 (25)	2.63 (26)	0.86 (27)
19	2.44 (19)	43.48 (10)	0.57 (16)	0.02 (18)	0.00 (16)	14.14(17)	5.97 (18)	18.01 (17)	3.26 (17)	5.35 (15)	3.12 (15)
20	2.54 (18)	20.83 (30)	0.50 (19)	0.04 (11)	0.01 (10)	30.75(7)	8.67 (12)	31.43 (10)	3.32 (16)	9.46 (8)	4.28 (11)

Table 5 Continued

Panel A: measurement value and ranking (in parentheses) of inputs											
Input fund	SD	LPM ₀	LPM ₁	LPM ₂	LPM ₃	MD	AD	SDD	VAR	CVAR	MVAR
21	2.56 (16)	32.00(26)	0.34 (26)	0.01 (23)	0.00 (24)	6.59(24)	3.07(25)	8.32(26)	3.06(20)	3.57(21)	-2.12 (30)
22	0.88 (30)	29.17(28)	0.12 (30)	0.00 (30)	0.00 (30)	2.42(30)	1.04(30)	2.89(30)	0.67(30)	1.63(29)	0.36 (29)
23	3.07 (11)	42.50(12)	0.77 (13)	0.02 (16)	0.00 (18)	23.08(10)	8.65(13)	26.68(12)	3.97(12)	4.36(17)	2.87 (16)
24	7.21 (2)	49.23(4)	0.82 (11)	0.03 (14)	0.00 (14)	16.30(16)	8.90(11)	23.85(15)	10.01(3)	7.40(12)	1.64 (21)
25	1.57 (25)	44.17(8)	0.38 (22)	0.01 (26)	0.00 (26)	6.32(25)	3.32(24)	8.46(25)	1.92(23)	2.77(25)	1.27 (24)
26	5.53 (6)	42.59(11)	1.66 (5)	0.10 (7)	0.01 (9)	22.12(12)	16.76(5)	38.13(8)	7.80(7)	8.98(9)	7.64 (6)
27	7.80 (1)	45.00(6)	1.78 (2)	0.13 (3)	0.01 (5)	35.12(5)	22.91(1)	53.15(1)	10.73(1)	14.60(3)	4.85 (10)
28	3.89 (9)	45.00(6)	1.30 (7)	0.06 (9)	0.00 (11)	26.82(8)	13.64(7)	35.29(9)	5.59(9)	7.77(11)	5.58 (8)
29	1.81 (21)	53.57(1)	0.48 (20)	0.01 (22)	0.00 (22)	12.79(18)	3.06(26)	12.87(22)	2.57(21)	3.73(20)	2.46 (18)
30	2.55 (17)	36.67(18)	0.52 (17)	0.01 (19)	0.00 (19)	10.71(21)	6.24(17)	16.23(19)	3.21(18)	4.23(19)	0.44 (28)
Panel B: rank correlation of inputs											
SD	1.00	0.38	0.92	0.91	0.87	0.81	0.90	0.87	0.99	0.86	0.77
LPM ₀	0.38	1.00	0.46	0.29	0.24	0.28	0.33	0.32	0.41	0.23	0.30
LPM ₁	0.92	0.46	1.00	0.94	0.89	0.81	0.95	0.91	0.93	0.83	0.88
LPM ₂	0.91	0.29	0.94	1.00	0.98	0.90	0.96	0.96	0.94	0.95	0.93
LPM ₃	0.87	0.24	0.89	0.98	1.00	0.91	0.93	0.95	0.91	0.97	0.93
MD	0.81	0.28	0.81	0.90	0.91	1.00	0.87	0.96	0.85	0.90	0.88
AD	0.90	0.33	0.95	0.96	0.93	0.87	1.00	0.96	0.93	0.87	0.89
SDD	0.87	0.32	0.91	0.96	0.95	0.96	0.96	1.00	0.91	0.91	0.92
VAR	0.99	0.41	0.93	0.94	0.91	0.85	0.93	0.91	1.00	0.91	0.82
CVAR	0.86	0.23	0.83	0.95	0.97	0.90	0.87	0.91	0.91	1.00	0.86
MVAR	0.77	0.30	0.88	0.93	0.93	0.88	0.89	0.92	0.82	0.86	1.00

Panel A of this table shows, for each of the 30 funds, the measurement results for the 11 inputs defined in Sect. 2.2. The numbers in parentheses indicate the ranking of the measurement values. Panel B shows the resulting Spearman's rank correlation between the inputs. The abbreviations are explained in Table 3. The formulas for calculation of the measures can be found in Appendix A

Table 6 Measurement value and ranking of outputs

Panel A: measurement value and ranking (in parentheses) of outputs						
Input fund	AER	GER	HPM ₀	HPM ₁	HPM ₂	HPM ₃
1	0.45 (17)	0.44 (13)	71.67 (2)	0.83 (19)	0.01 (23)	0.00 (24)
2	0.29 (24)	0.28 (19)	65.83 (9)	0.64 (23)	0.01 (25)	0.00 (26)
3	0.55 (13)	0.16 (24)	65.22 (10)	0.51 (26)	0.01 (24)	0.00 (23)
4	0.17 (26)	-0.06 (28)	53.73 (26)	0.26 (30)	0.00 (30)	0.00 (29)
5	1.02 (5)	0.90 (3)	63.79 (12)	1.92 (6)	0.09 (10)	0.00 (13)
6	0.48 (15)	0.29 (18)	62.62 (14)	1.69 (8)	0.09 (8)	0.01 (8)
7	0.39 (20)	-0.03 (26)	48.72 (28)	1.66 (9)	0.15 (6)	0.02 (5)
8	1.06 (4)	0.89 (4)	58.82 (16)	2.75 (3)	0.20 (4)	0.02 (4)
9	0.87 (7)	0.83 (5)	66.67 (6)	1.59 (10)	0.06 (12)	0.00 (14)
10	0.35 (22)	0.21 (22)	59.17 (15)	2.08 (5)	0.14 (7)	0.01 (7)
11	1.40 (3)	1.15 (2)	57.63 (18)	3.30 (2)	0.32 (2)	0.04 (3)
12	0.64 (10)	0.60 (8)	66.67 (6)	1.43 (13)	0.05 (15)	0.00 (16)
13	0.25 (25)	0.24 (20)	55.83 (22)	0.77 (21)	0.02 (21)	0.00 (22)
14	-0.14 (30)	-0.18 (30)	47.50 (29)	1.06 (15)	0.04 (16)	0.00 (17)
15	0.12 (27)	0.12 (25)	64.17 (11)	0.40 (27)	0.00 (29)	0.00 (30)
16	0.36 (21)	0.33 (16)	58.33 (17)	1.01 (16)	0.04 (17)	0.00 (15)
17	0.10 (28)	-0.04 (27)	66.67 (6)	0.40 (28)	0.01 (28)	0.00 (27)
18	0.59 (12)	0.23 (21)	69.33 (4)	0.52 (25)	0.02 (22)	0.00 (20)
19	0.41 (19)	0.21 (23)	56.52 (21)	0.80 (20)	0.03 (18)	0.00 (18)
20	0.51 (14)	0.47 (12)	79.17 (1)	1.00 (17)	0.02 (19)	0.00 (21)
21	0.80 (8)	0.58 (10)	68.00 (5)	0.94 (18)	0.05 (14)	0.01 (9)
22	0.42 (18)	0.42 (14)	70.83 (3)	0.54 (24)	0.01 (26)	0.00 (28)
23	0.73 (9)	0.68 (6)	57.50 (19)	1.50 (11)	0.08 (11)	0.01 (11)
24	1.50 (2)	0.53 (11)	50.77 (27)	1.47 (12)	0.26 (3)	0.07 (2)
25	0.31 (23)	0.30 (17)	55.83 (22)	0.69 (22)	0.02 (20)	0.00 (19)
26	0.94 (6)	0.68 (7)	57.41 (20)	2.48 (4)	0.18 (5)	0.02 (6)
27	1.75 (1)	1.48 (1)	55.00 (24)	3.53 (1)	0.50 (1)	0.13 (1)
28	0.46 (16)	0.38 (15)	55.00 (24)	1.76 (7)	0.09 (9)	0.01 (10)
29	0.05 (29)	-0.17 (29)	46.43 (30)	0.32 (29)	0.01 (27)	0.00 (25)
30	0.63 (11)	0.60 (9)	63.33 (13)	1.15 (14)	0.06 (13)	0.01 (12)
Panel B: rank correlation of outputs						
AER	1.00	0.88	0.17	0.68	0.72	0.71
GER	0.88	1.00	0.24	0.69	0.64	0.59
HPM ₀	0.17	0.24	1.00	-0.19	-0.30	-0.33
HPM ₁	0.68	0.69	-0.19	1.00	0.96	0.90
HPM ₂	0.72	0.64	-0.30	0.96	1.00	0.98
HPM ₃	0.71	0.59	-0.33	0.90	0.98	1.00

Panel A of this table shows, for each of the 30 funds, the measurement results for the six outputs defined in Sect. 2.2. The numbers in parentheses indicate the ranking of the measurement values. Panel B shows the resulting Spearman’s rank correlation between the outputs. The abbreviations are explained in Table 3. The formulas for calculation of the measures can be found in Appendix A

0.97, component B) have the highest impact in the principal component analysis. Components A and B explain 90.15% of the total variance.

Using the principal component analysis results, we calculate inputs and outputs for Applications 8–10. On the input side, the first principal component is the result of a linear combination of the observed variables SD to MVAR with the factor loadings 0.90–0.91. The second principal component uses the factor loadings -0.05 to -0.04. On the output side, the first principal component

Table 7 Results of principal component analysis

Inputs	Component A	Component B	Outputs	Component A	Component B
Standard deviation	0.90	-0.05	Arithmetic excess return	0.94	0.01
LPM ₀	0.45	0.86	Geometric excess return	0.89	0.39
LPM ₁	0.95	0.20	HPM ₀	0.11	0.97
LPM ₂	0.97	-0.06	HPM ₁	0.90	0.07
LPM ₃	0.91	-0.27	HPM ₂	0.95	-0.24
Maximum drawdown	0.87	-0.05	HPM ₃	0.86	-0.34
Average drawdown	0.96	0.16			
Standard deviation of drawdown	0.97	0.05			
Value at risk	0.93	-0.05			
Conditional value at risk	0.82	-0.36			
Modified value at risk	0.91	-0.04			
Total variance explained (%)	78.94	88.20	Total variance explained (%)	69.11	90.15

The left side of this table shows the principal component analysis results for the 11 inputs defined in Sect. 2.2. For each of the inputs, two factor loadings are calculated; the first one for component A and the second one for component B. The last row gives the total variance explained by the principal components. The right side of the table shows the corresponding results for the six outputs defined in Sect. 2.2. The abbreviations are explained in Table 3. The formulas for calculation of the measures can be found in Appendix A

combines the results of the observed variables AER to HPM₃ with the factor loadings 0.94–0.86, while the second principal component has the loadings 0.01 to -0.34.

3.3 DEA performance measurement

Table 8 (panel A) shows the performance measurement results for the classic Sharpe ratio and for the ten DEA applications. Furthermore, for each measure the resulting fund rankings are presented. The Spearman's rank correlation coefficients between the performance measures are set out in panel B of Table 8 in order to illustrate the relations between the applications.

We first compare the Sharpe ratio and DEA Application 1. Under the Sharpe ratio, fund 22 yields the best performance of 0.48. Fund 14 exhibits the lowest Sharpe ratio of -0.05. By definition, the Sharpe ratio and DEA Application 1 lead to the same ranking of investments and thus to a rank correlation of 1.00. Fund 22 achieves the highest possible efficiency score of 1.00; fund 14 has the worst efficiency score of 0.00.

The integration of variable returns to scale (DEA Application 2) leads to higher efficiency scores than Application 1 with constant returns to scale. On average, the efficiency score in Application 2 amounts to 0.64 compared to 0.41

Table 8 Performance measurement results

Panel A: performance measurement values and ranking (in parentheses) of funds

Fund	1	2	3	4	5	6	7	8	9	10
1	0.30 (6)	0.63 (6)	0.68 (15)	0.74 (15)	0.68 (17)	0.97 (6)	0.98 (7)	0.91 (5)	0.92 (11)	0.92 (11)
2	0.22 (12)	0.46 (12)	0.68 (16)	0.49 (23)	0.68 (18)	0.78 (10)	0.80 (12)	0.70 (9)	0.72 (15)	0.72 (15)
3	0.34 (3)	0.71 (3)	0.88 (10)	0.51 (22)	0.88 (11)	0.50 (23)	0.96 (8)	0.74 (7)	0.94 (10)	0.94 (10)
4	0.18 (18)	0.37 (18)	0.90 (7)	0.70 (16)	1.00 (1)	0.43 (27)	0.77 (15)	0.46 (26)	0.95 (8)	0.95 (8)
5	0.26 (7)	0.55 (7)	0.95 (6)	0.93 (10)	0.97 (9)	0.59 (13)	0.62 (19)	0.69 (12)	0.78 (12)	0.78 (12)
6	0.10 (25)	0.21 (25)	0.24 (28)	0.58 (19)	0.33 (26)	0.75 (12)	0.82 (11)	0.50 (22)	0.59 (20)	0.59 (20)
7	0.06 (27)	0.12 (27)	0.13 (30)	0.48 (24)	0.25 (30)	0.37 (28)	0.63 (18)	0.33 (28)	0.57 (21)	0.57 (21)
8	0.19 (17)	0.39 (17)	0.69 (14)	0.79 (12)	0.70 (16)	0.55 (17)	0.57 (25)	0.57 (16)	0.69 (18)	0.69 (18)
9	0.31 (5)	0.64 (5)	1.00 (1)	1.00 (1)	1.00 (1)	0.77 (11)	0.79 (13)	0.78 (6)	1.00 (1)	1.00 (6)
10	0.07 (26)	0.14 (26)	0.17 (29)	0.56 (21)	0.29 (29)	0.52 (20)	0.55 (28)	0.49 (24)	0.50 (28)	0.50 (28)
11	0.21 (14)	0.43 (14)	0.86 (11)	0.86 (11)	0.86 (12)	0.56 (14)	0.59 (21)	0.64 (14)	0.98 (7)	0.98 (7)
12	0.23 (10)	0.47 (10)	0.64 (17)	0.78 (13)	0.64 (19)	0.54 (18)	0.61 (20)	0.73 (8)	0.73 (14)	0.73 (14)
13	0.15 (22)	0.31 (22)	0.51 (20)	0.47 (25)	0.51 (22)	0.51 (22)	0.58 (23)	0.50 (23)	0.52 (27)	0.52 (27)
14	-0.05 (30)	0.00 (30)	0.30 (26)	0.34 (28)	0.30 (28)	0.35 (29)	0.39 (30)	0.28 (30)	0.35 (30)	0.35 (30)
15	0.12 (23)	0.26 (23)	0.88 (9)	0.43 (27)	0.88 (10)	1.00 (1)	2.62 (1)	0.67 (13)	0.69 (19)	0.69 (19)
16	0.15 (21)	0.32 (21)	0.37 (23)	0.57 (20)	0.37 (25)	0.52 (21)	0.56 (27)	0.54 (20)	0.55 (24)	0.55 (24)
17	0.04 (28)	0.08 (28)	0.34 (25)	0.29 (29)	0.83 (13)	0.53 (19)	0.78 (14)	0.69 (11)	1.00 (1)	1.09 (5)
18	0.37 (2)	0.77 (2)	1.00 (1)	1.00 (1)	1.00 (1)	0.84 (8)	1.00 (1)	0.92 (4)	1.00 (1)	1.15 (4)
19	0.17 (20)	0.35 (20)	0.36 (24)	0.45 (26)	0.40 (24)	0.48 (25)	0.58 (24)	0.49 (25)	0.56 (22)	0.56 (22)
20	0.20 (15)	0.42 (15)	0.49 (21)	0.75 (14)	1.00 (1)	1.00 (1)	2.50 (2)	1.00 (1)	1.00 (1)	2.63 (1)

Table 8 Continued

Fund	Panel A: performance measurement values and ranking (in parentheses) of funds											
	Sharpe ratio	DEA application	1	2	3	4	5	6	7	8	9	10
21	0.31 (4)	0.65 (4)	1.00 (1)	0.98 (5)	1.00 (1)	1.00 (1)	0.86 (7)	0.91 (9)	0.93 (3)	0.94 (9)	0.94 (9)	0.94 (9)
22	0.48 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.62 (4)	1.00 (1)	1.00 (1)	1.00 (1)	2.02 (2)
23	0.24 (9)	0.50 (9)	1.00 (1)	0.72 (12)	1.00 (1)	0.72 (14)	0.56 (15)	0.59 (22)	0.54 (19)	0.55 (26)	0.55 (26)	0.55 (26)
24	0.21 (13)	0.43 (13)	1.00 (1)	0.89 (8)	1.00 (1)	1.00 (1)	1.00 (1)	1.23 (6)	0.56 (17)	0.71 (16)	0.71 (16)	0.71 (16)
25	0.20 (16)	0.41 (16)	1.00 (1)	0.56 (19)	0.70 (17)	0.63 (20)	0.55 (16)	0.63 (17)	0.51 (21)	0.55 (25)	0.55 (25)	0.55 (25)
26	0.17 (19)	0.36 (19)	1.00 (1)	0.59 (18)	0.98 (8)	0.60 (21)	0.79 (9)	0.83 (10)	0.55 (18)	0.56 (23)	0.56 (23)	0.56 (23)
27	0.22 (11)	0.47 (11)	1.00 (1)	1.00 (1)	1.00 (1)	1.00 (1)	1.00 (1)	2.13 (3)	0.60 (15)	1.00 (1)	1.00 (1)	1.60 (3)
28	0.12 (24)	0.24 (24)	1.00 (1)	0.26 (27)	0.62 (18)	0.31 (27)	0.46 (26)	0.50 (29)	0.45 (27)	0.46 (29)	0.46 (29)	0.46 (29)
29	0.03 (29)	0.06 (29)	1.00 (1)	0.49 (22)	0.25 (30)	0.49 (23)	0.32 (30)	0.77 (16)	0.30 (29)	0.76 (13)	0.76 (13)	0.76 (13)
30	0.25 (8)	0.51 (8)	1.00 (1)	0.69 (13)	0.93 (9)	0.71 (15)	0.49 (24)	0.56 (26)	0.69 (10)	0.70 (17)	0.70 (17)	0.70 (17)
Panel B: rank correlation of performance measures												
SR	1.00	1.00	0.81	0.76	0.69	0.69	0.52	0.43	0.80	0.55	0.56	0.56
1	1.00	1.00	0.81	0.76	0.69	0.69	0.52	0.43	0.80	0.55	0.56	0.56
2	0.81	0.81	1.00	0.74	0.87	0.87	0.57	0.54	0.62	0.64	0.64	0.64
3	0.76	0.76	0.74	1.00	0.67	0.67	0.60	0.33	0.52	0.42	0.42	0.41
4	0.69	0.69	0.69	0.67	1.00	1.00	0.65	0.68	0.69	0.80	0.80	0.80
5	0.52	0.52	0.52	0.57	0.60	0.65	1.00	0.78	0.68	0.49	0.49	0.49
6	0.45	0.45	0.38	0.56	0.70	0.70	0.79	1.00	0.60	0.67	0.67	0.68
7	0.43	0.43	0.33	0.54	0.68	0.68	0.78	0.99	0.61	0.67	0.67	0.68
8	0.80	0.80	0.52	0.62	0.69	0.69	0.68	1.00	1.00	0.72	0.72	0.72
9	0.55	0.55	0.42	0.64	0.80	0.80	0.49	0.67	0.72	1.00	1.00	0.99
10	0.56	0.56	0.41	0.64	0.80	0.80	0.49	0.68	0.72	0.99	0.99	1.00

Panel A of this table shows, for each of the 30 funds, the measurement results for the Sharpe ratio and for the 10 DEA applications defined in Table 4. The numbers in parentheses indicate the ranking of the measurement values. Panel B shows the resulting Spearman's rank correlation between the performance measures

in Application 1. In Application 2, four funds achieve the maximum efficiency score of 1.00—funds 9, 18, 22, and 27. While funds 9, 18, and 22 were already among the best funds in Application 1, fund 27 gains ten places. Assuming decreasing returns to scale this fund seems to be more attractive because it has the highest return but also the highest standard deviation. Comparing rank correlations between the Applications 1 and 2, we find some changes in rankings; the rank correlation is 0.81.

The rank correlations are noticeably lower in the applications employing multi-dimensional performance analysis (Applications 3–10), thus resulting in a different evaluation of the investments. For example, according to Application 3, fund 15 achieves an efficiency score of only 0.43 and ranks at 27th place. Obviously, the low rankings with the higher partial moments of order 1–3 lead to a worse evaluation of this fund compared to Application 2. Fund 21 is now in the first place, apparently due to its good ranking with the lower partial moments. Hence, the rank correlation between the Sharpe ratio and Application 3 is 0.76.

The rank correlations are even lower with Applications 5–7. It amounts to 0.52 with Application 5, 0.45 with Application 6, and 0.43 with Application 7. Therefore, we find a significantly different fund evaluation when using these proposed applications.¹³ For example, fund 15, which is in 23rd place with the Sharpe ratio, is evaluated as efficient in these three applications. This might be due to its good ranking with the HPM_0 . Comparing Applications 5–7 with each other, we find small changes in rankings. Thus, on basis of our data it only makes a slight difference as to which of the three DEA modeling approaches (Charnes et al. 1978; Banker et al. 1984; Andersen and Petersen 1993) is used to assess hedge fund performance.

We also find small differences comparing Applications 8–10. The rank correlations compared to the Sharpe ratio are now 0.55 (Application 9), 0.56 (Application 10), and 0.80 (Application 8). Fund 15 (first place with Applications 5–7) is now either in 13th or 19th place, but fund 20 is in first place for all applications. In this context it is striking that Fund 20 has the highest HPM_0 . Also of interest is a comparison of Applications 9 and 10. With Application 9, six funds (9, 17, 18, 20, 22, and 27) achieve the maximum efficiency score of 1.00 and rank at first place. Using the super-efficiency approach of Application 10, it is possible to provide a ranking of these efficient funds: Fund 9 is sixth, Fund 17 is fifth, Fund 18 is fourth, Fund 27 is third, Fund 22 is second, and Fund 20 is in first place.

¹³ The statistical significance of the rank correlation can be checked with the Hotelling–Pabst statistic. In this test, the hypothesis of independence of the rankings is checked for the correlation coefficient. The test statistic D has the following form: $D = d_1^2 + \dots + d_7^2$; see Hotelling and Pabst (1936). The null hypothesis H_0 : “The measurement series are independent” must be rejected at level $\alpha = 0.01$ if $D < 2,402$ or $D > 6,110$. In this example these critical values are attained at a rank correlation of $r_{sp} = |0.47|$. Thus, if the rank correlation is smaller than 0.47 and higher than -0.47 we conclude that we find a significantly different fund evaluation.

To sum up, using the new DEA applications, we find differences in rankings compared to the classic Sharpe ratio and compared to previously proposed DEA applications. This is especially true when the selection rule based on Spearman's rank correlation is applied. Therefore, it seems that DEA is able to capture additional information contained in hedge fund returns and provides a new evaluation of the hedge funds under consideration.

4 Conclusion

In this paper, data envelopment analysis (DEA) has been suggested as a method for hedge fund performance measurement. DEA is a linear programming approach that can be used to model the relationship between inputs and outputs and to assess the efficiency of any decision-making unit. The resulting efficiency score can be interpreted as a performance measure. Recent literature confirms that using DEA to measure the performance of hedge funds yields some advantages compared to classic performance measures.

The contribution of this paper is to put the existing literature into a common framework and to provide rules for the selection of inputs and outputs when implementing DEA. We first presented the basic principle of DEA, that is, the general model approach and different model variants. Next, we systematized possible inputs and outputs and analyzed their eligibility for hedge fund performance measurement. Finally, two rules were developed that investors might use for their selection of inputs and outputs. The first rule is based on Spearman's rank correlation and the second on principal component analysis. Using the new DEA applications described in this paper, significant differences in the rankings of hedge funds have been found compared to the classic Sharpe ratio and compared to previously proposed DEA applications.

What are the conclusions that can be drawn from these results? We cannot conclude that DEA based on the suggested rules provides a better evaluation compared to classic performance measures such as the Sharpe ratio or compared to previously proposed DEA applications, as we do not have any criteria to evaluate the quality of different performance measures.¹⁴ Furthermore, as higher moments in financial returns are very unstable over time, it is questionable how stable DEA's ranking results are, as well as those of other performance measures that include higher moments. Given their dynamic trading strategies, the answer to this question will be particularly important to hedge funds. Future research should thus evaluate the implications of the instability of higher moments on performance measurement for hedge funds and other investments.

¹⁴ Moreover, the suitability of different performance measures depends on the preferences of the investor and on the concrete decision-making situation, which makes a generally accepted comparison impossible; see, e.g., [Scholz and Wilkens \(2003\)](#)

But, as illustrated by the large differences in rankings and the low rank correlations, our DEA applications provide additional information about the funds' risk and return characteristics, especially when the selection rule based on Spearman's rank correlation is applied. For an investor, it might be useful to have this additional information, e.g., for the selection and assessment of those funds, that provide comparable measurement values with classic two-dimensional measures. Thus, we propose that classic performance measurements should be supplemented with DEA to achieve a more complete view of hedge fund performance.

Acknowledgments The author thanks Wolfgang Bessler, Nadine Gatzert, Thomas Parnitzke, Hato Schmeiser, Denis Toplek, and two anonymous referees for their helpful comments and suggestions.

Appendix A: Formulas

Average drawdown

$$AD_{ik} = 1/K \sum_{k=1}^K -D_{ik} \tag{A1}$$

with K : number of drawdowns ($k = 1$: maximum drawdown, $k = 2$: second-largest drawdown, $k = 3$: third-largest drawdown, ...)
 D_{ik} : drawdown of fund i

Average excess return (computed arithmetically)

$$AER_i = r_{i\text{arithmetic}}^d - r_f = 1/T \sum_{t=1}^T r_{it} - r_f \tag{A2}$$

with r_{it} : discrete return of fund i in month t ($t = 1, \dots, T$)
 r_f : (constant) risk-free interest rate
 T : number of months

Average excess return (computed geometrically)

$$GER_i = r_{i\text{geometric}}^d - r_f = \sqrt[T]{\prod_{t=1}^T (1 + r_{it})} - 1 - r_f \tag{A3}$$

Conditional value at risk

$$CVAR_i = E[-r_{it} | r_{it} \leq -VAR_i] \tag{A4}$$

with VAR_i : value at risk

Higher partial moments

$$HPM_{in} = \begin{cases} \Pr(r_{it} \geq \tau) & \text{for } n = 0 \\ 1/T \sum_{t=1}^T \max[r_{it} - \tau, 0]^n & \text{for } n > 0 \end{cases} \tag{A5}$$

with n : order of the lower partial moment
 τ : minimal acceptable return

Lower partial moments

$$LPM_{in} = \begin{cases} \Pr(r_{it} < \tau) & \text{for } n = 0 \\ 1/T \sum_{t=1}^T \max[\tau - r_{it}, 0]^n & \text{for } n > 0 \end{cases} \tag{A6}$$

Modified value at risk

$$MVAR_i = - \left(r_{i \text{ arithmetic}}^d + SD_i \left(z_\alpha + (z_\alpha^2 - 1) S_i / 6 \right. \right. \\ \left. \left. + (z_\alpha^3 - 3 \cdot z_\alpha) E_i / 24 - (2z_\alpha^3 - 5z_\alpha) S_i^2 / 36 \right) \right) \tag{A7}$$

with

$$S_i : \text{skewness} \left(= \left(1/T \sum_{t=1}^T (r_{it} - r_{i \text{ arithmetic}}^d)^3 \right) / SD_i^3 \right)$$

$$E_i : \text{excess kurtosis} \left(= \left(1/T \sum_{t=1}^T (r_{it} - r_{i \text{ arithmetic}}^d)^4 \right) / SD_i^4 - 3 \right)$$

Standard deviation

$$SD_i = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_{it} - r_{i \text{ arithmetic}}^d)^2} \tag{A8}$$

Standard deviation of drawdown (square root of the squared maximum drawdowns)

$$SDD_{ik} = \sqrt{\sum_{k=1}^K D_{ik}^2} \tag{A9}$$

Value at risk

$$VAR_i = -\left(r_{i \text{ arithmetic}}^d + z_\alpha SD_i\right) \tag{A10}$$

with z_α : α -quantile of the standard normal distribution

Appendix B: Robustness of results

Table 9 Robustness of results

	Rank correlation compared to the Sharpe ratio for DEA Application									
	1	2	3	4	5	6	7	8	9	10
Section 3	1.00	0.81	0.76	0.69	0.52	0.45	0.43	0.80	0.55	0.56
Hedge fund strategy										
Convertible arbitrage	1.00	0.72	0.42	0.72	0.47	0.50	0.49	0.80	0.46	0.47
Distressed securities	1.00	0.88	0.52	0.72	0.29	0.34	0.32	0.63	0.55	0.48
Event driven multi strategy	1.00	0.78	0.72	0.79	0.68	0.44	0.43	0.79	0.52	0.52
Equity long only	0.99	0.60	0.73	0.43	0.57	0.42	0.41	0.44	0.30	0.28
Equity long short	1.00	0.88	0.69	0.77	0.48	0.48	0.49	0.80	0.73	0.77
Emerging markets	1.00	0.96	0.30	0.84	0.43	0.44	0.50	0.86	0.45	0.49
Equity market neutral	1.00	0.68	0.85	0.74	0.72	0.73	0.73	0.89	0.68	0.68
Fixed income arbitrage	1.00	0.73	0.72	0.17	0.49	0.42	0.42	0.56	0.34	0.38
Funds of funds	1.00	0.80	0.31	0.59	0.51	0.52	0.52	0.47	0.36	0.28
Global macro	1.00	0.87	0.71	0.68	0.45	0.32	0.32	0.69	0.39	0.41
Merger arbitrage	1.00	0.73	0.61	0.74	0.67	0.57	0.56	0.57	0.15	0.15
Mortgage backed securities	1.00	0.91	0.58	0.89	0.34	0.45	0.42	0.37	0.18	0.21
Multi strategy	1.00	0.88	0.37	0.75	0.39	0.35	0.35	0.32	0.13	0.00
Sector	1.00	0.77	0.15	0.60	0.51	0.50	0.47	0.57	0.17	0.16
Short bias	0.99	0.68	0.45	0.70	0.24	0.22	0.25	0.34	0.20	0.26
Mean	1.00	0.79	0.54	0.68	0.48	0.45	0.45	0.61	0.37	0.37

Table 9 Continued

	Rank correlation compared to the Sharpe ratio for DEA Application									
	1	2	3	4	5	6	7	8	9	10
Time horizon										
1996–2000	1.00	0.87	0.64	0.84	0.69	0.50	0.52	0.87	0.69	0.72
2001–2005	1.00	0.83	0.86	0.75	0.61	0.62	0.62	0.72	0.60	0.63
Mean	1.00	0.85	0.75	0.80	0.65	0.56	0.57	0.79	0.65	0.68
Portfolio perspective										
hedge fund portion 1%	1.00	0.75	0.31	0.14	0.56	0.05	0.09	0.65	0.16	0.18
hedge fund portion 5%	1.00	0.85	0.39	0.09	0.59	0.21	0.19	0.71	0.39	0.39
hedge fund portion 10%	0.98	0.81	0.47	0.66	0.75	0.25	0.22	0.78	0.53	0.52
Mean	0.99	0.80	0.39	0.30	0.63	0.17	0.16	0.72	0.36	0.36

This table shows the Spearman's rank correlations of the 10 DEA applications defined in Table 4 compared to the Sharpe ratio for 15 different hedge fund strategies, 2 different time horizons, and 3 different decision-making situations from a portfolio perspective. The 15 different hedge fund strategies are defined by the Center for International Securities and Derivatives Markets (CIS-DM). To investigate the effect of different time horizons, we broke down the total investigation period from 1996 to 2005 into two equal time periods. To investigate the portfolio perspective, we measured the performance of a portfolio comprised of traditional investments and a small portion of hedge funds (1, 5, and 10%). Detailed results are available upon request

References

- Adler, N., Golany, B.: Including principal component weights to improve discrimination in data envelopment analysis. *J. Oper. Res. Soc.* **53**(9), 985–991 (2002)
- Allen, R., Athanassopoulos, A., Dyson, R.G., Thanassoulis, E.: Weights restrictions and value judgments in data envelopment analysis: evolution, development and future directions. *Ann. Oper. Res.* **73**(1), 13–34 (1997)
- Ammann, M., Moerth, P.: Impact of fund size on hedge fund performance. *J. Asset Manage.* **6**(3), 219–238 (2005)
- Andersen, P., Petersen, N.C.: A procedure for ranking efficient units in data envelopment analysis. *Manage. Sci.* **39**(1), 1261–1264 (1993)
- Artzner, P., Delbaen, F., Eber, J.-M., Heath, D.: Coherent measures of risk. *Mathe. Finance* **9**(3), 203–228 (1999)
- Banker, R.D., Charnes, A., Cooper, W.W.: Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage. Sci.* **30**(9), 1078–1092 (1984)
- Banz, R.W., de Planta, R.: Hedge funds: all that glitters is not gold—seven questions for prospective investors. *Financ. Markets Portfolio Manage.* **16**(3), 316–336 (2002)
- Basso, A., Funari, S.: A data envelopment analysis approach to measure the mutual fund performance. *Euro. J. Oper. Res.* **135**(3), 477–492 (2001)
- Bessler, W., Norsworthy, J.R.: A dynamic model of bank production with adjustment costs and multiple risk effects. In: Hasan, I., Hunter, W.C. (eds.), *Research in Banking and Finance*, vol. 3. pp. 115–146. Elsevier, Oxford (2003)
- Borner, H.: Things investors should know about alternative investments. *Financ. Markets Portfolio Manage.* **18**(3), 324–335 (2004)
- Bowlin, W.F.: Measuring performance: an introduction to data envelopment analysis (DEA). *J. Cost Anal.* **3**(1), 3–28 (1998)
- Capocci, D., Hübner, G.: Analysis of hedge fund performance. *J. Empir. Finance* **11**(1), 55–89 (2004)
- Charnes, A., Cooper, W.W., Rhodes, E.: Measuring the efficiency of decision making units. *Eur. J. Oper. Res.* **2**(6), 429–444 (1978)

- Choi, Y.K., Murthi, B.P.S.: Relative performance evaluation of mutual funds: a non-parametric approach. *J. Bus. Finance Account.* **28**(7/8), 853–876 (2001)
- Cooper, W.W., Seiford, L.M., Tone, K.: *Introduction to Data Envelopment Analysis and Its Uses*. Springer, Berlin Heidelberg New York (2006)
- Edwards, F.R., Caglayan, M.O.: Hedge fund performance and manager skill. *J. Futures Mark.* **21**(11), 1003–1028 (2001)
- Eling, M.: *Hedgefonds-Strategien und ihre Performance*. Eul-Verlag, Lohmar (2006)
- Eling, M., Schuhmacher, F.: Does the choice of performance measure influence the evaluation of hedge funds? *J. Bank. Finance* (2006) (in press)
- Farell, M.J.: The measurement of productive efficiency. *J. R. Stat. Soc.* **120**(3), 253–281 (1957)
- Favre, L., Galéano, J.: The inclusion of hedge funds in swiss pension fund portfolios. *Financ. Mark. Portfolio Manage.* **15**(4), 450–472 (2001)
- Fecher, F., Pestieau, P.: Efficiency and competition in OECD financial services. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (eds.) *The Measurement of Productivity Efficiency: Techniques and Applications*, pp. 374–385. Oxford University Press, New York (1993)
- Ferris, S.P., Chance, D.M.: The effect of 12b-1 plans on mutual fund expense ratios: A note. *J. Finance* **42**(4), 1077–1082 (1987)
- Fung, W., Hsieh, D.A.: Empirical characteristics of dynamic trading strategies: the case of hedge funds. *Rev. Financ. Stud.* **10**(2), 275–302 (1997)
- Gregoriou, G.N.: Performance appraisal of funds of hedge funds using data envelopment analysis. *J. Wealth Manage.* **5**(4), 88–95 (2003)
- Gregoriou, G.N., Gueyie, J.-P.: Risk-adjusted performance of funds of hedge funds using a modified sharpe ratio. *J. Alterna. Invest.* **6**(3), 77–83 (2003)
- Gregoriou, G.N., Sedzro, K., Zhu, J.: Hedge fund performance appraisal using data envelopment analysis. *Eur. J. Oper. Res.* **164**(2), 555–571 (2005)
- Halme, M., Joro, T., Korhonen, P., Salo, S., Wallenius, J.: A value efficiency approach to incorporating preference information in data envelopment analysis. *Manage. Sci.* **45**(1), 103–115 (1999)
- Hotelling, H., Pabst, M.R.: Rank correlation and tests of significance involving no assumption of normality. *Ann. Math. Stat.* **7**(1), 29–43 (1936)
- Hwang, S., Satchell, S.E.: Modelling emerging market risk premia using higher moments. *Int. J. Finance Econ.* **4**(4), 271–296 (1999)
- Jackson, J.E.: *A user's Guide to principal Components*. Wiley, Hoboken (2003)
- Jolliffe, I.T.: *Principal Component Analysis*. Springer, Berlin Heidelberg New York (2002)
- Kassberger, S., Kiesel, R.: A fully parametric approach to return modelling and risk management of hedge funds. *Financ. Mark. Portfolio Manage.* **20**(4) (2006)
- Kendall, M., Gibbons, J.D.: *Rank Correlation Methods*, 5th edn. Edward Arnold, London (1990)
- Kouwenberg, R.: Do hedge funds add value to a passive portfolio? *J. Asset Manage.* **3**(4), 361–382 (2003)
- Kuosmanen, T., Cherchye, L., Sipiläinen, T.: The law of one price in data envelopment analysis: restricting weight flexibility across firms. *Eur. J. Oper. Res.* **170**(3), 735–757 (2006)
- Leland, H.E.: Beyond mean-variance: performance measurement in a nonsymmetrical world. *Financ. Anal. J.* **55**(1), 27–36 (1999)
- Le Moigne, C., Savaria, P.: Relative importance of hede funds characteristics. *Financ. Mark. Portfolio Manage.* **20**(4) (2006)
- Lhabitant, F.-S.: *Hedge Funds: Quantitative Insights*. Wiley, Chichester (2004)
- Mahajan, J.: A data envelopment analytical model for assessing the relative efficiency of the selling function. *Euro. J. Oper. Res.* **53**(2), 189–205 (1991)
- Malkiel, B.G., Saha, A.: Hedge funds: risk and return. *Financ. Anal. J.* **61**(6), 80–88 (2005)
- McMullen, P., Strong, R.A.: Selection of mutual funds using data envelopment analysis. *J. Busi. Econ. Stud.* **4**(1), 1–14 (1998)
- Morey, M.R., Morey, R.C.: Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. *Omega* **27**(2), 241–258 (1999)
- Murthi, B.P.S., Choi, Y.K., Desai, P.: Efficiency of mutual funds and portfolio performance measurement: a non-parametric approach. *Eur. J. Oper. Res.* **98**(2), 408–418 (1997)
- Nguyen-Thi-Thanh, H.: *On the use of data envelopment analysis in hedge fund selection*. Working Paper, Université d'Orléans (2006)
- Parkan, C.: Calculation of operational performance ratings. *Int. J. Prod. Econ.* **24**(1/2), 165–173 (1991)

- Pedraja-Chaparro, F., Salinas-Jimenez, J., Smith, P.: On the quality of the DEA model. *J. Oper. Res. Soc.* **50**(6), 636–644 (1999)
- Scheel, H.: EMS: efficiency measurement system user's manual, version 1.3. Working Paper, University of Dortmund (2000)
- Scholz, H., Wilkens, M.: Zur Relevanz von Sharpe Ratio und Treynor Ratio: Ein investorspezifisches Performancemaß. *Zeits. Bankrecht Bankwirtschaft* **15**(1), 1–8 (2003)
- Sherman, H.D.: Hospital efficiency measurement and evaluation. Empirical test of a new technique. *Med. Care* **22**(10), 922–938 (1984)
- Sherman, H.D., Gold, F.: Bank branch operating efficiency: evaluation with data envelopment analysis. *J. Bank. Finance* **9**(2), 297–315 (1985)
- Smith, P.: Model misspecification in data envelopment analysis. *Ann. Oper. Res.* **73**(1), 233–252 (1997)
- Spearman, C.: The proof and measurement of association between two things. *Am. J. Psychol.* **15**(1), 72–101 (1904)
- Steinmann, L.: Konsistenzprobleme der Data Envelopment Analysis in der empirischen Forschung. Socioeconomic Institute, Zürich (2002)
- Wilkens, K., Zhu, J.: Classifying hedge funds using data envelopment analysis. In: Gregoriou, G.N., Rouah, F., Karavas, V.N. (eds.) *Hedge Funds: Strategies, Risk Assessment, and Returns*. Beard Books, Washington pp. 161–175 (2005)



Dr. Martin Eling is currently a postdoctoral research fellow and academic assistant at the University of St. Gallen, Switzerland. He received his doctoral degree from the University of Münster, Germany. His research interests include empirical finance, asset liability management, and alternative investments.