# On the Use of Data Envelopment Analysis in Hedge Fund Performance Appraisal

Huyen Nguyen-Thi-Thanh[*†]

This Draft: December 2006

## Abstract

This paper aims to show that Data Envelopment Analysis (DEA) is an efficient tool to assist investors in multiple criteria decision-making tasks like assessing hedge fund performance. DEA has the merit of offering investors the possibility to consider simultaneously multiple evaluation criteria with direct control over the priority level paid to each criterion. By addressing main methodological issues regarding the use of DEA in evaluating hedge fund performance, this paper attempts to provide investors sufficient guidelines for tailoring their own performance measure which reflect successfully their own preferences. Although these guidelines are formulated in the hedge fund context, they can also be applied to other kinds of investment funds.

JEL CLASSIFICATION: G2, G11, G15

KEYWORDS: hedge fund, mutual fund, data envelopment analysis, performance measures, Sharpe ratio.

[*]Laboratoire d'Economie d'Orléans (LEO), Université d'Orléans, Rue de Blois, B.P. 6739, 45 067 Orléans Cedex 2. E-mail: huyen.nguyen-thi-thanh@univ-orleans.fr.

# Introduction

The highly successful performance of the so-called hedge funds over the past two decades, notably during the long bull equity market of the 1990s, has made them quickly well-known to financial communities as well as to the public. While hedge funds still manage only \$1 trillion at the end of 2004, a fraction of the \$8 trillion invested by mutual funds, their assets have ballooned from only about \$150 billion a decade ago. With over 8,000 hedge funds now available, fund selecting is quite challenging for investors. Hence, before any due-diligence process, investors first need an efficient tool to assist them in screening task in which the most important evaluation is undoubtedly fund (historical) performance.

In general, the historical performance of funds is defined as their return adjusted for risk. According to traditional financial theories, the risk is measured either by the standard deviation of returns or by the correlation of fund returns with market factors via different betas[1]. Most of these measures, even though validated in "buy-and-hold" portfolios of mutual funds and pension funds, are irrelevant within the context of hedge funds. On the one hand, hedge fund returns are documented as usually asymmetric and kurtotic, a feature largely imputed to the intensive use of short sales, leverage, derivative instruments and to the free call-option like incentive structure, all specific to only the hedge fund industry. On the other hand, their short-term movements across diverse asset categories and the market neutral absolute investment objective of hedge fund managers make it really delicate to identify market factors necessary to the use of multi-factorial models[2]. Recent techniques enlarge the evaluation dimension to the skewness (Stutzer 2000), the skewness and the kurtosis (Gregoriou & Gueyie 2003) or to the whole distribution of returns (Keating & Shadwick 2002) in order to take into account the non normality of return distributions. Despite this significant progress, these measures do not allow considering after-net-returns fees paid by investors if only. Besides, most of them are restrictive in the sense that they often assume very simplistic decision-making rules which are common to all investors.

Yet, it is well documented that actual evaluation criteria, in fact, may be more complicated and differ significantly from theoretical formulations. Not only are there many

---

[1] When fund's risk is measured by betas, fund performance is simply the alpha.

[2] Unlike other kinds of investment funds, hedge funds are loosely regulated, and in many cases, are largely exempted from legal obligations as the case of offshore hedge funds. Hedge fund managers thus have a broad flexibility in determining the proportion of securities they hold, the type of positions (long or short) they take and the leverage level they make. As a consequence, they are free to make very short-term movements across diverse asset categories involving frequent use of short sales, leverage and derivatives to attempt to time the market.

attributes to consider, each one being associated with a priority level, but also these attributes and their importance level are usually quite specific to each investor. The need to consider simultaneously multiple criteria while incorporating investors' own preferences is natural since they are do not always share the same financial objective, risk aversion, investment horizon, etc. From such perspectives, the Data Envelopment Analysis approach (hereafter, DEA) seems particularly appealing as it provides the possibility of incorporating many criteria at the same time, together with a direct control over the importance level paid to each criterion by means of a tailor-made optimizing system.

DEA can be roughly defined as a mathematical optimizing technique first introduced by Charnes, Cooper & Rhodes (1978), based on Farrell (1957)'s efficiency concept, to measure the efficiency (technical, allocative, economic, etc.) of decision-making units (hereafter, DMU) whose objective consists in transforming multiple inputs into multiple outputs. The merits of the DEA method lies in providing an unique aggregate measure for each DMU from a system of multiple inputs and multiple outputs and in putting emphasis on the "best observed practices" in a comparative perspective. In addition, DEA allows considering inputs and outputs whose measure units are different, a property known as "units invariance"[3]. Furthermore, it makes no assumption on the form of the relation between inputs and outputs.

Because of its many advantages, DEA has been applied in various fields including public administration (to evaluate hospitals, administrative offices, educational establishments or to resolve siting problems), engineering (to evaluate airplanes and engines), commerce (to evaluate supermarkets), finance (to evaluate bank branches, micro-finance institutions, assurance companies, to identify dominant financial assets and recently to assess investment funds' performance). The application of DEA is generally proceeded in two main perspectives: (1) to evaluate the efficiency of DMUs whose activities are to employ inputs to produce outputs; and (2) to solve decision-making problems with multiple criteria. It is in the second perspective that DEA can be applied to assess hedge fund performance. Initiated by Murthi et al. (1997) to evaluate empirically the performance of mutual funds, this idea has been applied and revisited by several studies, including those on hedge fund performance. However, this literature is composed essentially of empirical applications, methodological issues remain either ignored or discussed in a simplistic and superficial manner with little directive value. To the best of my knowledge, none of methodological studies investigates the use of DEA in the hedge fund context.

---

[3]This is true provided that unit measures are the same for all DMUs in the sample. For example, one person can measure outputs in mile and inputs in gallons of gasoline and quarts of oil while another measures these same outputs and inputs in kilometers and liters with the same collection of automobiles .

Following this literature, this paper is devoted to methodological issues in applying DEA to hedge fund performance appraisal. Specifically, I focus on the choice of evaluation criteria (DEA's inputs and outputs), the choice of DEA models with and without negative data on returns and performance, and on "transcribing" specific evaluating preferences of investors into mathematical constraints. By doing so, this study attempts to offer investors sufficient guidelines in order to apply successfully the DEA method to assessing hedge fund performance. Although it only addresses the hedge fund context, the whole framework is completely applicable to mutual funds, pension funds, ethical funds, etc.

The remainder of the paper is organized as follows. Section 1 reviews briefly the literature related to this study. Section 2 introduces basic concepts of the DEA method. Section 3 addresses methodological issues of applying DEA to screening hedge funds via their performance. Section 4 provides several numerical illustrations on a sample including 38 hedge funds. The last section summaries and concludes the paper.

# 1   Related literature

This study emanates from two main streams of literature. The first one concerns DEA's use in making a selection when decision-makers have multiple criteria. The second evolves evaluating the performance of investment funds by means of the DEA method.

With respect to the first literature, three studies can be enumerated: Thompson et al. (1986), Tone (1999) and Powers & McMullen (2000). Thompson et al. (1986) dealt with identifying feasible sites among six candidate sites for location of a very high-energy physics lab in Texas. A comparative analysis between six sites was conducted by applying the basic DEA model, incorporating project cost, user time delay, and environmental impact data as selection criteria. These criteria are those evaluators want to minimize, they thus form exclusively the DEA's inputs. Being absent, the output is assumed to be unique and equal to unity so that DEA can be applied. This setting is naturally plausible as it is equivalent to considering inputs per one unity of output[4]. In the same spirit, Tone (1999) described a japanese governmental project applying DEA to select a city to take over some political functions of Tokyo as a new capital. In this study, the selection criteria are composed of distance from Tokyo, safety indexes (regarding earthquakes and volcanoes), access to an international airport, ease of land acquisition, landscape, water supply, matters with historical associations; they form exclusively DEA's outputs. The

---

[4]Inputs (outputs) include all criteria that evaluators want to minimize (maximize).

input is thus set to be equal to unity[5]. It is important to note that in these studies, only inputs (outputs) are available and thus output (input) is assumed to be unique and equal to 1. Another common interesting point is that the evaluators, with prior expert knowledge about the relative importance of chosen criteria, fixe lower and upper bounds to the weights associated with each criterion in the mathematical optimization. In finance, Powers & McMullen (2000) suggested using DEA to select dominant stocks among the 185 american largest capitalization stocks because this technique makes it possible to incorporate multiple selection attributes such as the Price-Earnings Ratio, the systematic risk and the total risk (DEA's inputs), the Earnings Per Share ratio and the mean return over 1 year, 3 years, 5 years and 10 years (DEA's outputs).

The second literature relates to studies using DEA to evaluate the performance of mutual funds, ethical funds and more recently hedge funds. Studies on mutual funds include Murthi et al. (1997), McMullen & Strong (1998), Choi & Murthi (2001), Basso & Funari (2001), Tarim & Karan (2001) and Sengupta (2003). All these studies assume that fund performance is a combination of multiple attributes such as mean returns (DEA's outputs), total or systematic risk, expenses[6], and sometimes even fund size, turnover speed and minimum initial investment (DEA's inputs). In the same vein, Basso & Funari (2003) suggested putting in the DEA's outputs, together with the mean return, an indicator measuring funds' ethical level fulfillments since according to them, "the solidarity and social responsibility features that characterize the ethical funds satisfy the fulfillment of humanitarian aims, but may lower the investment profitability".

The application of DEA in evaluating hedge funds emerged from the work of Gregoriou (2003). It was then supported by Gregoriou et al. (2005)[7] and discussed in Kooli et al. (2005). A common feature of these studies is that they only consider risk–return performance without referring to fees. Besides, risks and returns are approximated respectively by lower variations (what investors seek to minimize) and upper variations (what investors seek to maximize) compared to a threshold defined by mean return. Specifically, the inputs are composed of lower mean monthly semi-skewness, lower mean monthly semi-variance and mean monthly lower return; the outputs include upper mean monthly semi-skewness, upper mean monthly semi-variance and mean monthly upper return. Another common feature is that they put emphasis on fund's absolute rankings by

---

[5]I did not have access to documents related to this project. All the information mentioned here is extracted from Cooper et al. (2000, p.169).

[6]The concept of expenses differs from study to study. It might include transaction costs and administration fees (totaled in expense ratio) and loads (subscription or/and redemption costs).

[7]Gregoriou et al. (2005) is an extended version of Gregoriou (2003) and more complete while employing the same DEA methodology with Gregoriou (2003). Therefore, I refer only to Gregoriou et al. (2005).

employing modified DEA techniques: super–efficiency (Andersen & Petersen 1993) and cross–efficiency (Sexton et al. 1986). By comparing DEA results with rankings provided by Sharpe and modified Sharpe ratios via rank correlation coefficients, they observed a weak consistency between DEA and these measures. In particular, Kooli et al. (2005) found quite low correlation between DEA rankings and rankings given by the stochastic dominance technique and concluded to a weak relevancy of DEA to fund performance evaluation context. With regard to super-efficiency and cross-efficiency models, despite their appealing property, i.e. providing fund absolute rankings, their technical caveats cast doubts about their efficacy[8]. Hence, in what follows, I will only introduce the basic DEA model and its dichotomic classification into assessing hedge fund performance.

## 2   DEA's approach

### 2.1   DEA as a measure of technical efficiency

Before introducing the general approach of DEA and the basic DEA model, it is important to distinguish the "technical efficiency", on which is based this study, from the "economic efficiency" usually applied in production context. According to Fried, Lovell & Schmidt (1993, p.9-10), "productive efficiency has two components. The purely technical, or physical, component refers to the ability to avoid waste by producing as much output as input usage allows, or by using as little input as output production allows.... The allocative, or price, or economic, component refers to the ability to combine inputs and outputs in optimal proportions in light of prevailing prices.". Consequently, technical efficiency measurement is based solely on quantity information on the inputs and the outputs whereas the economic efficiency necessitates the recourse to information on prices as well as on economic behavioral objectives of producers (cost minimization, profit maximization or revenue maximization). Conceptually, the efficiency of each DMU under evaluation is determined by the distance from the point representing this DMU to the efficient frontier (production frontier in the case of technical efficiency; cost, revenue or profit frontier in the case of cost, revenue or profit efficiency respectively). In figure 1, the isoquant $L(y)$ represents various combinations of two inputs that a perfectly efficient

---

[8]The super-efficiency model has two main caveats. First, it allocates so excessively high efficiency score to efficient DMUs having extreme values of inputs and outputs that optimal values can sometimes "explode". Second, it is infeasible in some circumstances (Zhu 1996, Seiford & Zhu 1999). The pitfall of the cross-efficiency model is that it penalizes DMUs whose the combination of inputs and outputs is different from the others while it highly praises average DMUs. Besides, the use of the mean, the variance, the mode or the median, etc. of scores to completely rank DMUs is too ambiguous, especially when different indicators provide different rankings.
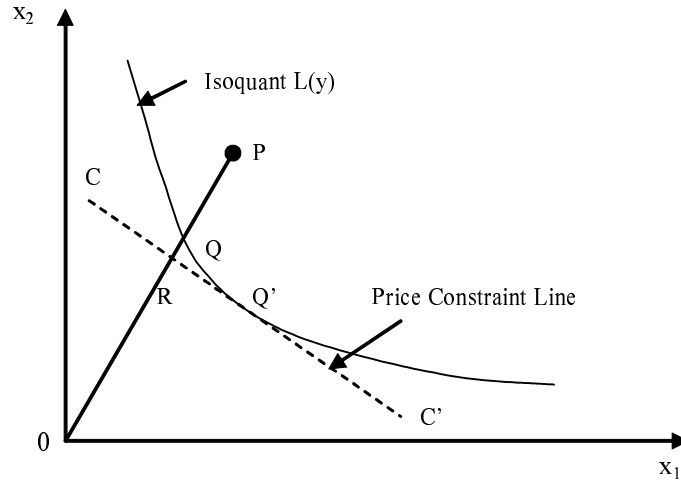
Figure 1: Technical efficiency versus economic efficiency with two inputs (Farrell, 1957)

firm like $Q$ or $Q'$ might use to produce an unit of output. The line $CC'$ whose slope is equal to the ratio of the prices of the two inputs represents the price constraint that all the firms must face. Farrell (1957) defined $OQ/OP$ as the technical efficiency level, $OR/OQ$ as the price (cost) efficiency and $OR/OP$ as the overall efficiency of the firm $P$. In DEA, the production frontier against which the technical efficiency of each DMU is derived is empirically constructed from observed DMUs, and thus without any assumption on the functional relation between inputs and outputs[9]. In other words, it is formed by a set of best practices (the most efficient DMUs) and the other DMUs are enveloped by this frontier, which explains the origin of the name "Data Envelopment Analysis" of this method. For the shake of brevity, hereafter I will use the term "efficiency" to refer to the technical efficiency and the term "efficiency frontier" to denote the production frontier.

## 2.2 DEA's basic model — CCR (1978)

### 2.2.1 The general formulation

Consider $n$ DMUs under evaluation that use $m$ inputs ($X$) to produce $s$ outputs ($Y$) with $X$ and $Y$ are semipositive[10]. The efficiency score $h_k$ assigned to the DMU $k$ is the solution

---

[9]In econometric methods, the efficient frontier is estimated by supposing a particular form of the production function (e.g., Cobb-Douglas, translog, etc.).

[10]The semipositivity signifies that all data are nonnegative but at least one component of every input and output vector is positive.

of the following optimizing system:

$$\max_{u,v} \quad h_k = \frac{\sum\limits_{r=1}^{s} u_r y_{rk}}{\sum\limits_{i=1}^{m} v_i x_{ik}} \tag{1}$$

$$\text{subject to:} \quad \frac{\sum\limits_{r=1}^{s} u_r y_{rj}}{\sum\limits_{i=1}^{m} v_i x_{ij}} \leq 1, j = 1, \dots, n \tag{2}$$

$$u_r, v_i \geq \varepsilon, r = 1, \dots, s; i = 1, \dots, m \tag{3}$$

where $k$ is the DMU under evaluation, $y_{rj}$ is the amount of the output $r$ of the DMU$_j$, $x_{ij}$ is the amount of the input $i$ of the DMU$_j$, $u_r$ and $v_i$ (also called "absolute weights") are the weights assigned respectively to the output $r$ and the input $i$, $\varepsilon$ is an infinitesimal positive number imposed to assure that no input or output is ignored in the optimization, $v_i x_{ij}$ and $u_r y_{rj}$ are called "virtual weights" of respectively the input $i$ and the output $r$ of the DMU$_j$.

Mathematically, the model's objective is to seek for the most favorable (positive) weight system associated with each input and each output which maximizes the weighted sum of the outputs over the weighted sum of the inputs of the DMU$_k$ ($h_k$), provided that this ratio does not exceed 1 for any DMU in the sample (reflected by constraint (2)). Given that the efficiency frontier contains efficient DMUs and envelopes inefficient ones, and that the efficiency level of each DMU is, by definition, the distance from its position to the efficiency frontier, it is natural to fix the maximal value of the objective function to unity[11]. Thus efficient DMUs will obtain a score of 1 and inefficient DMUs a score smaller than 1.

Conceptually, each DMU is free to choose its own combination of inputs and outputs so that it is as desirable as possible compared to other DMUs in the same category. Obviously, this combination must also be technically "feasible" for others, that is the efficiency level of any other DMU using this combination should not exceed the maximum attainable bounded by the efficiency curve (the constraint (2) is thus also applied to $j = 1, \dots, n$ with $j \neq k$). The idea is that if one DMU can not attain an efficiency rating of 100% under this set of weights, then it can never be attained from any other set. It should be noted that in practice, more constraints on weight systems can be imposed to take into account

---

[11]Mathematically, the maximal value of the objective function can be given any other number without changing the relative efficiency of the DMUs. The choice of unity is to assure the coherence between mathematical calculations and efficiency definitions.

specific preferences of decision-makers. This point will be illustrated further.

Alternatively, the DEA original problem can be formulated as the following system:

$$\min_{u,v} \quad h_k = \frac{\sum\limits_{i=1}^{m} v_i x_{ik}}{\sum\limits_{r=1}^{s} u_r y_{rk}} \tag{4}$$

$$\text{subject to} \quad \frac{\sum\limits_{i=1}^{m} v_i x_{ij}}{\sum\limits_{r=1}^{s} u_r y_{rj}} \geq 1, j = 1, \dots, n \tag{5}$$

$$u_r, v_i \geq \varepsilon, r = 1, \dots, s; i = 1, \dots, m \tag{6}$$

where the objective is to seek for optimal weights so as to minimize the ratio of the weighted sum of inputs to the weighted sum of outputs. The smaller this ratio, the better. In this case, efficient DMUs have a score of 1 and inefficient ones have a score greater than 1. Note however that the system (4-6) is less familiar within DEA's applications in finance than the system (1-3).

It is important to keep in mind that basic DEA models do provide a dichotomic classification, not a complete ranking of DMUs as all efficient DMUs have the same score equal to 1. Besides, efficiency or inefficiency of DMUs is solely relative to the sample under consideration. Hence, once the sample is modified, results may be very different.

### 2.2.2 The primal program

The optimizing systems (1-3) and (4-6) are fractional problems, non convex with fractional constraints, which are quite difficult to solve. According to Charnes & Cooper (1962, 1973) and Charnes et al. (1978), the fractional problem (1-3) (or 4-6) can be conveniently converted into an equivalent linear programming problem by normalizing the denominator to 1 and maximizing (minimizing) the nominator. By doing so, we obtain the input-oriented version (system (7-10)) and the output-oriented version (system (11-

14)) of the so-called CCR model — the seminal model of the DEA method:

$$\text{Input-oriented: } \max_{u,v} \quad h_k = \sum_{r=1}^{s} u_r y_{rk} \tag{7}$$

$$\text{subject to: } \quad \sum_{i=1}^{m} v_i x_{ik} = 1 \tag{8}$$

$$\sum_{r=1}^{s} u_r y_{rj} \leq \sum_{i=1}^{m} v_i x_{ij}, j = 1, \ldots, n \tag{9}$$

$$u_r, v_i \geq \varepsilon, r = 1, \ldots, s; i = 1, \ldots, m \tag{10}$$

or

$$\text{Output-oriented: } \min_{u,v} \quad h_k = \sum_{i=1}^{m} v_i x_{ik} \tag{11}$$

$$\text{subject to} \quad \sum_{r=1}^{s} u_r y_{rk} = 1 \tag{12}$$

$$\sum_{r=1}^{s} u_r y_{rj} \leq \sum_{i=1}^{m} v_i x_{ij}, j = 1, \ldots, n \tag{13}$$

$$u_r, v_i \geq \varepsilon, r = 1, \ldots, s; i = 1, \ldots, m \tag{14}$$

The input-oriented (output-oriented) version assumes that only inputs (outputs) can be adjusted, outputs (inputs) being fixed.

### 2.2.3 The dual program

According to linear programming theories, each primal program is associated with a dual program which provides the same optimal value of the objective function as the primal. The system (7-10) thus has a dual below:

$$\text{Input-oriented: } \min_{\theta,\lambda} \quad \theta \tag{15}$$

$$\text{subject to: } \quad \theta x_{ik} \geq \sum_{j=1}^{n} \lambda_j x_{ij}, i = 1, \ldots, m \tag{16}$$

$$y_{rk} \leq \sum_{j=1}^{n} \lambda_j y_{rj}, r = 1, \ldots, s \tag{17}$$

$$\lambda_j \geq 0, \theta \text{ unconstrained in sign} \tag{18}$$

with $\theta$ and $\lambda$ are dual variables. Note that $\theta$ can not, by construction, exceed unity[12].

---

[12]We can easily see that $\theta = 1$, $\lambda_k = 1$, $\lambda_j = 0$ ($j \neq k$) is a feasible solution to (15-18). Hence, the optimal value of $\theta$ can not be greater than 1. Besides, the constraint (16) implies that $\theta$ must be positive as $X$ is

In a similar fashion, the dual of the system (11-14) is defined by:

$$\text{Output-oriented: } \max_{\eta, \gamma} \quad \eta \tag{19}$$

$$\text{subject to: } \quad x_{ik} \geq \sum_{j=1}^{n} \gamma_j x_{ij}, i = 1, \ldots, m \tag{20}$$

$$\eta y_{rk} \leq \sum_{j=1}^{n} \gamma_j y_{rj}, r = 1, \ldots, s \tag{21}$$

$$\gamma_j \geq 0, \eta \text{ unconstrained in sign} \tag{22}$$

where $\eta$ and $\gamma$ are dual variables and $\eta$ can not be, by construction, lower than 1.

In fact, the primal program can be solved directly to obtain the optimal efficiency score. However, the dual program is usually preferred for the following reasons. On the one hand, it is mathematically easier to find the optimal solution via the dual because of a considerable reduction of constraints: from $n + s + m + 1$ constraints in the primal to only $s + m$ constraints in the dual. This calculating parsimony is of particularly appealing when dealing with large samples. On the other hand, the dual formulation has an interesting economic interpretation. In economic terms, under the input-oriented form (output-oriented form), the dual looks for a feasible activity — a *virtual* DMU which is a linear combination of the best practices — that guarantees (uses) the output level $y_k$ (the input level $x_k$) of the $DMU_k$ in all components while using only a proportion of the $DMU_k$'s inputs $\theta x_{ik}$ (producing higher outputs than $DMU_k$'s outputs, $\eta y_{rk}$ with $\eta \geq 1$ ). Hence, $\theta$ (or $\eta$) is defined as a measure of the efficiency level of the $DMU_k$. Graphically, in the input-output plan depicted in figure 2, under the input-oriented or input contraction setting, $\theta$ of the DMU $A$ is the ratio $DC/DA$, with $C$ being the virtual DMU which serves as benchmark to measure the efficiency of $A$; under the output-oriented or output expansion setting, $\eta$ of the DMU $A$ is measured by $FH/AH$ with $F$ being the reference DMU for $A$ now.

In order to obtain the efficiency scores of $n$ DMUs, the optimizing system (primal or dual) must be run $n$ times with each time the DMU under evaluation changes.

**Theorem 1** *(Connexion between the CCR input-oriented version and the CCR output-oriented version) Let $(\theta^*, \lambda^*)$ be an optimal solution for the CCR input-oriented version. Then $(1/\theta^*, \lambda^*/\theta^*) = (\eta^*, \gamma^*)$ is optimal for the corresponding CCR output-oriented version. Similarly, if $(\eta^*, \gamma^*)$ is optimal for the CCR output-oriented version, then $(1/\eta^*, \gamma^*/\eta^*) = (\theta^*, \lambda^*)$ is optimal for the corresponding CCR output-oriented version (Seiford et al. 2004, p.17).*
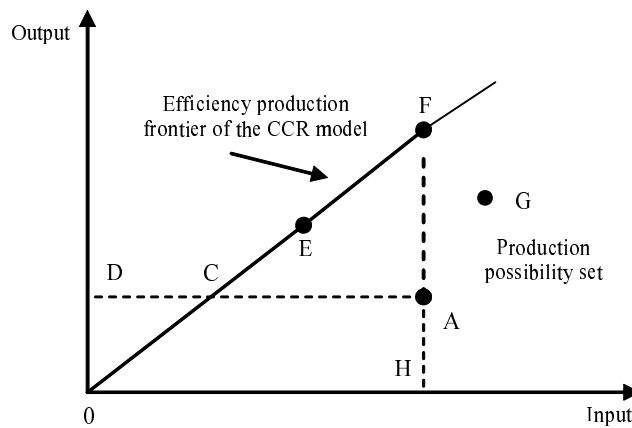
—————————————————————————

assumed to be semipositive.

Figure 2: CCR efficiency frontier

# 3   A DEA framework for hedge fund performance appraisal

In financial literature, funds' historical performance is often measured by the ratio of return to risk. Traditionally founded on the "mean-variance" basis, the evaluation dimension has been recently extended to the skewness (Stutzer index — Stutzer (2000)), to the skewness and the kurtosis (the modified Sharpe ratio — Gregoriou & Gueyie (2003)), even to the whole distribution of returns (Omega index — Keating & Shadwick (2002)) in an attempt to take into account the non normality features of returns. Despite this improvement, most of these measures are highly restrictive in the sense that they usually assume simplistic decision-making rules common to all investors.

Yet, there are suggestions that actual individual decisions differ significantly from theoretical formulations since they are much more complicated and quite specific to investors. Often there are more attributes to consider and for each investor, each attribute does not necessarily have the same priority level. While some investors are more concerned with central tendencies (mean, variance), others may care more about extreme values (skewness, kurtosis). One kind of such preferences is summarized by the positive preference for skewness first invoked by Arditti (1967) and then supported by Jean (1971), Kraus & Litzenberger (1976), Francis & Archer (1979), Scott & Horvath (1980), Kane (1982), Broihanne et al. (2004). It implies that individuals prefer portfolio A to portfolio B with higher mean return if both portfolios have the same variance, and if portfolio A has greater positive skewness, all higher moments being the same. In other words, individuals may attach more importance to the skewness than to the mean of returns. Despite the diversity of preferences for moments of returns, most measures assume the

12

same preference structure for all investors. Consider for example the modified Sharpe ratio (hereafter, M-Sharpe) (Gregoriou & Gueyie 2003) computed by the following equation:

$$M - Sharpe = \frac{\bar{r} - \overline{r_f}}{MVAR} = \frac{\bar{r} - \overline{r_f}}{W \left[ \mu - \left\{ z_c + \frac{1}{6} \left( z_c^2 - 1 \right) S + \frac{1}{24} \left( z_c^3 - 3z_c \right) K - \frac{1}{36} \left( 2z_c^3 - 5z_c \right) S^2 \right\} \sigma \right]}$$

(23)

where $\bar{r}$ is the mean return, $\overline{r_f}$ is the average risk-free rate, $W$ is the amount of portfolio at risk, $\mu$ is the mean return and naturally equal to $\bar{r}$, $\sigma$ is the standard deviation of returns, $S$ is the skewness, $K$ is the kurtosis excess, $z_c$ is the critical value for probability $(1 - \alpha)$ ($z_c = -1.96$ for a 95% probability), MVAR (modified value-at-risk) is introduced by Favre & Galeano (2002). According to Favre & Galeano (2002) and Gregoriou & Gueyie (2003), all investors are certainly concerned about the skewness and the kurtosis of returns but they share the same preference structure which is necessarily in the form of MVAR. This rigidity is not only restrictive but might bias significantly investors' choice of funds as their true evaluation criteria are not considered at all or considered but in a biased manner.

In addition to that, investors may need to take account of sales loads charged by the fund on their entrance into (front-end sales load) or/and on their exit of the fund (back-end or deferred sales load). Unlike management fees which are directly deducted from the fund's value, sales loads are charged on the net returns paid to investors[13]. As a result, a fund with good performance and a high percentage of loads is not necessarily more attractive than another fund which has lower performance but charges lower loads.

Moreover, as argued by McMullen & Strong (1998), Morey & Morey (1999) and Powers & McMullen (2000), investors may also be concerned about fund's performances over various time horizons (over the last year, the last 3 years, the last 5 years and sometimes the last 10 years). Such information is undoubtedly valuable as it provides much more informative insight into fund's perspective than the performance over only one horizon.

Furthermore, even when investors care about the return and the risk, or the performance of funds over only one horizon, it is often quite difficult to choose an absolutely

---

[13]Murthi et al. (1997), McMullen & Strong (1998), Tarim & Karan (2001), Choi & Murthi (2001) and Sengupta (2003) advocated incorporating also expense ratio (in percentage of fund assets, covering various operating expenses incurred by the fund management such as management fees, administrative fees, advisory fees) in evaluating fund performance. This element which is obviously necessary to appraise the performance of funds in a productivity perspective, i.e. their capacity to exploit efficiently input resources (fund expenses are considered here as a production factor), is irrelevant in this context where inputs and outputs are selection criteria chosen by investors. In this regard, investors are not likely concerned by these expenses as they are directly deducted before calculating funds' net asset value — the real value of investors' investments. Hence, such expenses are generally invisible to investors.

suitable measure among a wide range of existing measures in the literature. This difficulty is particularly true for the choice of risk and performance indicators because of inexistence or deficiency of mechanisms to validate empirically them. Consequently, investors are sometimes in need of considering simultaneously several measures. Here again, they do not necessarily share the same preferences for such and such measures.

Given these specificities in performance evaluating practices, the DEA's approach seems very appealing. In fact, the application of DEA into hedge fund performance appraisal can be made in two perspectives. The first one consists in evaluating the *productive* performance of funds where the latter are considered as a particular type of production units which employs multiple resources (risks, various operating expenses, turnover speed, etc.) to realize profits (returns). The second, which is in the spirit of Thompson et al. (1986), Tone (1999) and Powers & McMullen (2000), aims to assess funds as decision–making units whose inputs and outputs are evaluation criteria chosen by decision-makers. It is the second perspective that interests investors as DEA, with its broad flexibility, allows investors to tailor their own evaluation tools corresponding the most to their own preferences. Since each investor naturally has different risk aversion levels, performance objectives and other distinct constraints, the tailor-made possibility is essential to correctly screen fund.

In this context, the DEA method can be applied to evaluate either the "local" performance or the "global" performance of hedge funds. By the "local performance", I imply the performance measured by the weighted sum of several criteria of gain (or return) on the weighted sum of several criteria of risk and possibly certain types of expenses. In contrast, the term "global performance" denotes the performance synthesized from either several measures of "local" performance, or elementary performances over several temporal horizons. Within this framework, the application of DEA (in its basic form) raises four main questions: (1) how to choose inputs and outputs, (2) what version to choose (input-oriented or output-oriented), (3) how to deal with negative values in the inputs or/and the outputs if they exist[14], and (4) how to incorporate more specific preferences of investors into the mathematical formulation. If the first, the second and the fourth questions are relevant to any application fields, the third one is quite specific to data of returns and performances. These issues will be addressed successively in what follows.

---

[14]Inputs and outputs of DEA are originally assumed to be semipositive.

## 3.1 Evaluation criteria and the choice of inputs and outputs

Unlike applications of DEA in production fields where inputs and outputs are tangible elements, the choice of inputs and outputs is not straightforward when dealing with fund performance. Nevertheless, in a multiple criteria decision-making framework, it is logical to consider inputs as criteria that investors want to minimize and outputs as those they want to maximize. Hence, if investors seek to evaluate the funds' "local" performance, i.e. returns[15] to risks, the inputs can be (1) several measures of risk (standard deviation, kurtosis, beta, various measures of value-at-risk) over one (or several) horizon(s), (2) possibly the sales loads; the outputs can be composed of (1) several measures of returns (mean, skewness) — over one (or several) horizon(s). The difference between the configuration suggested here (to evaluate fund "local" performance) and that assumed by standard performance indicators is that according to the former, each investor knows perfectly his relevant evaluation criteria but does not know the functional relation between these criteria as well as the exact trade-off between them, which is not the case of the modified Sharpe ratio as previously described. The case where investors know the relative trade-off between these criteria will be discussed further.

Otherwise, if investors want to evaluate funds by considering several elementary performances simultaneously, they can calculate the *global* performance by (1) including in the outputs either the performances measured by the same technique on several periods, or the performances on the same period but measured by several indicators, (2) setting the input equal to one. It is important to notice that in this setting, all selection criteria are those investors want to maximize, they thus form exclusively DEA's outputs; meanwhile, there is no input. Assuming the presence of one input equal to 1 makes it possible to apply DEA without any modification of results. As explained earlier, we are in a basic configuration in which there is one input and several outputs and the quantity of each output is often "standardized" by the quantity of the input to obtain the unit outputs (per one unit of the input) in order to facilitate calculations. This setting is employed by Thompson et al. (1986) and Tone (1999).

Following the principles evoked above, each investor will determine, according to his own preferences, the inputs and outputs for DEA while complying with general rules as:

 – Inputs and outputs must be criteria indispensable to the appraisal of fund perfor-

---

[15]The term "return" should be understood here in broad sense. In traditional language of portfolio theories, the concept of return is always associated to the arithmetic mean of elementary returns over a given period. By "return", I imply in what follows any measure, in addition to the mean return, that is indicative of fund's expected returns such as the skewness.

mance.

– The number of inputs and outputs should be lower than the number of funds. In general, the number of funds should be at least three times larger than the number of inputs and outputs.

Any violation of these rules will lead to a deficiency of the discriminatory power of DEA. As a result, we risk obtaining an excessive number of dominant (efficient) funds whereas some of them are not rightly so[16].

## 3.2 Input-oriented or output-oriented versions ?

In general, when inputs and outputs are semipositive, the choice between the CCR input-oriented version and the CCR output-oriented version can be simply made at users' discretion following their preferences. Note that the input-oriented version (output-oriented version) assumes that outputs (inputs) are fixed, only inputs (outputs) can be adjusted. This assumption conditions the reference fund on the efficient frontier to which is compared the target fund and thus determines the distance between the former and the latter, this distance measuring the efficiency level of the latter. The theorem 1 describes the correspondence between the optimal solutions of the two versions. We can easily see that the two versions of the CCR model provides the same classification of inefficient DMUs[17] (efficient DMUs always obtain a full score of 1 under any version). Nevertheless, it is interesting to notice that all studies which apply DEA to evaluating fund performance adopted the input-oriented version whatever the DEA model is used. This popularity is undoubtedly due to the fact that this mathematical form shares the same logic as Markowitz's efficient frontier construction, that is to minimize the risks (inputs) for a defined level of returns (outputs).

However, when there are only outputs (inputs), the input-oriented (output-oriented) version is required as in this case, we assume the existence of one input (output) whose quantity is fixed equal to 1.

---

[16]The terms "dominant funds" and "efficient funds" will be used interchangeably hereafter to indicate funds having a full efficiency score of 1.

[17]It is important to specify that this equivalence between the input-oriented version and the output-oriented version is only valid under the constant returns-to-scale technology assumed by the CCR model.

## 3.3   Dealing with negative inputs and outputs

DEA models as originally designed require that inputs and outputs are semipositive, i.e. all inputs and all outputs are non negative and at least one input and one output are positive. In many application fields like production economics, negative inputs and outputs naturally make no sense. However, in fund performance appraisal context, it is likely that we sometimes have negative values like mean, skewness of returns, or some performance indicators, etc.

Although in the CCR model, or more generally in basic DEA models, inputs and outputs are systematically required to be semipositive, we can easily see that negative values in inputs and outputs are tolerated in following ways without any incidence on the solubility of DEA optimizing systems (Cooper et al. 2000, p.304-305):

– If there are at least one input and one output positif, either the input-oriented version or the output-oriented version can be used;

– If all outputs (inputs) are negative and at least one input (output) is positive, the input-oriented (output-oriented) version is required;

– If there is no (effective) input (output) and all outputs (inputs) are negative, the input-oriented (output-oriented) version is required;

– The case where all inputs and all outputs are negative at the same time, which is extremely rare in fund performance appraisal context, can not be dealt with within the DEA framework.

Note that in the second and the third cases, the optimal value of the objective function will be negative.

## 3.4   Taking account of investors' more specific preferences

The CCR model as presented earlier allows a quasi-absolute freedom in the determination of the weights $\{u, v\}$ so that each funds obtains a maximum score of efficiency, given its input and output level. Specifically, $\{u, v\}$ are only required to be equal to or greater than an infinitesimal positive number $\varepsilon$. This constraint is essential to assure that all selected evaluation criteria are considered in the evaluating process. Nevertheless, such flexibility level also implies that important, even excessive, weights can be assigned to the input(s) or/and the output(s) which make the funds as efficient as possible compared to others.

As a result, this setting is only plausible when investors have no idea about the trade-off between the selected criteria. When such information is available, it can be easily incorporated in DEA optimizing systems by restricting the absolute weights $\{u, v\}$ or the virtual weights $\{uy, vx\}$ associated with each input and each output.

An investor in full knowledge of the "price" range for each evaluation criterion — e.g. the coefficient of aversion to the mean, the variance, the skewness or the kurtosis of returns — can have recourse to constraints like:

$$u_r \leq \xi_r \tag{24}$$

$$v_i \geq \psi_i \tag{25}$$

$$\alpha_r \leq u_r \leq \beta_r \tag{26}$$

$$\gamma_i \leq v_i \leq \delta_i \tag{27}$$

An investor who knows more or less his personal or conventional trade-off or substitution rate between evaluation criteria can add following constraints into original DEA program:

$$\frac{u_r}{v_i} \leq \zeta \tag{28}$$

$$\frac{v_i}{v_{i+1}} \leq \kappa \tag{29}$$

$$\pi_i v_i + \pi_{i+1} v_{i+1} \leq v_{i+2} \tag{30}$$

An investor who wants to control the relative importance of each criterion in the performance appraisal process will formulate additional constraints on the virtual weights as follows:

$$a_r \leq \frac{u_r y_{rj}}{\sum_{r=1}^{s} u_r y_{rj}} \leq b_r \tag{31}$$

$$c_i \leq \frac{v_i x_{ij}}{\sum_{i=1}^{m} v_i x_{ij}} \leq d_i \tag{32}$$

$$u_r y_{rj} \leq u_{r+1} y_{r+1,j} \tag{33}$$

$$v_i x_{ij} \leq v_{i+1} x_{i+1,j} \tag{34}$$

where $\xi, \psi, \alpha, \beta, \zeta, \kappa, a, b, c, d$ are values pre-defined by investors to bound absolute and virtual weights.

There are certainly many other forms of additional constraints because of a broad

variety of investors' preferences. The constraints mentioned above are to give examples of "transcribing" more specific preferences into mathematical formulations. A numerical illustration will be provided further.

This possibility of exerting a direct control on the relative importance of each evaluation criteria in assessing fund performance, along with the choice of evaluation criteria (inputs and outputs), makes it possible for each investor to conceive a customized measure corresponding to his preferences. With such quality, the DEA approach is an efficient and complementary tool to other existing measures.

# 4    Illustrative applications

## 4.1    Data

To illustrate the use of DEA in assessing the performance of hedge funds, I used a sample of 38 hedge funds belonging to the category Equity Hedge[18]. Data includes 60 monthly returns covering the period of January 2000 to December 2004. Table 1 reports some descriptive statistics of these funds.

As we can see, return distributions of many funds show highly positive (negative) skewness signifying higher probability of extreme positive (negative) values compared to that implied by the normal distribution. Besides, many of them possess high kurtosis excess, which indicates more returns close to the central value but also more regular large positive or negative returns than in a normal distribution. The normality assumption of return distributions is tested by means of three tests: Shapiro-Wilk, Kolmogorov-Smirnov, and Jarque-Bera. Results provided by the Shapiro-Wilk and Jarque-Bera tests are quite similar although they are rather different from those provided by the Kolmogorov-Smirnov test. This divergence is likely due to the sample's limit size as the Kolmogorov-Smirnov test is more appropriate to large samples. According to the Shapiro-Wilk test, documented as the most reliable for small samples, the normality assumption is rejected in 14 out of 38 cases at the confidence level of 95%. These findings imply much higher return or risk of these funds than those approximated under normality assumption. They thus highlight the importance of incorporating moments of order higher than the mean and the variance when appraising funds' return and risk profiles.

---

[18]These 38 funds are extracted from a database provided by the company Standard & Poor's. Equity Hedge covers several different strategies whose investments are focused on the equity markets. Its two large categories are Global Macro and Relative Value.

Table 1: Descriptive statistics

| Funds | Min (%) | Max (%) | Me (%) | SD (%) | SK | KU | S-W | K-S | J-B |
|---|---|---|---|---|---|---|---|---|---|
| 1 | -6.16 | 10.20 | 0.68 | 3.77 | 0.34 | -0.35 | 0.98 | 0.11 | 1.43 |
| 2 | -7.89 | 7.69 | 0.16 | 3.33 | -0.07 | -0.25 | 0.99 | 0.06 | 0.20 |
| 3 | -12.51 | 19.66 | -0.30 | 5.59 | 0.32 | 1.86 | 0.95** | 0.12 | 9.72*** |
| 4 | -7.25 | 5.63 | 0.25 | 3.04 | -0.27 | -0.49 | 0.98 | 0.07 | 1.31 |
| 5 | -11.37 | 11.95 | 0.10 | 4.65 | -0.10 | 0.09 | 0.99 | 0.06 | 0.12 |
| 6 | -6.50 | 5.92 | 0.06 | 2.34 | -0.36 | 0.29 | 0.98 | 0.08 | 1.50 |
| 7 | -14.67 | 24.36 | -0.11 | 6.23 | 0.91 | 3.37 | 0.95*** | 0.11 | 36.60*** |
| 8 | -22.96 | 33.89 | -0.18 | 8.86 | 0.71 | 2.69 | 0.96** | 0.06 | 23.09*** |
| 9 | -7.87 | 8.59 | -0.01 | 3.92 | -0.01 | -0.43 | 0.98 | 0.07 | 0.46 |
| 10 | -11.84 | 13.05 | 0.03 | 5.40 | 0.16 | -0.12 | 0.99 | 0.08 | 0.29 |
| 11 | -8.19 | 17.11 | 1.08* | 4.95 | 1.18 | 2.13 | 0.92*** | 0.16* | 25.40*** |
| 12 | -13.49 | 9.11 | 0.01 | 4.35 | -0.43 | 0.66 | 0.98 | 0.08 | 2.98 |
| 13 | -6.77 | 7.23 | -0.57 | 3.36 | 0.12 | -0.75 | 0.98 | 0.07 | 1.53 |
| 14 | -40.85 | 19.45 | -0.73 | 9.02 | -1.42 | 5.88 | 0.90*** | 0.15 | 106.6*** |
| 15 | -12.04 | 14.17 | 0.23 | 5.53 | -0.05 | 0.06 | 0.98 | 0.07 | 0.03 |
| 16 | -5.76 | 6.58 | 0.24 | 2.83 | 0.13 | -0.22 | 0.99 | 0.06 | 0.29 |
| 17 | -7.10 | 6.27 | 0.33 | 3.25 | -0.15 | -0.49 | 0.98 | 0.06 | 0.83 |
| 18 | -6.33 | 5.94 | 0.15 | 2.65 | 0.00 | -0.33 | 0.99 | 0.08 | 0.28 |
| 19 | -7.21 | 8.18 | 0.25 | 3.43 | 0.17 | -0.46 | 0.99 | 0.06 | 0.81 |
| 20 | -9.90 | 14.45 | 0.12 | 4.54 | 0.25 | 0.71 | 0.99 | 0.05 | 1.91 |
| 21 | -6.81 | 9.77 | 0.64* | 2.75 | 0.48 | 2.08 | 0.96* | 0.10 | 13.19*** |
| 22 | -9.20 | 7.57 | -0.23 | 3.95 | -0.09 | -0.38 | 0.99 | 0.05 | 0.44 |
| 23 | -5.31 | 6.82 | 0.33 | 2.48 | 0.62 | 0.59 | 0.96** | 0.15 | 4.66*** |
| 24 | -13.75 | 15.03 | -0.81 | 5.72 | -0.10 | 0.30 | 0.98 | 0.09 | 0.33 |
| 25 | -9.78 | 19.59 | 0.09 | 5.46 | 1.35 | 3.85 | 0.90*** | 0.12 | 55.13*** |
| 26 | -16.34 | 25.90 | 1.34 * | 5.31 | 1.35 | 10.02 | 0.72*** | 0.24*** | 269.3*** |
| 27 | -1.24 | 5.45 | 0.52*** | 1.10 | 1.64 | 5.81 | 0.88 *** | 0.17* | 111.2*** |
| 28 | -2.37 | 15.86 | 0.74 ** | 2.39 | 4.36 | 27.27 | 0.63 *** | 0.19** | 2050*** |
| 29 | -15.48 | 22.37 | 0.64 | 5.07 | 1.23 | 7.06 | 0.85 *** | 0.16* | 139.7*** |
| 30 | -13.76 | 17.90 | -0.35 | 5.29 | 0.50 | 1.86 | 0.97 | 0.09 | 11.22*** |
| 31 | -14.30 | 17.93 | 0.28 | 5.00 | 0.21 | 2.69 | 0.95** | 0.10 | 18.58*** |
| 32 | -6.93 | 11.54 | 0.00 | 3.15 | 0.73 | 2.47 | 0.96** | 0.08 | 20.64*** |
| 33 | -7.88 | 11.53 | 0.08 | 4.07 | 0.58 | 0.71 | 0.97 | 0.08 | 4.57 |
| 34 | -7.12 | 8.67 | -0.07 | 4.12 | 0.20 | -0.64 | 0.97 | 0.08 | 1.42 |
| 35 | -5.68 | 10.93 | 1.10*** | 2.97 | 0.38 | 1.30 | 0.96 * | 0.12 | 5.66* |
| 36 | -10.13 | 7.95 | 0.58 | 3.75 | -0.29 | -0.05 | 0.98 | 0.08 | 0.82 |
| 37 | -6.48 | 11.01 | 0.38 | 2.60 | 0.71 | 4.25 | 0.93 *** | 0.10 | 50.14*** |
| 38 | -9.93 | 12.48 | -0.13 | 4.10 | 0.04 | 0.71 | 0.98 | 0.07 | 1.26 |

Me = Mean, SD = Standard deviation, SK = Skewness, KU = Kurtosis excess relatively to the normal distribution. S-W = Shapiro-Wilk, K-S = Kolmogorov-Smirnov, J-B = Jarque-Bera are normality tests on return distributions. ***, **, ** denote the rejection of the normality assumption at respectively the 99%, 95% and 90% confidence levels.

## 4.2 Assessing hedge fund local performance

### 4.2.1 Settings

Due to unavailable data on sales loads charged by the funds in the sample, illustrations are limited to considering their historical return and risk profiles. Since the distribution of hedge fund returns is documented as usually non gaussian, it is important to incorporate these features into the selection of evaluation criteria (DEA's inputs and outputs). Several settings are likely.

The first setting assumes the case where investors have a positive preference for odd moments and a negative preference for even moments. Given these preference, it is logical to include in the inputs the standard deviation and the kurtosis of returns, and in the outputs the mean and the skewness. In this configuration, the problem of negative outputs raises. More specifically, 11 out of 38 funds under consideration have a negative mean, 12 funds have a negative skewness, 4 funds among them have simultaneously negative mean and negative skewness.

The second setting is designed in the spirit of Gregoriou et al. (2005) and Kooli et al. (2005), following which it is more clever to reason in terms of partial variations. As documented in the literature, investors are likely to be averse only to volatility under the Minimum Accepted Return (MAR)[19], which are called lower variations. In contrast, they appreciate volatility above this value, which are called upper variations. Thus, the composition of inputs and outputs can be determined in the following manner. The inputs include lower mean, lower semi-standard deviation, lower semi-skewness and lower semi-kurtosis which are obtained from returns lower than the MAR represented by the average rate over the period january 2000 to december 2004 of the US 3-month T-bill. The outputs contain upper mean, upper semi-standard deviation, upper semi-skewness and upper semi-kurtosis obtained from returns greater than the MAR. In addition to his financial finesse, this configuration has the clear-cut advantage to avoid the problem of negative inputs and outputs.

Now assume furthermore that investors are more concerned for extreme values than central ones. Hence, they naturally pay more attention to the skewness and the kurtosis than to the mean and the standard deviation. Mathematically, they will require that the contribution of the upper (lower) skewness and kurtosis to the efficiency score of the fund must be greater than or equal to the contribution of the upper (lower) mean and

---

[19]The determination of the Minimum Accepted Return is purely subjective and specific to each investor. It can be a risk-free rate or any rate required by investors.

standard deviation. This preference can be taken into consideration by adding four more constraints on virtual weights into the optimization system:

$$y_{3j}u_3 \geqslant y_{1j}u_1 \quad ; \quad x_{3j}v_3 \geqslant x_{1j}v_1$$
$$y_{3j}u_3 \geqslant y_{2j}u_2 \quad ; \quad x_{3j}v_3 \geqslant x_{2j}v_2$$
$$y_{4j}u_4 \geqslant y_{1j}u_1 \quad ; \quad x_{4j}v_4 \geqslant x_{1j}v_1$$
$$y_{4j}u_4 \geqslant y_{2j}u_2 \quad ; \quad x_{4j}v_4 \geqslant x_{2j}v_2$$

where $y_{1j}, y_{2j}, y_{3j}, y_{4j}$ are the amount of upper mean, upper standard deviation, upper skewness and upper kurtosis of the fund $j$ under consideration; $x_{1j}, x_{2j}, x_{3j}, x_{4j}$ are the amount of its lower mean, lower standard deviation, lower skewness and lower kurtosis; $u_1, u_2, u_3, u_4, v_1, v_2, v_3, v_4$ are the weights associated respectively with these outputs and inputs.

Otherwise, if investors are more or less markowitzian, i.e. they rely essentially on the mean and standard deviation to assess fund performance, the following constraints are necessary so that this preference is incorporated:

$$y_{1j}u_1 \geqslant y_{3j}u_3 \quad ; \quad x_{1j}v_1 \geqslant x_{3j}v_3$$
$$y_{1j}u_1 \geqslant y_{4j}u_4 \quad ; \quad x_{1j}v_1 \geqslant x_{4j}v_4$$
$$y_{2j}u_2 \geqslant y_{3j}u_3 \quad ; \quad x_{2j}v_2 \geqslant x_{3j}v_3$$
$$y_{2j}u_2 \geqslant y_{4j}u_4 \quad ; \quad x_{2j}v_2 \geqslant x_{4j}v_4$$

The third setting illustrates another case where investors need to reconcile funds' local performance over several horizons, from a long period to a more recent period in the past. To this end, DEA inputs are modeled by the MVAR (described by the denominator of equation 23) representing the loss limits over three horizons: 1 year, 3 years and 5 years; outputs are the mean returns over these three horizons. Again, many cases of negative outputs are found: 12 cases over the one-year horizon, 22 cases over the three-year horizon and 11 cases over the five-year horizon, among them 5 funds have all negative outputs.

It is important to keep in mind that the above settings are only some standard configurations used by investors. Given the diversity of investors' preferences, many other configurations are also expected.

### 4.2.2  Choice of CCR version

After inputs and outputs corresponding to investors' preferences are specified, the next step consists in running the foregoing inputs and outputs under the CCR model. Then what version to choose, input-oriented or output-oriented? Following principles highlighted in the section 3.3, we are constraint to adopt the input-oriented version for the first and the third settings where outputs are sometimes all negative. Regarding the second setting, either version is possible. However, in this study, the input-oriented version is chosen for all settings. Its primal and dual programs are described respectively by the systems (7-10) and (15-18).

Note that the weights assigned to each output and input are constrained to be equal to or greater than 0.001 ($\varepsilon = 0.001$)[20] to assure that all criteria are considered in the optimization program.

### 4.2.3  Results

Table 2 displays detailed results on DEA score, absolute weights ($u$, $v$) and virtual weights ($uy$, $vx$) obtained under a CCR input-oriented setting with mean and skewness as outputs, standard deviation and kurtosis as inputs. Funds with negative scores are those having simultaneously negative mean and negative skewness. Given the difference of measure scale between mean, standard deviation on the one hand and skewness, kurtosis on the other hand, virtual weights rather than absolute weights are more informative about key factors (inputs and outputs) that make some funds dominant compared to others in the sample. Each of the five funds qualified as dominant (1, 11, 27, 28, 35) has its own combination of evaluation criteria to attain the full efficiency. For fund 27 and fund 35, the virtual weights associated with the mean and the standard deviation are much higher than those associated with the skewness and the kurtosis. By referring to the statistics of returns given in table 1, we find that they effectively have fairly high mean and small standard deviation in comparison with the others. Their profiles are thus well adapted to markowitzian investors. By contrast, the dominance of fund 28 is primarily due to its positive skewness. In fact, this fund has the highest positive skewness in the sample. With fund 1, the dominance is mainly based on the mean and the kurtosis while with fund 11, dominant factors are the mean, the skewness and the kurtosis. These findings imply that not all dominant funds are necessarily adapted to an investor having a precise

---

[20]In fact, all calculations were already tested with several values of $\varepsilon$: 0, 0.0001, 0.001 and 0.01. However, performance scores changed very slightly while the relative rank between funds remains unchanged. Thus, $\varepsilon$ was fixed to be equal to 0.001 to facilitate result presentation.

preference. Consequently, when no additional constraint is formulated like in this setting, an investor who is not completely indifferent among evaluation criteria should identify the factors determining the efficiency of dominant funds and select only those whose profiles correspond the most to his preferences.

Results on DEA scores across various settings are summarized in table 3. Note that in the first and the third settings (respectively in columns 2 and 6), funds with negative scores are those whose all outputs are negative. Several points are noteworthy. In general, results are rather sensitive to the specification of evaluation criteria and supplementary constraints. Not only the number of dominant funds varies (from 1 to 5) but also dominant members differ across settings. Look at for example fund 26 which is qualified as dominant only when its returns and risks over three horizons are considered simultaneously.

Related to the second setting, as would be expected, the introduction of additional constraints on virtual weights naturally deteriorates efficiency scores and the short list of dominant funds becomes more selective. When preferences for extreme values (represented by the skewness and the kurtosis) are explicitly formulated, only fund 28 (among five funds 1, 8, 11, 25, 28 qualified as dominant without any additional constraints) satisfies this requirement. Similarly, when more importance is explicitly attached to central values (represented by the mean and the standard deviation), there are only three funds 8, 11, 28 in the dominant list. These results highlight the importance of correct specification of relevant DEA inputs and outputs as well as additional constraints which reflect best investors' evaluation preferences.

At empirical level, one may notice persistent dominance of several funds across settings like the case of fund 28, which stays dominant whatever preferences are considered. This feature can be regarded as a sign of the robustness of fund 28's performance relatively to other funds in the sample.

Since we are examining funds' local performance without sales loads, it could be interesting at this point to contrast DEA results in the first and the second settings with fund rankings provided by the traditional Sharpe ratio (Sharpe 1966) and the M-Sharpe ratio. The latter is computed following equation 23 while the former is calculated by the formula below:

$$Sharpe = \frac{\bar{r} - \bar{r_f}}{\sigma} \tag{35}$$

where $\bar{r}$ is the average return of the fund, $\bar{r_f}$ is the average risk-free rate approximated here by the US 3-month T-bill rate, $\sigma$ is the standard deviation of fund returns. Note that

Table 2: Performance with standard deviation-kurtosis as inputs, mean-skewness as outputs

| Funds | Score[a] | Absolute weights $(u, v)$[b] | | | | Virtual weights $(uy, vx)$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Me | SK | SD | KU | Me | SK | SD | KU |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| 1 | 1 | 147.427 | 0.001 | 0.167 | 0.375 | 0.9997 | 0.0003 | 0.0063 | 0.9937 |
| 2 | 0.228 | 142.003 | 0.001 | 0.161 | 0.361 | 0.2284 | -0.0001 | 0.0054 | 0.9946 |
| 3 | 0.286 | 0.001 | 0.891 | 0.001 | 0.206 | 0.0000 | 0.2864 | 0.0001 | 0.9999 |
| 4 | 0.389 | 155.891 | 0.001 | 0.176 | 0.396 | 0.3888 | -0.0003 | 0.0054 | 0.9946 |
| 5 | 0.127 | 126.330 | 0.001 | 0.001 | 0.323 | 0.1270 | -0.0001 | *0.0000* | *1.0000* |
| 6 | 0.077 | 119.256 | 0.001 | 0.136 | 0.303 | 0.0774 | -0.0004 | 0.0032 | 0.9968 |
| 7 | 0.618 | 0.001 | 0.681 | 0.001 | 0.157 | 0.0000 | 0.6181 | 0.0001 | 0.9999 |
| 8 | 0.537 | 0.001 | 0.762 | 0.001 | 0.176 | 0.0000 | 0.5374 | 0.0001 | 0.9999 |
| 9 | -0.000 | 0.001 | 0.001 | 25.456 | 0.001 | -0.0000 | -0.0000 | 0.9974 | 0.0026 |
| 10 | 0.240 | 0.001 | 1.507 | 0.001 | 0.348 | 0.0000 | 0.2403 | 0.0001 | 0.9999 |
| 11 | 1 | 66.707 | 0.237 | 1.281 | 0.182 | 0.7193 | 0.2807 | 0.0635 | 0.9365 |
| 12 | 0.006 | 106.752 | 0.001 | 0.122 | 0.272 | 0.0062 | -0.0004 | 0.0053 | 0.9947 |
| 13 | 0.222 | 0.001 | 1.925 | 0.001 | 0.444 | 0.0000 | 0.2219 | *0.0000* | *1.0000* |
| 14 | -0.001 | 0.001 | 0.001 | 10.984 | 0.001 | 0.0000 | -0.0014 | 0.9911 | 0.0089 |
| 15 | 0.295 | 127.735 | 0.001 | 0.001 | 0.327 | 0.2947 | -0.0001 | 0.0001 | 0.9999 |
| 16 | 0.347 | 122.726 | 0.436 | 2.357 | 0.336 | 0.2899 | 0.0566 | 0.0668 | 0.9332 |
| 17 | 0.514 | 155.511 | 0.001 | 0.176 | 0.396 | 0.5140 | -0.0002 | 0.0057 | 0.9943 |
| 18 | 0.215 | 146.828 | 0.001 | 0.166 | 0.373 | 0.2148 | 0.0000 | 0.0044 | 0.9956 |
| 19 | 0.412 | 126.397 | 0.556 | 0.001 | 0.394 | 0.3183 | 0.0937 | *0.0000* | *1.0000* |
| 20 | 0.297 | 0.001 | 1.168 | 0.001 | 0.269 | 0.0000 | 0.2966 | *0.0000* | *1.0000* |
| 21 | 0.684 | 68.216 | 0.508 | 15.009 | 0.116 | 0.4380 | 0.2463 | 0.4122 | 0.5878 |
| 22 | -0.000 | 0.001 | 0.001 | 25.250 | 0.001 | -0.0000 | -0.0001 | 0.9974 | 0.0026 |
| 23 | 0.793 | 0.001 | 1.289 | 6.883 | 0.231 | 0.0000 | 0.7929 | 0.1706 | 0.8294 |
| 24 | -0.000 | 0.001 | 0.001 | 17.422 | 0.001 | -0.0000 | -0.0001 | 0.9967 | 0.0033 |
| 25 | 0.888 | 0.001 | 0.659 | 3.518 | 0.118 | 0.0000 | 0.8877 | 0.1922 | 0.8078 |
| 26 | 0.695 | 29.615 | 0.221 | 6.516 | 0.050 | 0.3973 | 0.2977 | 0.3461 | 0.6539 |
| 27 | 1 | 193.390 | 0.001 | 90.268 | 0.001 | 0.9984 | 0.0016 | 0.9912 | 0.0088 |
| 28 | 1 | 33.048 | 0.173 | 40.581 | 0.001 | 0.2437 | 0.7563 | 0.9697 | 0.0303 |
| 29 | 0.592 | 0.001 | 0.482 | 2.576 | 0.086 | 0.0000 | 0.5918 | 0.1306 | 0.8694 |
| 30 | 0.448 | 0.001 | 0.891 | 0.001 | 0.206 | 0.0000 | 0.4482 | 0.0001 | 0.9999 |
| 31 | 0.212 | 60.499 | 0.215 | 1.162 | 0.165 | 0.1664 | 0.0460 | 0.0581 | 0.9419 |
| 32 | 0.639 | 0.001 | 0.871 | 4.652 | 0.156 | 0.0000 | 0.6393 | 0.1463 | 0.8537 |
| 33 | 0.675 | 0.001 | 1.170 | 0.001 | 0.270 | 0.0000 | 0.6747 | *0.0000* | *1.0000* |
| 34 | 0.361 | 0.001 | 1.840 | 0.001 | 0.424 | 0.0000 | 0.3607 | *0.0000* | *1.0000* |
| 35 | 1 | 91.251 | 0.001 | 31.662 | 0.014 | 0.9996 | 0.0004 | 0.9390 | 0.0610 |
| 36 | 0.773 | 132.376 | 0.001 | 0.150 | 0.337 | 0.7735 | -0.0003 | 0.0056 | 0.9944 |
| 37 | 0.502 | 55.541 | 0.414 | 12.220 | 0.094 | 0.2083 | 0.2940 | 0.3175 | 0.6825 |
| 38 | 0.044 | 0.001 | 1.170 | 0.001 | 0.270 | 0.0000 | 0.0439 | *0.0000* | *1.0000* |

Note: Me = Mean, SK = Skewness, SD = Standard deviation, KU = Kurtosis. Values in italics are approximative.

[a]Funds with negative scores are those whose mean and skewness are simultaneously negative.

[b]$u$ and $v$ are required to be equal to or greater than 0.001 ($\varepsilon = 0.001$).

## Table 3: Local performance

| Funds | Standard[a] moments | DEA scores Partial moments[b] | | | Horizons[c] | Rank Sharpe | Rank M-Sharpe |
| | | Standard | Preference SK & KU | Preference Me & SD | | | |
|---|---|---|---|---|---|---|---|
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1 | 1 | 1 | 0.52 | 0.93 | 0.46 | 7 | 7 |
| 2 | 0.23 | 0.71 | 0.38 | 0.69 | 0.46 | 18 | 18 |
| 3 | 0.29 | 0.83 | 0.69 | 0.71 | -0.00 | 33 | 34 |
| 4 | 0.39 | 0.66 | 0.37 | 0.64 | 0.19 | 14 | 15 |
| 5 | 0.13 | 0.78 | 0.36 | 0.77 | 0.08 | 21 | 20 |
| 6 | 0.08 | 0.64 | 0.35 | 0.58 | 0.93 | 29 | 28 |
| 7 | 0.62 | 0.86 | 0.67 | 0.81 | 0.21 | 27 | 29 |
| 8 | 0.54 | 1 | 0.62 | 1 | 0.11 | 25 | 26 |
| 9 | 0.00 | 0.72 | 0.44 | 0.67 | 0.22 | 28 | 27 |
| 10 | 0.24 | 0.85 | 0.36 | 0.84 | 0.14 | 23 | 23 |
| 11 | 1 | 1 | 0.60 | 1 | 0.83 | 5 | 4 |
| 12 | 0.01 | 0.68 | 0.32 | 0.66 | 0.24 | 26 | 25 |
| 13 | 0.22 | 0.79 | 0.38 | 0.74 | -0.00 | 38 | 38 |
| 14 | 0.00 | 0.62 | 0.21 | 0.61 | 1 | 34 | 31 |
| 15 | 0.29 | 0.89 | 0.45 | 0.75 | 0.10 | 17 | 17 |
| 16 | 0.35 | 0.72 | 0.41 | 0.70 | 0.24 | 16 | 16 |
| 17 | 0.51 | 0.80 | 0.32 | 0.78 | 0.63 | 12 | 12 |
| 18 | 0.21 | 0.71 | 0.39 | 0.68 | 0.14 | 22 | 21 |
| 19 | 0.41 | 0.78 | 0.44 | 0.76 | 0.19 | 15 | 14 |
| 20 | 0.30 | 0.85 | 0.45 | 0.84 | 1 | 19 | 19 |
| 21 | 0.68 | 0.75 | 0.44 | 0.72 | 0.70 | 6 | 6 |
| 22 | 0.00 | 0.85 | 0.31 | 0.82 | -0.00 | 36 | 35 |
| 23 | 0.79 | 0.88 | 0.37 | 0.83 | 0.26 | 11 | 11 |
| 24 | 0.00 | 0.72 | 0.46 | 0.70 | 0.48 | 37 | 37 |
| 25 | 0.89 | 1 | 0.72 | 0.97 | 0.04 | 20 | 22 |
| 26 | 0.70 | 0.72 | 0.65 | 0.59 | 1 | 4 | 5 |
| 27 | 1 | 0.92 | 0.92 | 0.81 | 0.96 | 2 | 2 |
| 28 | 1 | 1 | 1 | 1 | 1 | 3 | 1 |
| 29 | 0.59 | 0.75 | 0.59 | 0.66 | 0.41 | 9 | 8 |
| 30 | 0.45 | 0.89 | 0.45 | 0.85 | 0.35 | 35 | 36 |
| 31 | 0.21 | 0.83 | 0.43 | 0.82 | 0.71 | 13 | 13 |
| 32 | 0.64 | 0.70 | 0.58 | 0.66 | 0.17 | 30 | 32 |
| 33 | 0.67 | 0.94 | 0.45 | 0.91 | 0.26 | 24 | 24 |
| 34 | 0.36 | 0.91 | 0.44 | 0.88 | -0.00 | 31 | 30 |
| 35 | 1 | 0.76 | 0.52 | 0.74 | 1 | 1 | 3 |
| 36 | 0.77 | 0.84 | 0.30 | 0.82 | 0.37 | 8 | 9 |
| 37 | 0.50 | 0.63 | 0.58 | 0.59 | 0.44 | 10 | 10 |
| 38 | 0.04 | 0.80 | 0.43 | 0.76 | -0.00 | 32 | 33 |
| Rank correlation (Sharpe & M-Sharpe) | | | | | | | 0.995 |

Note: Results are obtained from the CCR input-oriented version with $\varepsilon = 0.001$. Funds with negative scores are those whose all outputs are simultaneously negative. Me = Mean, SK = Skewness, SD = Standard deviation, KU = Kurtosis.

[a]In the first setting, inputs are standard deviation and kurtosis, outputs are mean and skewness.

[b]In the second setting, inputs are composed of lower mean, lower semi standard deviation, lower skewness and lower kurtosis; outputs contain upper mean, upper semi standard deviation, upper skewness and upper kurtosis.

[c]In the third setting, inputs include the MVAR over the previous year, the 3 previous years and the 5 previous years; outputs include mean returns over these three periods.

the Sharpe ratio is based on the mean-variance paradigm while the modified Sharpe ratio takes account of the skewness and the kurtosis.

Fund rankings according to these two ratios are reported in the columns 7 and 8 of table 3. Several main observations can be drawn from these results. We can see easily that despite differences in the approach taken by the two measures, fund rankings are surprisingly quite similar, both in terms of correlation coefficient (0.995) and in terms of direct contrasting from fund to fund. Does this strong similarity imply that the return distribution of all funds is quite close to the normal one? The answer according to the Shapiro-Wilk normality test is rather negative because the normality assumption is rejected in 14 among 38 cases at the confidence level of 95% (see table 1). However, finding explanations to such problem is beyond the scope of this paper.

Regarding the connection of DEA classifications (except for the third setting) with Sharpe and M-Sharpe rankings, the results show that most dominant funds are generally among the seven funds the most highly ranked by Sharpe and M-Sharpe ratios. Nevertheless, funds 8 (dominant once) and 25 (dominant twice) in the second setting are only placed respectively at the $25^{th}$ and $20^{th}$ rank by Sharpe, $26^{th}$ and $22^{th}$ by M-Sharpe. This disfavor is certainly related to the slightly negative mean of fund 28 (-0.18%) and to the quite low positive mean of fund 25 (0.09%). A closer examination of their return distributions reveals much wider dispersal of returns and higher frequency of extreme positive values in these two distributions than in those of other funds ranked before them by Sharpe and M-Sharpe ratios. It is undoubtedly the reason why these funds are highly praised by the second setting of DEA. An investor who likes good surprises would find his interests in these profiles. Yet, if he used only Sharpe and M-Sharpe ratios, he would have missed his chance, at least in the case of this sample. Such result provides evidence that DEA can be an efficient supplementary tool to assist investors in selecting correctly funds satisfying their preferences.

## 4.3 Assessing hedge fund global performance

### 4.3.1 Settings

As argued previously, investors may sometimes want to evaluate local performance of funds (1) on several horizons simultaneously or (2) by using several measures at the same time. In these cases, how will they reconcile between elementary performances ? On which basis they can assign a final note to each fund so as to rank them ? This choice is

particularly difficult when elementary performances provide divergent rankings of funds. Meanwhile, by means of optimizing the weighted sum of elementary performances, DEA offers an aggregate measure allowing investors to identify funds having the best combination of these performances. In other words, by combining multiple performance criteria simultaneously, DEA provides an exhaustive image of funds. In order to illustrate the first setting, the M-Sharpe performance ratio over three horizons — the previous year, the three previous years, the five previous years — are used as elementary performances. In the second setting, three performance measures which consider the non normal features of returns are selected, namely modified Stutzer index (Stutzer 2000, Kaplan & Knowles 2001), M-Sharpe ratio and Omega index (Keating & Shadwick 2002). The formulas to compute the modified Stutzer and the Omega indices are given below:

$$\text{M-Stutzer} = sign(\bar{r})\sqrt{2\text{Stutzer}} \text{ with} \tag{36}$$

$$\text{Stutzer} = \max_{\theta}\left[-\ln\frac{1}{T}\sum_{t=1}^{T}e^{\theta(r_t-r_{ft})}\right] \tag{37}$$

$$\text{Omega} = \frac{\int_{\tau}^{\infty}[1-F(r)]dr}{\int_{-\infty}^{\tau}F(r)dr} \tag{38}$$

where $r_t$ is fund return on month $t$, $sign(\bar{r})$ is the sign of the mean return, $\theta$ is a negative number, $T$ is the number of monthly returns, $r_{ft}$ is risk-free rate (approximated by the US 3-month T-bill rate) on month $t$, $\tau$ is the MAR pre-determined by investors (approximated by the US 3-month T-bill rate's average over the study period). It is necessary to specify that the modified Stutzer index (hereafter, M-Stutzer) considers up to the skewness of returns, the M-Sharpe ratio takes into account both the skewness and the kurtosis while the Omega index regards the whole (empirical) distribution of returns.

These two settings are characterized as having only outputs. Since there are no inputs, it is possible to assume existence of one input equal to one so that DEA can be applied. Given this feature, the input-oriented version is required. Besides, investors are also assumed to be indifferent among horizons and performance indicators so that no additional constraints are needed.

### 4.3.2 Results

Table 4 reports detailed results of the two settings under consideration. In panel A presenting fund classification over three horizons, empirical results confirm that fund performance varies strongly from one horizon to another. Indeed, coefficients of rank corre-

# Table 4: Global performance

| | Panel A: Perf. over 3 horizons | | | | | Panel B: Perf. over 5 years with 3 measures | | | | |
| | M-Sh over | | | Global Perf. | | M-St | Ω | M-Sh | Global Perf. | |
| Funds | 5 years | 3 years | 1 year | Rank | Score | | | | Rank | Score |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 7 | 21 | 27 | 16 | 0.2028 | 7 | 7 | 7 | 7 | 0.657 |
| 2 | 18 | 17 | 11 | 13 | 0.2854 | 22 | 18 | 18 | 18 | 0.430 |
| 3 | 34 | 26 | 35 | 33 | -0.0003 | 33 | 34 | 34 | 34 | 0.334 |
| 4 | 15 | 28 | 28 | 29 | 0.0153 | 24 | 13 | 15 | 13 | 0.466 |
| 5 | 20 | 25 | 36 | 31 | -0.0002 | 19 | 21 | 20 | 21 | 0.416 |
| 6 | 28 | 8 | 3 | 6 | 0.6827 | 11 | 29 | 28 | 29 | 0.374 |
| 7 | 29 | 23 | 18 | 22 | 0.0806 | 29 | 30 | 29 | 30 | 0.365 |
| 8 | 26 | 14 | 23 | 27 | 0.0419 | 28 | 28 | 26 | 28 | 0.380 |
| 9 | 27 | 11 | 25 | 24 | 0.0734 | 30 | 26 | 27 | 26 | 0.389 |
| 10 | 23 | 16 | 20 | 25 | 0.0568 | 17 | 23 | 23 | 23 | 0.409 |
| 11 | 4 | 22 | 17 | 12 | 0.3600 | 5 | 5 | 4 | 5 | 0.745 |
| 12 | 25 | 10 | 26 | 21 | 0.0951 | 13 | 25 | 25 | 25 | 0.391 |
| 13 | 38 | 37 | 37 | 37 | -0.0005 | 38 | 38 | 38 | 38 | 0.260 |
| 14 | 31 | 6 | 1 | 1 | 1 | 35 | 36 | 31 | 36 | 0.310 |
| 15 | 17 | 24 | 31 | 30 | 0.0034 | 27 | 16 | 17 | 16 | 0.450 |
| 16 | 16 | 29 | 19 | 23 | 0.0749 | 26 | 15 | 16 | 15 | 0.459 |
| 17 | 12 | 27 | 6 | 9 | 0.5109 | 16 | 12 | 12 | 12 | 0.494 |
| 18 | 21 | 35 | 34 | 36 | -0.0003 | 18 | 19 | 21 | 19 | 0.419 |
| 19 | 14 | 32 | 30 | 28 | 0.0162 | 25 | 14 | 14 | 14 | 0.461 |
| 20 | 19 | 13 | 2 | 4 | 0.9856 | 21 | 20 | 19 | 20 | 0.416 |
| 21 | 6 | 7 | 8 | 7 | 0.5718 | 6 | 6 | 6 | 6 | 0.661 |
| 22 | 35 | 33 | 32 | 35 | -0.0003 | 36 | 33 | 35 | 33 | 0.341 |
| 23 | 11 | 12 | 22 | 20 | 0.1046 | 14 | 11 | 11 | 11 | 0.497 |
| 24 | 37 | 18 | 7 | 11 | 0.4276 | 37 | 37 | 37 | 37 | 0.278 |
| 25 | 22 | 31 | 29 | 32 | -0.0002 | 20 | 24 | 22 | 24 | 0.408 |
| 26 | 5 | 1 | 4 | 1 | 1 | 4 | 4 | 5 | 4 | 0.877 |
| 27 | 2 | 5 | 13 | 8 | 0.5565 | 1 | 2 | 2 | 1 | 1 |
| 28 | 1 | 3 | 9 | 1 | 1 | 3 | 3 | 1 | 1 | 1 |
| 29 | 8 | 15 | 16 | 15 | 0.2217 | 9 | 9 | 8 | 9 | 0.530 |
| 30 | 36 | 20 | 14 | 17 | 0.1691 | 34 | 35 | 36 | 35 | 0.318 |
| 31 | 13 | 4 | 10 | 10 | 0.4409 | 23 | 17 | 13 | 17 | 0.438 |
| 32 | 32 | 36 | 21 | 26 | 0.0504 | 10 | 31 | 32 | 31 | 0.358 |
| 33 | 24 | 34 | 15 | 18 | 0.1415 | 15 | 22 | 24 | 22 | 0.409 |
| 34 | 30 | 38 | 38 | 38 | -0.0006 | 31 | 27 | 30 | 27 | 0.382 |
| 35 | 3 | 2 | 5 | 5 | 0.8112 | 2 | 1 | 3 | 1 | 1 |
| 36 | 9 | 19 | 24 | 19 | 0.1411 | 8 | 8 | 9 | 8 | 0.568 |
| 37 | 10 | 9 | 12 | 14 | 0.2675 | 12 | 10 | 10 | 10 | 0.500 |
| 38 | 33 | 30 | 33 | 34 | -0.0003 | 32 | 32 | 33 | 32 | 0.355 |
| Correlation | | | | | | | | | | |
| M-Sh 3 years | | | 0.71 | | | | | | | |
| M-Sh 5 years | 0.49 | 0.34 | | | | 0.82 | 0.99 | | | |
| Ω | | | | | | 0.82 | | | | |

Note: M-Sh = modified Sharpe ratio, M-St = modified Stutzer index, Ω = Omega index. Funds with negative scores are those whose elementary performances are all negative.

lation between classifications are rather weak: 0.49 (5 years versus 3 years), 0.34 (5 years versus 1 year) and 0.71 (1 year versus 3 years). The most striking example is fund 14 classified at the $31^{th}$ position over five-year horizon but at $6^{th}$ and $1^{th}$ ranks over respectively three-year and one-year horizons. According to DEA, this fund is classified as dominant with an aggregate score of 1. In detail, it is found to have the performance over one-year horizon not only classified at the $1^{st}$ rank but also sufficiently high to compensate for the slightly negative performance over five-year horizon so as to arrive finally at the head of the sample. Unlike fund 14, two other dominant funds according to the DEA global performance score — funds 26 and 28 — have very stable performance profiles over time.

With regard to fund rankings according to the three selected performance indicators (Panel B), contrary to the preceding case, classifications are overall very coherent between them. This coherence is confirmed by high coefficients of rank correlation: 0.99 (M-Sharpe with Omega), 0.82 (M-Sharpe with M-Stutzer) and 0.82 (Omega with M-Stutzer). Such finding certainly does not provide an ideal illustration of the problem which this study aims to illustrate here. Nevertheless, in detail, rankings given by the M-Stutzer index and those provided by the M-Sharpe ratio and the Omega index are quite divergent on several occasions. It is particularly true for fund 6 which is ranked at the $11^{th}$ place by the M-Stutzer index but ranked only at the $29^{th}$ and the $28^{th}$ places in the classification of respectively the Omega index and the M-Sharpe ratio. The other examples are funds 4, 10, 12, 15, 16, 19 and 32. In such cases, applying DEA to determine definitive ranks of these funds presents an undeniable interest.

Finally, three funds globally qualified as dominant are funds 27, 28 and 35. All of them are the most highly ranked of the sample, no matter what measure is used. In these cases, the dominance of these funds is obvious, DEA can only confirm it. Only when performance measures disagree on fund rankings that DEA proves its perspicacity by providing for each fund an aggregate indicator of performance allowing a global and definitive classification. Such is the case of funds 14, 18, 22, 25, 31, 33 and 34. Consider for example fund 25. It is ranked $20^{th}$ by the M-Stutzer index, $24^{th}$ by the Omega index and $22^{th}$ by the M-Sharpe ratio. However, according to DEA, it is globally placed only $24^{th}$ when all three performance indicators are considered. Similarly, fund 34 is at the $31^{th}$ position of the list according to M-Stutzer, $27^{th}$ according to Omega, $30^{th}$ according to M-Sharpe but globally ranked $27^{th}$ by DEA.

# Conclusion

Previous studies documented that DEA could be a good tool to solve decision-making problems with multiple criteria, including investment fund performance evaluating task. This paper shows that DEA is particularly adapted to assess hedge fund performance for the following reasons. First, it can incorporate multiple risk-return attributes of non normal returns in an unique aggregate score so as to rank funds. Hence, DEA can be used to evaluate local and global performances of hedge funds. The local performance is obtained when evaluation criteria include risks, eventually sales loads (DEA's inputs) and returns (DEA's outputs). The global performance is defined as the aggregate score of several elementary performances which could be performances over several temporal horizons, or performances over one temporal horizon but measured by different indicators. Second, unlike other performance measures, DEA offers investors the possibility to exert direct control on the importance level paid to each evaluation criteria. Thus, each investor can tailor his own performance measure to select funds corresponding the most to his own preferences. This flexibility is very important as in reality, each investor usually has his own preferences and constraints. Third, by putting emphasis on the best observed funds, DEA makes no assumption on the functional relation between evaluation criteria.

To this end, this paper focuses on the most important methodological issues concerning the application of the basic CCR model to hedge fund performance appraisal, namely (1) the choice of evaluation criteria as DEA's inputs and outputs, (2) the choice between input-oriented or output-oriented version of the CCR model, (3) dealing with negative inputs and outputs, and (4) transcribing investors' specific preferences into mathematical constraints. These elements are presented in such a way to provide investors with a general framework to apply DEA in assessing fund performance. In order to make these guidelines more intuitive, several numerical illustrations with thorough discussion of results are provided on a sample of 38 hedge funds. The illustrations also highlight the importance of correct specification of evaluation criteria and preference structure for efficient application of DEA. A comparison between DEA classification and rankings provided by traditional Sharpe and modified Sharpe ratios indicates that they are sometimes radically inconsistent. Further examination of funds' return distributions suggests that these latter two measures might not price properly good surprises (extremely high positive returns). In such case, DEA proves to be a good supplement to improve the precision of selection tasks. Although this paper only addresses the application of DEA in the hedge fund context, its guidelines are also applicable to other types of investment funds like mutual funds, pension funds or ethical funds, etc.

Like any other tools, DEA also has its caveats. One of the main weaknesses arises from the fact that DEA basic models do not provide complete rankings of dominant funds. Nevertheless, this weakness can be mitigated by either adding more restrictive preferences (additional mathematical constraints) so that the short list becomes more and more selective, or applying other qualitative and quantitative criteria on dominant funds so as to rank them. Besides, the dominance or efficiency of funds is only relative to the other funds in the sample and thus can be changed once the sample is modified. However, relative evaluation is a well-established concept in economic literature (Holmstrom 1982). In addition, the relative property of fund evaluation is still quite valuable because in the investment industry, funds are often rated relatively to others in the same category. A broad literature documented that the investment fund market is a tournament and the managers compete against each other in the same category to attract investors (Brown et al. 1996, Agarwal et al. 2003, Kristiansen 2005).

## References

Agarwal, V., Daniel, N. D. & Naik, N. Y. (2003), Flows, performance, and managerial incentives in the hedge fund industry, Working paper.

Andersen, P. & Petersen, N. (1993), 'A procedure for ranking efficient units in data envelopment analysis', *Management Science* **39**(10), 1261–1294.

Arditti, F. (1967), 'Risk and the required return on equity', *Journal of Finance*.

Basso, A. & Funari, S. (2001), 'A data envelopment analysis approach to measure the mutual fund performance', *European Journal of Operational Research* **135**(3), 477–492.

Basso, A. & Funari, S. (2003), 'Measuring the performance of ethical mutual funds: A DEA approach', *Journal of the Operational Research Society* **54**(5), 521–531.

Broihanne, M.-H., Merli, M. & Roger, P. (2004), *Finance Comportementale*, Gestion, Economica.

Brown, K. C., Harlow, W. V. & Starks, L. T. (1996), 'Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry', *Journal of Finance* **51**(1), 85–110.

Charnes, A. & Cooper, W. (1962), 'Programming with linear fractional functionals', *Naval Research Logistics Quarterly* **9**(3, 4), 181–185.

Charnes, A. & Cooper, W. (1973), 'An explicit general solution in linear fractional programming', *Naval Research Logistics Quarterly*.

Charnes, A., Cooper, W. & Rhodes, E. (1978), 'Measuring the efficiency of decision-making units', *European Journal of Operational Research* **2**, 429–444.

Choi, Y. K. & Murthi, B. P. S. (2001), 'Relative performance evaluation of mutual funds: A non-parametric approach', *Journal of Business Finance & Accounting* **28**(7/8), 853.

Cooper, W. W., Seiford, L. M. & Tone, K. (2000), *Data Envelopment Analysis - A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, Kluwer Academic Publishers.

Farrell, M. (1957), 'The measurement of productive efficiency', *Journal of the Royal Statistical Society* **120**(3), 253–281.

Favre, L. & Galeano, J.-A. (2002), 'Mean-modified value-at-risk optimization with hedge funds', *Journal of Alternative Investments* **5**(2), 21–25.

Francis, J. C. & Archer, S. H. (1979), *Portfolio Analysis*, Foundation of Finance, Prentice - Hall.

Fried, H. O., Lovell, C. K. & Schmidt, S. S. (1993), *The measurement of productive efficiency: Techniques and Applications*, Oxford University Press.

Gregoriou, G. N. (2003), 'Performance appraisal of funds of hedge funds using data envelopment analysis', *Journal of Wealth Management* **5**, 88–95.

Gregoriou, G. N. & Gueyie, J.-P. (2003), 'Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio.', *Journal of Wealth Management* **6**(3), 77–83.

Gregoriou, G. N., Sedzro, K. & Zhu, J. (2005), 'Hedge fund performance appraisal using data envelopment analysis', *European Journal of Operational Research* **164**(2), 555.

Holmstrom, B. (1982), 'Moral hazard in teams', *Bell Journal of Economics* **13**(2), 324–340.

Jean, W. H. (1971), 'The extension of portfolio analysis to three or more parameters', *Journal of Financial & Quantitative Analysis*.

Kane, A. (1982), 'Skewness preference and portfolio choice', *Journal of Financial & Quantitative Analysis* **17**(1), 15–25.

Kaplan, P. & Knowles, J. (2001), The Stutzer performance index: Summary of rationale, mathematics and behaviour, Technical report, Morningstar.

Keating, C. & Shadwick, W. (2002), 'A universal performance measure', *Journal of Performance Measurement* **8**(3), 59–84.

Kooli, M., Morin, F. & Sedzro, K. (2005), Evaluation des mesures de performance des hedge funds, Communication at International Conference of the French Association of Finance, Paris, June 2005.

Kraus, A. & Litzenberger, R. (1976), 'Skewness preference and valuation of risk assets', *Journal of Finance*.

Kristiansen, E. (2005), Agency, performance, and selection: the market for fund managers, Working paper.

McMullen, P. R. & Strong, R. A. (1998), 'Selection of mutual funds using data envelopment analysis', *Journal of Business & Economic Studies* **4**(1), 1.

Morey, M. R. & Morey, R. C. (1999), 'Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking', *Omega, The international journal of management science* **27**(2), 241–258.

Murthi, B. P. S., Choi, Y. K. & Desai, P. (1997), 'Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach', *European Journal of Operational Research* **98**(2), 408.

Powers, J. & McMullen, P. (2000), 'Using data envelopment analysis to select efficient large market cap securities', *Journal of Business and Management* **7**(2), 31–42.

Scott, R. C. & Horvath, P. A. (1980), 'On the direction of preference for moments of higher order than the variance', *Journal of Finance* **35**(4), 915–919.

Seiford, L. & Zhu, J. (1999), 'Infeasibility of super-efficiency data envelopment analysis', *INFOR* **37**(2), 174–187.

Seiford, L. M., Cooper, W. W. & Zhu, J. (2004), *Handbook on Data Envelopment Analysis*, Kluwer Academic Publisher.

Sengupta, J. K. (2003), 'Efficiency tests for mutual fund portfolios', *Applied Financial Economics* **13**, 869–876.

Sexton, T., Silkman, R. & Hogan, A. (1986), Data envelopment analysis: Critique and extensions, *in* R. Silkman, ed., 'Measuring Efficiency: An Assessment of Data Envelopment Analysis', Jossey-Bass, San Francisco, CA, pp. 73–105.

Sharpe, W. (1966), 'Mutual fund performance', *Journal of Business* pp. 119–138.

Stutzer, M. (2000), 'Portfolio performance index', *Financial Analysts Journal* pp. 52–61.

Tarim, S. A. & Karan, M. B. (2001), 'Investment fund performance measurement using weight-restricted data envelopment analysis', *Russian and East European Finance and Trade* **37**(5), 64–85.

Thompson, R., Singleton, F., Thrall, R. & Smith, B. (1986), 'Comparative site evaluations for locating a high-energy physics lab in Texas', *Interfaces* **16**(6), 35–49.

Tone, K. (1999), 'A consensus making method for group decisions', Proposal at the Committee Meeting, National Land Agency.

Zhu, J. (1996), 'Robustness of the efficient DMUs in data envelopment analysis', *European Journal of Operational Research* **90**, 451–460.