serie ec

WP-EC 2011-08

On the informativeness of persistence for evaluating mutual funds performance using partial frontiers

Juan Carlos Matallín, Amparo Soler and Emili Tortosa-Ausina

Ivie

Working papers

Working papers

Working papers

Edita / Published by: Instituto Valenciano de Investigaciones Económicas, S.A.

Depósito Legal / Legal Deposit no.: V-2806-2011

Impreso en España (agosto 2011) / Printed in Spain (August 2011)

WP-EC 2011-08

On the informativeness of persistence for evaluating mutual funds performance using partial frontiers*

Juan Carlos Matallín, Amparo Soler and Emili Tortosa-Ausina**

Abstract

The last few years have witnessed a rapid evolution in the literature evaluating mutual fund performance using frontier techniques. The instruments applied, mostly DEA (Data Envelopment Analysis) and, to a lesser extent, FDH (Free Disposal Hull), are able to encompass several dimensions of performance, but they also have some disadvantages that might be preventing a wider acceptance. The recently developed order-m and order-α partial frontiers overcome some of the disadvantages (they are robust with respect to extreme values and noise, and do not suffer from the well-known curse of dimensionality) while keeping the main virtues of DEA and FDH (they are fully-nonparametric). In this article we apply not only the non-convex counterpart of DEA, namely, FDH but also order-m and order-α partial frontiers to a sample of Spanish mutual funds. The results obtained for both order-m and order-α are useful, since a full ranking of mutual funds' performance is obtained. We combine these methods with the literature on mutual fund performance persistence. By combining the two literatures we derive an algorithm for guiding the choice of m and α parameters intrinsic to order-m and order-α (respectively) based on mutual fund performance persistence.

Keywords: efficiency, mutual funds, partial frontiers, persistence.

Resumen

Los últimos años han sido testigos de una rápida evolución de la literatura que evalúa el rendimiento de fondos de inversión utilizando la metodología del enfoque frontera. Los instrumentos aplicados, principalmente DEA (Data Envelopment Analysis) y, en menor medida, FDH (Free Disposable Hull), son capaces de abarcar varios aspectos del rendimiento, pero también poseen algunas desventajas que podrían impedir una mayor aceptación. El recientemente desarrollado enfoque de las fronteras parciales de orden-m y de orden-alfa supera algunos de los inconvenientes (estos procedimientos son robustos con respecto a los valores extremos y perturbaciones aleatorias o ruido, y no sufren la conocida "maldición de la dimensionalidad" o curse of dimensionality), manteniendo las principales virtudes de DEA y FDH (ambas técnicas son absolutamente no paramétricas). En este artículo se aplica no sólo la versión no convexa de DEA, es decir, FDH, sino también para fronteras de orden-m y de orden-alfa cuya utilidad es notable, ya que se obtiene una clasificación completa del rendimiento de los fondos de inversión. En este trabajo se combinan estos métodos con la literatura existente relativa a la persistencia en el rendimiento de los fondos de inversión. Mediante la combinación de ambas literaturas deducimos un algoritmo capaz de guiar (o que sirva de referencia) en la elección de los parámetros intrínsecos m y alfa correspondientes a orden-m y a orden-alfa (respectivamente) en base a la persistencia en el rendimiento de los fondos de inversión.

Palabras clave: Eficiencia, fondos de inversión, enfoque de fronteras parciales, persistencia.

---

*The authors acknowledge the financial support provided by the Ivie (Instituto Valenciano de Investigaciones Económicas). This study is part of research projects SEJ2007-67204/ECON, ECO2008-03813/ECON and ECO2008-05908-C02-01/ECON supported by the Spanish Ministry of Science and Technology, P1-1B2009-54 and P1.1B2008-46 by Fundació Caixa Castelló-Bancaixa and Universitat Jaume I, and PROMETEO/2009/066 by the Generalitat Valenciana. We are also grateful for comments from participants at the International Risk Management Conference (Florence, June 2010), the EURO XXIV (Lisbon, July 2010) and the International Conference on Operations Research (Munich, September 2010).

**J.C. Matallín and A. Soler: Universitat Jaume I. E. Tortosa-Ausina: Universitat Jaume I and Ivie. Corresponding author: E. Tortosa-Ausina, tortosa@uji.es

1. Introduction

Investors are increasingly interested in sound performance evaluation of available investment funds and, in this regard, they rely on risk-adjusted measures to make their choices. The development of mutual fund industries has given rise to a large body of literature. In this specific field, one issue of particular interest to investors, managers, and academics, and which has been extensively analyzed is, precisely, the performance of funds. From a methodological point of view, the existing literature dates back to Treynor (1965), Jensen (1968) and Sharpe (1966). Since these pioneering contributions, the literature has evolved to propose newer approaches to performance measurement. Some of them have been surveyed by Ippolito (1993), Grinblatt and Titman (1995), Cesari and Panetta (2002) or, in the particular field of hedge funds, Eling and Schuhmacher (2007).

In contrast to what we might call *traditional* approaches for mutual fund evaluation, since the late 1990s interest has been growing in applying the so-called frontier techniques, both parametric and nonparametric (see Murillo-Zamorano, 2004, for a survey) to evaluate the performance of mutual funds. The number of proposals, both from theoretical and empirical points of view, is already substantial, including Murthi *et al.* (1997), McMullen and Strong (1998), Morey and Morey (1999), Wilkens and Zhu (2001), Basso and Funari (2001), or Choi and Murthi (2001), among others. Indeed, due to the now remarkable number of proposals, some initiatives have been taken to review early contributions, such as those by Joro and Na (2002) and, more recently, Eling (2006), Glawischnig and Sommersguter-Reichmann (2010), or the monograph by Gregoriou and Zhu (2005) in the specific fields of hedge fund and commodity trading advisors (CTAs) performance evaluation.

These studies examine the advantages and disadvantages of applying nonparametric frontier techniques—mostly Data Envelopment Analysis (DEA)—to evaluate the performance of mutual funds. The main advantage of these approaches is one of the features which has led to the popularity of DEA, i.e., the ability to deal simultaneously with several inputs and outputs, and to combine them in a single performance indicator—namely, the so-called *efficiency score*. This ability fits conveniently into the context of mutual fund performance evaluation, where one may be interested in extending the approach to include other dimensions apart from mean and variance, thus allowing the inclusion not only of skewness, but also of other relevant dimensions. DEA also has the ability to weight easily, by selecting the optimal weight for each dimension.

However, one a disadvantage may have prevented some academics and practitioners from using

DEA and related techniques such as Free Disposable Hull (FDH), namely, the so-called "curse of dimensionality", which is related to problems associated with a low number of DMUs (Decision Making Units) relative to the number of input-output variables. This phenomenon not only affects both FDH and DEA estimators, but is also shared by other nonparametric approaches in statistics and econometrics. Although the issue was reported a while ago, few empirical applications have actually acknowledged its severity. However, some authors have taken the problem very seriously, claiming that "a number of applied papers using relatively small numbers of observations with many dimensions have appeared in the literature, but we hope that no more will appear" (Simar and Wilson, 2008, p.441). The curse of dimensionality severely affects those cases in which the number of inputs and outputs might vary, as well as the number of units under analysis.

From a theoretical point of view, the literature has evolved to provide solutions to the curse of dimensionality. The order-m (Cazals *et al.*, 2002) and order-α (Daouia and Simar, 2007) estimators are robust indicators not only to the curse of dimensionality itself, but also to the presence of outliers and noise in the data, to which both DEA and FDH are particularly sensitive. Neither order-m nor order-α require convexity assumptions and, in addition, they both have several desirable properties that are useful for drawing inferences about efficiency. As indicated by Wheelock and Wilson (2009), while keeping the fundamental advantages of DEA and FDH (i.e., being fully-nonparametric), they overcome some of their shortcomings, since they are $\sqrt{n}$ consistent, do not suffer from the curse of dimensionality and are robust to outliers and noise.

However, empirical applications are still scarce. In the particular context of mutual fund performance evaluation, only Daraio and Simar (2005, 2006, 2007b) have considered these robust methods. Although their theoretical contributions are highly valuable, they confine their analysis almost entirely to order-m estimators. In our paper, we update the contributions by Daraio and Simar (2005, 2006, 2007b) in several directions. First, we stretch the data to more recent dates, i.e., we focus on the period 1998–2007 whereas Daraio and Simar (2005, 2006, 2007b) consider sample periods for the early 2000s only; in addition, we have a much tighter focus on the application than on the specific details of the techniques. Second, our analysis is not confined to order-m techniques only. Taking into account the recent developments in the theoretical literature on efficiency and productivity analysis, we perform a comparison of *classical* approaches (FDH) with the new contributions, considering both Cazals *et al.*'s order-m estimators and Daouia and Simar's (2007) order-α estimators. This robustness analysis has relevant implications, since the analyst (especially from a practitioner's point of view) might be puzzled if different methodologies yield different results.

Related to this, one of the main contributions of the paper is to analyze the robustness of results when applying partial frontiers. Should the applied methodologies be robust, it would be possible to forecast mutual fund efficiencies where they persist over time. Given that the partial frontiers methodologies provide us with funds' rankings, the practitioner, or an individual investor, could use this information to buy the best (winner) and sell the worst (loser) funds. It would therefore be possible to evaluate which method is best able to discriminate between best and worst funds. This approach stands along with the large body of literature devoted to measuring whether certainss fund managers consistently achieve higher (or lower) returns than their competitors. As one key component of the fund selection process, most individual investors and their advisors spend a significant amount of time studying historical performance of mutual funds, since it contains useful information about future performance. As indicated by Droms (2006), "winners in one year tend to remain winners in the following year and losers have an even stronger tendency to remain losers" (Droms, 2006, p.60). This particular topic has gained importance in the mutual fund performance evaluation literature, and several significant studies have been published since the early 1990s acknowledging this reality (see, for instance Grinblatt and Titman, 1992; Brown and Goetzmann, 1995; Carhart, 1997; Hendricks *et al.*, 1993; Elton *et al.*, 1996; Hendricks *et al.*, 1993, among others). More recently, Pätäri (2009) has provided an extensive literature review of mutual fund performance persistence, and Cremers and Petajisto (2009) and Loon (2011) have proposed new methods reporting evidence of persistence and also on how investors respond to previous performance rankings.

As Pätäri (2009) points out, analysis of persistence is often sensitive to methodological choices, especially in the case of equity funds. These choices are either parametric or nonparametric methods that focus on the analysis of persistence as a static association between the performance of different time periods. To avoid this sensitivity to the method and provide enhanced robustness, our paper focuses on the economic relevance of mutual fund persistence rather than adopting a static approach. Related to this, Carhart (1997) proposed a framework in which the most relevant result is the economic value added of persistence. Following this approach, we will construct equally weighted portfolios that follow a buy-and-hold strategy based on the past efficiencies of mutual funds obtained using partial frontiers. This strategy will help guide investors' choices, based on the assumption that a good methodology to measure mutual fund efficiency is one that provides investment recommendations which, when followed, yield good results; in other words, a methodology that captures the *persistence* of managers' skills over time.

The paper is structured as follows. In section 2 we discuss the advantages and disadvantages of the most popular nonparametric techniques for efficiency measurement, namely, DEA and FDH, along with the new partial frontiers. Section 3 presents the underpinnings of the persistence analysis. Section 4 and 5 report the data and results, respectively. Section 6 concludes.

2. Mutual fund evaluation using frontier techniques

As noted above, the literature on the evaluation of mutual fund performance using frontier techniques has grown considerably. Apart from the nonparametric approaches referred to in the previous section, contributions have also come from the parametric field, where the most popular method is Stochastic Frontier Analysis, SFA (Lovell and Kumbhakar, 2000). These approaches must to specify a functional form for the frontier, and choose a distribution for the inefficiency. None of these requirements have to be met in the case of nonparametric frontier methods. Studies applying parametric frontier analysis methods to mutual funds include Briec *et al.* (2004), or Annaert *et al.* (2003), who considered stochastic Bayesian techniques. Although these approaches have several advantages, their drawbacks (not only having to specify a functional form for the frontier and distributions for the inefficiency, but also the assumption of independence for the inefficiency term) have led many authors to lean towards nonparametric methods.

Within the nonparametric field, we can distinguish between a theoretical view (Sengupta, 1991; Sengupta and Park, 1993; Briec *et al.*, 2001) or a more applied perspective (apart from the references provided in the introduction, see also Sengupta, 2000). From a theoretical point of view, Sengupta (1991) and Sengupta and Park (1993) provide links between the Capital Asset Pricing Model (CAPM) and nonparametric estimation of frontiers, whereas Briec *et al.* (2001) analyze the relation between the hypothesis of the basic Markowitz (1952) model and efficiency analysis theory, by developing a dual framework for assessing the degree to which investors' preferences are satisfied. From a more applied perspective, the first specific application of DEA for evaluating the performance of mutual funds was Murthi *et al.* (1997), whose main motivation was to overcome the shortcomings of the *classical* two dimensional (mean-variance) performance measures.

A careful review of the literature assessing performance of traditional and alternative investment funds using DEA is provided by Glawischnig and Sommersguter-Reichmann (2010). Their survey implicitly suggests that the amount of studies applying nonparametric frontier methods such as DEA greatly outnumbers others using parametric methods. They conclude that DEA applications in the investment fund industry can be classified into two categories, namely, *traditional* and *alternative* fund performance evaluation studies. Their survey also implicitly recognizes that the studies applying FDH to mutual fund evaluation are virtually non-existent. The paper by Daraio and Simar (2006) is also mentioned in their survey, but only to briefly indicate that their approach was "computationally demanding" (Glawischnig and Sommersguter-Reichmann, 2010, p.297).[1]

2.1. Linear programming techniques: DEA and FDH for mutual fund evaluation

Several measures of efficiency can be used comprehensively way in order to make a rigorous comparative efficiency analysis. The DEA efficiency score is a performance indicator obtained by comparing each mutual fund with the best performers of its objective group. The same underpinnings of DEA are shared by the FDH estimator. In his early proposals, Tulkens (1993) stressed the relevance of the main difference between DEA and FDH, namely, DEA rests on the hypothesis of convexity of the attainable set, whereas FDH does not. If the convexity hypothesis is questionable, DEA might be a wrong measure (i.e. statistically inconsistent). We may even consider that DEA is closer to parametric methods than FDH, since imposing convexity in DEA actually *is* an assumption. However, both DEA and FDH methods share the same advantage compared with parametric methods, i.e., no hypotheses are required (they are free from the "parametric straitjacket") and they are relatively straightforward to compute.

In contrast to other performance measures, both DEA and FDH benefit from the ability to incorporate many factors that are associated with the fund performance in a very flexibly way. In particular, both approaches allow definition of mutual fund performance indexes that can take into account different risk measures and the costs of investment (e.g., fees). Following Banker *et al.* (1984) only one minimum assumption is required for DEA: the convexity of the efficient frontier (convexity implies that any convex combination of inputs and outputs is feasible in the production function). The efficiency of a fund can be determined by the relative distance between the observed output and the 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.

Banker and Maindiratta (1986) compared the advantages of using DEA over parametric methods. In the context of mutual fund performance evaluation, DEA has the advantage of being a nonparametric analysis and, as such, does not require any theoretical model as a benchmark, such as the CAPM or the Arbitrage Pricing Theory (APT). 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

[1] Thanks to the `FEAR` package for `R` by Paul W. Wilson this claim is no longer valid (Wilson, 2008). See also `URL:http://http://www.clemson.edu/economics/faculty/wilson/`.

loads, as well as returns. Basso and Funari (2001) measured the efficiency of a sample of mutual funds between 1997 and 1999. Their contribution was to develop a generalized DEA-based performance measure that can integrate both classic performance measures (such as Sharpe, Treynor, and Jensen) and the approach of Murthi *et al.* (1997).

Another positive characteristic is that DEA measures its Sharpe measure relative to that of the best-performing fund in the same category; in other words, DEA measures the performance of a fund in reference to the best set of funds within the declared objective category. In Banker and Morey (1986) and Kamakura (1988) controllable categorical variables in the form of outputs are only treated as hierarchically ordered, e.g. outputs are classified in categories or similar orderings, with respect to attributes. Basso and Funari (2003) proposed a DEA categorical variable model in order to find an appropriate model to obtain an indicator of ethical fund performance. Fund performance is a combination of multiple fund attributes such as mean returns (outputs), risk (total or systematic) and expenses, and sometimes even fund size, turnover speed and minimum initial investment (inputs). Employing essentially basic DEA models like CCR (Charnes *et al.*, 1978) or BCC (Banker *et al.*, 1984), they sought to compare the efficiency of funds within a category or between several different categories of funds.

This nonparametric approach allows one to estimate an efficient frontier combining mean-variance and cost efficiency and to further estimate returns to scale for each mutual fund, implying that with DEA the effect of returns to scale on performance is controlled for (Choi and Murthi, 2001). The next advantage is that DEA measures efficiency with respect to the efficient frontier, which measures the best performance that can be achieved in practical terms. Another important point is the consideration that DEA provides an efficient index (the so-called *efficiency scores*) for each mutual fund, which enables calculation of the optimal weights for each attribute into different time periods. And the last feature described is that DEA not only measures inefficiency, but also the magnitude of the inefficiency in the different dimensions. This is considered the greatest advantage of using the DEA method over other approaches for measuring fund performance: namely that DEA reveals the reason why a fund is inefficient and shows how to restore the fund to its optimum level of efficiency. Choi and Murthi (2001) and Kuosmanen *et al.* (2006) argue that 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. Therefore DEA not only measures efficiency, but can also provide guidance as to how to improve the efficiency of inefficient funds.

2.2. Measuring performance via DEA and FDH

In the first stage of the estimation process of this paper we evaluate the performance of the mutual funds in our sample considering the common non-convex FDH frontier. Although the preceding paragraphs have focused more closely on DEA, we have chosen the FDH frontier because of its higher flexibility and its asymptotic properties (Park *et al.*, 2000). Previously, the set of attainable combinations of inputs ($\mathbf{x}$) and outputs ($\mathbf{y}$), which defines the frontier of the set of possibilities, must be defined. To define the efficiency of a given fund we will then measure the distance between the observed value of the fund variables and the frontier. The $\mathbf{\Psi}$ set of possibilities is the set of attainable points $(\mathbf{x}, \mathbf{y})$, defined as:

$$\mathbf{\Psi} = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^{p+q} | (\mathbf{x}, \mathbf{y}) \text{ are attainable}\} \tag{1}$$

where $\mathbf{x} \in \mathbb{R}_+^p$ is the vector of inputs and $\mathbf{y} \in \mathbb{R}_+^q$ is the vector of outputs. For all possible output values we may define the section of possible values of $\mathbf{x}$ as

$$X(\mathbf{y}) = \{\mathbf{x} \in \mathbb{R}_+^p | (\mathbf{x}, \mathbf{y}) \in \mathbf{\Psi}\} \tag{2}$$

and its efficient boundary would be the subset of $X(\mathbf{y})$ defined by

$$\partial X(\mathbf{y}) = \{\mathbf{x} | \mathbf{x} \in X(\mathbf{y}), \theta\mathbf{x} \notin X(\mathbf{y}), \forall\theta \in (0, 1)\}. \tag{3}$$

In this particular setting the Farrell (1957) measure of input-oriented efficiency of a given mutual fund $(\mathbf{x}, \mathbf{y})$ is defined as

$$\theta(\mathbf{x}, \mathbf{y}) = \inf\{\theta : (\theta\mathbf{x}, \mathbf{y}) \in \mathbf{\Psi}\} = \min\{\theta : \theta\mathbf{x} \in X(\mathbf{y})\}, \tag{4}$$

where $\theta(\mathbf{x}, \mathbf{y}) \leq 1$ is the proportionate reduction of inputs required for a mutual fund with the input-output mix $(\mathbf{x}, \mathbf{y})$ to become efficient, i.e., to achieve the value of 1, since the efficient frontier corresponds to those funds whose $\theta(\mathbf{x}, \mathbf{y}) = 1$. The Farrell (1957) output-oriented efficiency score would be defined analogously.

The main nonparametric estimators to measure efficiency, namely, DEA and FDH, are based on envelopment techniques. As indicated previously, the main difference between DEA and FDH is that the former imposes convexity of the set $\mathbf{\Psi}$, whereas the latter drops this assumption. Both consider that the attainable set is defined by the set of minimum volume containing all the observations. Therefore, one may think of FDH as a "pure" nonparametric estimator—no assumption is imposed.

The FDH estimator of $\boldsymbol{\Psi}$, based on a sample of n observations $(\mathbf{x}_i, \mathbf{y}_i)$ is the free disposal closure of the reference set $\{(\mathbf{x}_i, \mathbf{y}_i) | i = 1, \ldots, n\}$, and it can be defined as:

$$\hat{\boldsymbol{\Psi}}_{\mathrm{FDH}} = \{(\mathbf{x}, \mathbf{y}) | \mathbf{x} \geq \mathbf{x}_i, \mathbf{y} \leq \mathbf{y}_i, i = 1, \ldots, n\}. \tag{5}$$

The FDH methodology is particularly recommended if the intention is to uncover the most blatant cases of *inefficiency*, since this technique is very demanding with regard to inefficiency measurement. For each fund labeled as FDH-inefficient, at least one other fund with superior performance can be found in the sample. However, for the convex DEA model, the results found for some *inefficient* funds might depend entirely on the convexity assumption.

2.3. Partial frontiers for mutual fund evaluation: order-m and order-α

2.3.1. The expected order-m frontier

The sensitivity of the deterministic DEA and FDH to measurement errors, outliers, sampling errors, and missing variables is an ongoing concern. In this regard, it is worth noting that the return data from financial markets are typically much more reliable and accurate than empirical production data usually studied with DEA. Therefore the problem of measurement error could seem *a priori* a less serious concern in the present context. However, the problem of outliers can actually occur in this setting if the return possibilities' set includes assets that, for whatever reason, are infeasible investment alternatives for the fund manager. By careful modeling of the investment alternatives as well as the investment criteria and constraints facing the fund managers, the problem of outliers can be alleviated. For instance, Kuosmanen (2007) constructs the benchmark portfolios directly from stocks and other assets, and his results indicate that heterogeneity of the evaluated funds were not obscuring the efficiency measures—although it could affect their ranking. Neither does sampling error seem to be a major problem: return data for stocks, bonds, and other investment alternatives are available, and modeling the fund manager's entire investment universe is technically feasible. Moreover, the sampling theory of the DEA and FDH estimators is nowadays well understood and those insights can also be directly applied in the present context.

The sensitivity of both DEA and FDH to the presence of outliers is caused by the fact that both methods envelop all observation points quite closely. All the nonparametric envelopment estimators of frontiers are particularly sensitive to extreme observations, or outliers, which may disproportionately influence the evaluation of mutual fund performance. In addition, both DEA and FDH estimators, like other nonparametric measures, are affected by the curse of dimensionality due to their slow convergence rate (Simar and Wilson, 2008, p.441). As indicated in the introduction, the problem is especially severe when the number of inputs and outputs is low with respect to the sample size, or when the number of inputs and outputs is unclear.

Taken together, the aforementioned problems may be serious enough to jeopardize the FDH estimates. To solve these problems, some additional procedures are required in order to make FDH estimates more robust. Several approaches have already been proposed in the literature. For instance, Wilson (1993, 1995) introduced descriptive methods to detect influential observations in nonparametric efficiency calculations. More recently, Cazals *et al.* (2002), Daraio and Simar (2005), Aragon *et al.* (2005) and Daouia and Simar (2007) have developed robust alternatives to the DEA and FDH estimators. Specifically, the nonparametric estimation order-m method developed by Cazals *et al.* (2002) is much more robust to both outliers and the curse of dimensionality. These authors introduced the concept of expected maximal output (or minimum input) frontier. It reflects a more realistic benchmark because it is constructed by comparing the performance of each fund (in terms of its use of inputs, i.e. risk and transaction costs) not with the best performing funds of the group, but considering the *expected* value of the minimum level of inputs of m funds drawn from the distribution of funds with a level of output equal to or higher than that of the analyzed fund. The order-m also allows for statistical inference while keeping its nonparametric nature. We briefly describe this approach below.

Let us consider the conditional distribution function $F_{\boldsymbol{x}|\boldsymbol{y}}(\boldsymbol{x}_0|\boldsymbol{y}_0) = Pr(\boldsymbol{x} \leq \boldsymbol{x}_0|\boldsymbol{y} \geq \boldsymbol{y}_0)$. For a given level of inputs $\boldsymbol{y}_0$ in the interior of the support of $\boldsymbol{y}$, consider the m iid random variables $\{\boldsymbol{\nu}_j\}_{j=1}^m$, drawn from the conditional distribution $F_{\boldsymbol{x}|\boldsymbol{y}}(\cdot|\boldsymbol{y}_0)$.[2]

Formally, the proposed algorithm (algorithm I) to compute the order-m estimator has the following steps:

1. For a given level of $\boldsymbol{y}_0$, draw a random sample of size m with replacement among those $\boldsymbol{y}_i$, such that $\boldsymbol{y}_i \geq \boldsymbol{y}_0$.

2. Compute Program (4) and estimate $\widetilde{\theta}_i$.

3. Repeat steps 1 and 2 B times and obtain B efficiency coefficients $\widetilde{\theta}_i^b (b = 1, 2, \ldots, B)$. The quality of the approximation can be tuned by increasing B, but in most applications

[2]Full technical details can be found in, for instance, Daraio and Simar (2007a).

$B = 200$ seems to be a reasonable choice.

4. Compute the empirical mean of B samples as:

$$\theta_i^m = \frac{1}{B}\sum_{b=1}^{B}\widetilde{\theta}_i^b \tag{6}$$

As m increases, the number of observations considered in the estimation approaches the observed units that meet the condition $\boldsymbol{y}_i \geq \boldsymbol{y}_0$ and the expected order-m estimator in each one of the b iterations ($\widetilde{\theta}_i^b$) tends toward the FDH estimator. Thus, m is an arbitrary positive integer value, but it is always convenient to observe the fluctuations of the $\widetilde{\boldsymbol{\theta}}_i^b$ coefficients depending on the level of m. For acceptable values of m, θ_i^m will normally present values smaller than unity. When $\theta_i^m > 1$, the i unit can be labeled as superefficient (Andersen and Petersen, 1993), as the order-m frontier exhibits a higher total cost. In addition, from an economic perspective, the order-m efficiency score has its own interest, since it does not provide the output-efficient frontier, but rather another reasonable benchmark value of the output for a fund with an $\boldsymbol{x}_0$ level of input: it is the *expected* maximal level of output achievable among a fixed number of m funds drawn from the population of funds with at most the same $\boldsymbol{x}_0$ level of input Simar (2003). Please note that this interpretations would correspond to an input orientation.

As indicated above, three aspects of the FDH methodology deserve special attention, namely, *efficiency by default*, the presence of *outliers* and the *curse of dimensionality*. In the absence of a sufficient number of *similar* mutual funds for a comparison, a particular fund is labeled as efficient *by default*. This efficiency ranking does not result from any effective superiority, but rather is due to the lack of information that would allow pertinent comparisons. In addition to this, the FDH concept of efficiency, by construction, applies both to the fund that presents the lowest level of inputs and to those with the highest values for at least one output indicator. This extreme form of the sparsity bias that characterizes the FDH technique leads ultimately to a lack of discrimination (i.e., an inability to *rank*) among production units and constitutes a shortcoming of the FDH approach.

Regarding outliers, nonparametric frontiers are defined by the extreme values of the dimensional space of inputs and outputs. Therefore, the existence of outliers (atypical observations that differ significantly from the rest of the data) may considerably influence the estimation of efficiency, so it is important to verify that the divergences do not result from evaluation errors. Due to the trimming nature of the order-m frontier, this estimator does not envelop all the observed data points (even for large values of m) and, therefore, it is more robust to outliers and/or extreme values. Finally, as for the curse of dimensionality, order-m estimators are much less affected than either DEA or FDH because of some of their statistical properties—they are $\sqrt{n}$-consistent and asymptotically normal.

2.4. The order-α quantile-type frontiers

Apart from the order-m estimators, there is another family of partial frontiers that has been proposed to overcome both the influence of outliers and the curse of dimensionality, namely, the order-α quantile frontiers (Daouia and Simar, 2007). The idea behind the order-α quantile-type frontier is to go the other way round, i.e., to determine the frontier by first fixing the probability $(1-\alpha)$ of observing points above this order-α frontier. Therefore, the order-α quantile frontiers reverse the causation of order-m and choose the proportion of the data lying above the frontier directly.

Order-α estimators also have better properties than the usual nonparametric frontier estimators (either DEA or FDH). They are consistent estimators of the full frontier, since the "order" (in this case the α order) of the frontier is allowed to grow with sample size. They have also the advantage, shared with order-m, that the asymptotic properties are the same as those of FDH. But perhaps the main advantage, also shared with order-m, is that in finite samples, the new estimators do not envelop all the data, and they are therefore more robust to outliers than FDH or DEA. As indicated by Simar and Wilson (2008), they have the side benefit of detecting outliers (Simar and Wilson, 2008, p.480).

The order-m ideas can easily be adapted to order-α quantile type-frontiers. The underpinnings of order-α were initially developed for the univariate case by Aragon *et al.* (2005) and extended to the multivariate setting by Daouia and Simar (2007), and are similar to those of quantile regression (Koenker, 2001).

Recall that in the context of order-m partial frontiers, a mutual fund operating at $(\boldsymbol{x}, \boldsymbol{y})$ is benchmarked against the expected minimum input among m peers drawn randomly from the population of funds with output levels of at least $\boldsymbol{y}$. In contrast, order-α quantile frontiers benchmark the mutual fund considered at $(\boldsymbol{x}, \boldsymbol{y})$ against the input level not exceeded by $(1-\alpha) \times 100\%$ of funds among the population of funds providing output levels of at least $\boldsymbol{y}$.

Following Simar and Wilson (2008), for $\alpha \in (0,1]$, the α-quantile input efficiency score for the mutual fund operating at $(\boldsymbol{x}, \boldsymbol{y}) \in \boldsymbol{\Psi}$ can be defined as

$$\theta_\alpha(\boldsymbol{x}, \boldsymbol{y}) = \inf\{\theta | F_{\boldsymbol{x}|\boldsymbol{y}}(\theta \boldsymbol{x} | \boldsymbol{y}) > 1-\alpha\} \tag{7}$$

Clearly, $\theta_\alpha(\boldsymbol{x},\boldsymbol{y})$ converges to the usual Farrell-Debreu input efficiency score $\theta(\boldsymbol{x},\boldsymbol{y})$ (i.e., to the FDH estimator) when $\alpha \to 1$. As pointed out by Daraio and Simar (2007a), the order-α efficiency score has an interesting interpretation. In cases where $\theta_\alpha(\boldsymbol{x},\boldsymbol{y}) = 1$, the fund is "efficient" at the level $\alpha \times 100\%$, since it is dominated by mutual funds providing more output than $\boldsymbol{y}$ with probability $1-\alpha$. In those cases where $\theta_\alpha(\boldsymbol{x},\boldsymbol{y}) < 1$ then the unit $(\boldsymbol{x},\boldsymbol{y})$ has to reduce its input to the level $\theta_\alpha(\boldsymbol{x},\boldsymbol{y})\boldsymbol{x}$ to achieve the input efficient frontier of level $\alpha\times 100\%$. Analogously to the case of the partial order-m frontiers, the case where $\theta_\alpha(\boldsymbol{x},\boldsymbol{y}) > 1$ *is* feasible, indicating that a particular fund $(\boldsymbol{x},\boldsymbol{y})$ can increase its input by a factor $\theta_\alpha(\boldsymbol{x},\boldsymbol{y})$ to reach the same frontier, a case in which this fund would be labeled as super-efficient (Andersen and Petersen, 1993) with respect to the order-α frontier level. Finally, we can apply the plug-in principle to obtain an intuitive nonparametric estimator of $\theta_\alpha(\boldsymbol{x},\boldsymbol{y}) = 1$ by replacing $F_{\boldsymbol{x}|\boldsymbol{y}}(\cdot|\cdot)$ with its empirical counterpart to obtain:

$$\widehat{\theta}_{\alpha,n}(\boldsymbol{x},\boldsymbol{y}) = \inf\{\theta|\widehat{F}_{\boldsymbol{x}|\boldsymbol{y},n}(\theta\boldsymbol{x}|\boldsymbol{y}) > 1-\alpha\} \tag{8}$$

Again, it is clear that when $\alpha \to 1$, then $\widehat{\theta}_{\alpha,n}(\boldsymbol{x},\boldsymbol{y})$ converges to the FDH input efficiency score $\widehat{\theta}_{FDH}(\boldsymbol{x},\boldsymbol{y})$.

As indicated by Daouia and Simar (2007), in practice the choice of the "tuning" parameters, both m and α, may be governed by their economic interpretation. Whereas in the case of order-m the benchmark could be against the best of m virtual competitors, in the case of order-α it would be against a level of input with a probability $(1-\alpha)\times 100\%$ of being dominated.

Regarding the choice of partial frontier estimator, Daouia and Simar (2007) conclude that both approaches (order-m and order-α) provide nonparametric estimators of the efficient frontier which are more robust than the usual envelopment estimators (like FDH/DEA estimators). It could be argued that the α-quantile approach is easier to interpret, since the parameter α is just the selected level of the quantile. The choice of the m parameter is more intricate although, in our particular setting, it can be interpreted as the number of potential funds against which the benchmark is set to determine the performance of a particular fund. As indicated by Daouia and Simar (2007), although the choice of m can also be indirectly piloted by the percentage of observed funds staying above the frontier for a given m, the α-quantile approach seems to be more direct.

3. Partial frontiers and persistence analysis of mutual fund performance

In the previous sections we have proposed and described the order-m and order-α estimators to evaluate mutual fund performance. We now test the performance of these methods in guiding the selection of funds, for which we have to evaluate the performance of each method along with the choice of *tuning* parameters, i.e., how choosing a given level of m and α might influence the results. In this task, we must bear in mind that an important use of mutual fund performance measures is to assess the possible value added by fund managers. This helps investors, both individual and institutional, to choose from the wide universe of mutual funds in the market. It is therefore reasonable to assume that a good methodology to measure mutual fund efficiency is one offering investment recommendations that, when followed, provide good results; in other words, a methodology that captures the *persistence* of managers' skills over time.

Following the methodology from the performance persistence literature (see, for instance Carhart, 1997; Bollen and Busse, 2005), we construct equally-weighted portfolios that follow a buy and/or sell strategy based on the past efficiency of mutual funds using results yielded by order-m and order-α methods. These portfolios are rebalanced at the end of each semester, investing (selling) the best (worst) mutual funds in the previous annual periods according to the efficiency ranking. Different quantile levels are computed: deciles, quintiles and half of the mutual funds' efficiency distribution. To achieve enhanced robustness in this analysis, three strategies based on past efficiency are considered: (i) buying past top quantile and selling past bottom quantile (buy + sell); (ii) buying past top quantile (buy); and (iii) selling bottom quantile (sell).

Previous work has considered the latter two strategies to analyze possible asymmetries in managers' persistence ability. Carhart (1997), Lynch and Musto (2003) and Bollen and Busse (2005) show different levels of persistence for the best and worst mutual funds. In this study, 36 portfolios are formed that follow dynamic strategies based on past efficiency (3 strategies $\times$ 2 efficiency methods—order-m and order-α—and $\times$ 6 levels for each one). After computing the daily return of each portfolio we estimate their performance.

We should bear in mind that each portfolio is linked to a buy and/or sell strategy in which funds are bought or sold depending on their past efficiency, which we estimate by selecting different parameters for the partial frontiers—m, in the case of order-m, and α, in the case of order-α. The objective is now to evaluate the performance of these portfolios, which we do via partial frontiers. However, this could entail a endogeneity problem because this methodology involves evaluating the results of portfolios constructed considering the information obtained with the same methodology. In order to avoid this, we evaluate the portfolios with an "independent" method which has been used intensely by the financial literature on portfolio management, namely, the multifactor linear model, which can be represented as:

$$r_{pt} = \alpha_p^J + \sum_j^J \beta_{pj} r_{jt} + \varepsilon_{pt} \tag{9}$$

where r_{pt} is the excess return, over the risk-free return, of portfolio p which follows a strategy based on past efficiency. This return is adjusted to risk factors β_{pj} with respect to r_{jt}, the excess return of the benchmark j. Then, as an extended version of Jensen's (1968) α, α_p^J measures performance. Considering the nature of the mutual funds analyzed, and to avoid benchmark omission bias (Pástor and Stambaugh, 2002; Matallín-Sáez, 2006) the following benchmarks will be considered: (i) the Ibex 35 index, for broad stock investment; (ii) MSCI Spanish market index, for Small Caps, Value and Growth styles; (iii) the AFI Government Debt index; and (iv) MSCI World index.[3]

4. Data

4.1. Data sources

The empirical analysis used a sample of Spanish mutual funds (FIM) from July 1998 to March 2007. The sample is made up of all the domestic equity mutual funds with a net asset value during this period.[4] Following the Spanish Stock Market National Commission (*Comisión Nacional del Mercado de Valores*, CNMV), two types of funds can be distinguished: equity funds (EF) and balanced equity-bonds funds (BF). The daily return was calculated as the variation relative to the net asset value.

Mutual funds data such as the net asset value, size, fees and loads were provided by the CNMV (Comisión Nacional del Mercado de Valores). Table 1 reports some basic statistics for the mutual funds sample.

[3]Further versions of the model are also applied, one without the world index and another only with the stock market (CAPM). For reasons of space, we only present the results for the broadest model.

[4]The Spanish mutual fund industry essentially evolved in the second half of the 1990s. For this reason, if we had selected an earlier starting date for the sample period, the number of funds in the sample would have been drastically reduced. Thus, the use of daily data has allowed us to analyze a large volume of information for all the funds existing on 30th June 1998.

Table 1: Descriptive statistics for inputs and outputs, mutual funds (1999–2007)[a]

Class.: EF	Min.	Max.	Mean	Median
Mean daily return	–0.01%	0.07%	0.02%	0.02%
Std.dev. daily return	0.52%	1.37%	1.13%	1.22%
Skewness daily return	–0.675	0.437	–0.151	–0.114
Kurtosis daily return	2.593	11.248	3.947	3.551
Average annual loads	1.23%	7.15%	4.97%	5.40%
Average annual fees	0.45%	2.79%	1.88%	2.02%
Average size	3.167	417.572	81.586	47.903
Number of funds	74			
Class.: BF	**Min.**	**Max.**	**Mean**	**Median**
Mean daily return	–0.01%	0.05%	0.01%	0.01%
Std.dev. daily return	0.15%	1.24%	0.62%	0.61%
Skewness daily return	–2.755	4.043	–0.111	–0.118
Kurtosis daily return	1.815	37.175	5.442	3.568
Average annual loads	1.15%	15.53%	4.54%	4.52%
Average annual fees	0.00%	2.52%	1.68%	1.75%
Average size	1.994	547.488	59.02	25.758
Number of funds	131			

[a] The table presents some descriptive statistics for the mutual fund sample. The sample period runs from July 1st, 1998 to March 31st, 2007. The size is measured by the assets in millions of euros and management fees and loads costs are shown as percentages of the assets. EF represents equity funds and BF, balanced funds.

4.2. Inputs and outputs selection

As indicated in previous sections, one of the main benefits of using frontier techniques to evaluate the performance of mutual funds is their ability for handling multiple inputs and outputs in the model. As indicated by Basso and Funari (2001), "DEA approach allows defining mutual fund performance indexes that can take into account several inputs and thus consider different risk measures (standard deviation, standard-semi deviation and beta) and redemption cost."

DEA, FDH or the partial frontiers order-m and order-α approaches may include other outputs apart from the traditional mean return measure in this framework, such as expected return or the expected excess return. In computing their portfolio efficiency index Murthi *et al.* (1997) considered the standard deviation of returns, expense ratio, loads and turnover as inputs, and mean gross return as output. Choi and Murthi (2001) applied the same inputs and outputs as Murthi *et al.* (1997) although adopting a different DEA formulation. Wilkens and Zhu (2001) performed their study with standard deviation and percentage of periods with negative returns as inputs, and mean return, minimum return and skewness as outputs. In Joro and Na (2002) there is an extension of the traditional mean-variance framework using DEA, and their methodology includes the risk and cost associated with the transaction as inputs, while return and skewness are included as outputs. Chang (2004) proposed a new non-standard DEA formulation based on minimum convex input requirement set: the standard deviation, β, total assets and loads, while the output was the traditional mean return.

The *right* selection of inputs and outputs is crucial when using frontier techniques. 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). Briec *et al.* (2004) developed a quadratic-constrained (mean-variance) DEA model applying a mean-variance approach with variance as input and mean return as output. And Lozano and Gutiérrez (2007) proposed a quadratic-constrained DEA model consistent with Third-degree Stochastic Dominance (TSD) in order to obtain an optimal portfolio benchmark for any rational risk-averse investor. Briec and Kerstens (2009) present a quadratic program that extends the multi-horizon analysis by Morey and Morey (1999) in several ways. Joro and Na (2006) suggested a cubic-constrained a mean-variance-skewness framework similarly to Briec *et al.* (2007), who consider both skewness and mean return as outputs.

To apply our methodological approach we must therefore define some variables as inputs and outputs. We consider the daily mean return over the sample period as the main output. Other outputs such as skewness have also been computed from the daily returns distribution. As inputs, the risk of the fund is measured by the standard deviation of the daily returns, as well as kurtosis, also computed from the daily returns. In some of the proposed models the management costs of the fund are also considered as input. In order to include these costs, we consider two variables. The first one is the fees paid from the fund to managers; the second one is the loads, including fees and other costs incurred for operational management, e.g., for turnover. Both variables are measured as percentages (average of the sample period) of costs over the managed portfolio size. Finally, we consider size as a possible source of economies of scale in mutual fund management. We measured size as the average of the amount of the managed assets over the sample period. Descriptive statistics for inputs and outputs are reported in table 1.

5. Results

5.1. FDH efficiency measures

Table 2 reports summary statistics for FDH efficiency. Results are reported for all mutual funds evaluated jointly and also for different size categories. The joint evaluation, for all 205 mutual funds, is reported in the last row of table 2. The different size categories have been constructed to give the same number of funds in each class, so each of them contains 20% of the observations, and each category contains 41 mutual funds.

As one might expect, given the characteristics of FDH, the number of efficient funds is very high. This is partly suggested by the high overall mean efficiency value, which indicates that the average fund has an efficiency value of 96.30%. This result indicates that the total amount of inputs could only be reduced, on average, by 3.70% to catch up with the best practice funds. This result holds for all size categories, whose mean efficiencies range from 95.44%, for the category comprising the largest mutual funds (5^{th} size class), to 98.31%, for the second size class. Indeed, efficiency does not apparently increase with size, thereby suggesting the absence of economies of scale in mutual fund performance.

The high mean values for the different size categories could have been caused by multi-modal distributions, with many efficient funds (in the vicinity of 1) and many inefficient funds, with a thinning in the middle. Table 2 reports different summary statistics of the distributions of efficiency so as to provide a more comprehensive view of the results. Although it could *a priori* seem remarkable that, from the first quartile onwards, all funds are efficient, this trend is actually not surprising if one takes into account that FDH drops the convexity assumption of DEA which, in practical terms, implies that when a given unit cannot be compared with others because of their

Table 2: Descriptive statistics for FDH efficiencies, mutual funds (1999–2007)

Size category (upper limit)[a]	# of funds	Mean	1st quartile	Median	3rd quartile	Std.dev.
9,324,009.54	41	0.9824	0.9801	1.0000	1.0000	0.0328
23,054,509.30	41	0.9867	0.9921	1.0000	1.0000	0.0286
41,828,423.13	41	0.9898	0.9930	1.0000	1.0000	0.0225
101,586,883.90	41	0.9849	0.9798	1.0000	1.0000	0.0259
547,487,533.60	41	0.9857	0.9843	1.0000	1.0000	0.0265
Total	205	0.9859	0.9879	1.0000	1.0000	0.0273

[a] In €.

input/output combinations, it is classified as efficient *by default*. Therefore, the useful property of FDH of being more flexible than DEA comes at the cost of a lower ability to discriminate among efficient funds.

This result is especially apparent from table 3, which reports additional details on FDH results. The last row in table 3 indicates that, out of a total number of 205 funds, 132 *were* efficient, representing a hefty 64.39% of the sample. This implies that, despite the attractive asymptotic properties of FDH referred to above, and despite being more flexible than DEA, FDH has difficulties in both *discriminating* and, more importantly (especially in the context of mutual fund performance evaluation) in *ranking* units. As indicated earlier, it is very well suited to those contexts in which the analyst wants to ascertain the most obvious cases of inefficiency. In our case, it is able to rank 35.61% of the observations, but cannot discriminate among the remaining 64.39% which are included in the efficient category.

Table 3 also provides information on FDH results related to *dominance*. Recall that, under FDH, the frontier is obtained by comparing inputs and outputs in order to establish the *dominant* points (Sampaio de Sousa and Schwengber, 2005). If we define domination as the ability to produce more output with less input, then an observation is declared inefficient if it is dominated by at least another observation. Therefore, if an observation is not dominated by any other, it is classified as FDH efficient and, by construction, inefficient observations are necessarily dominated by one or more other observations. As reported in table 3, a significant number of funds are efficient by default, which constitutes an extreme form of the sparsity bias that ultimately leads to the overestimation of the number of efficient units.

5.2. Expected order-m efficiency estimates

The order-m estimators overcome the difficulties of FDH and DEA for ranking *efficient* funds—i.e., those with a value of 1. We have computed the order-m estimates for different values of m ($m = 25$, $m = 75$ and $m = 150$). These frontiers are nested and, therefore, for $m' > m$, the order-m frontier is below the order-m' frontier. Although the choice of the m parameter might seem somewhat arbitrary, Sampaio de Sousa and Schwengber (2005) have shown that the impact of the decision might not be so relevant when plotting the order-m efficiencies for different values of m, which usually show they are highly correlated. Indeed the choice should not be intricate if one follows Cazals *et al.* (2002), who suggest that "a few values of m could be used to guide the manager of the production unit to evaluate its own performance". In addition, as indicated by Simar (2003), it is also important to notice the difference between m and n. Whereas m is

Table 3: FDH efficiencies, mutual funds (1999–2007)

Size category (upper limit)[a]	Efficient funds						Inefficient funds	
	Efficient funds		Efficient and dominating funds		Funds efficient by default			
	#	%	#	%	#	%	#	%
9,324,009.54	27	65.85	1	3.70	26	96.3	14	34.15
23,054,509.30	29	70.73	0	0.00	29	100	12	29.27
41,828,423.13	27	65.85	1	3.70	26	96.3	14	34.15
101,586,883.90	23	56.10	1	4.35	22	95.65	18	43.90
547,487,533.60	26	63.41	0	0.00	26	100	15	36.59
Total	132	64.39	3	1.46	129	62.93	73	35.61

[a] In €.

a "trimming" parameter fixed at any desired level defining the level of the benchmark, n is the sample size and accordingly, there are no *a priori* links between m and n. This idea of trimming is not new in statistics, although its use in boundary estimation is.

Table 4 reports summary statistics for efficiencies estimated using order-m, considering $m = 25$, $m = 75$ and $m = 150$. Regardless of the value of m, the trimming parameter which allows one to tune the percentage of points that will lie above the order-m frontier, the mean is always much lower than for the FDH case—here 83.10%, 81.85% and 80.14% for $m = 25$, $m = 75$ and $m = 150$, respectively. This could suggest a superior ability of order-m to rank observations, as it turns out to be the case. Overall, and regardless of the size class considered, these results indicate that, on average, the performance of mutual funds could improve much more than what FDH predicts, since many of the *efficient* funds under FDH are not efficient under order-m.

One of the results reported in table 4 which may surprise the reader unfamiliar with partial frontiers is the presence of efficiency scores higher than 1. Although these cases of super-efficiency (Andersen and Petersen, 1993) are present regardless of the m parameter, the only summary statistics' displaying such values are those corresponding to $m = 25$. This occurs because as m increases, the order-m estimator converges to the FDH estimator and, therefore, order-m efficiencies become more similar to FDH efficiencies and those cases above unity tend to disappear, as indicated by the results corresponding to $m = 75$ and $m = 150$. As explained in Simar (2003), for large values of m the two frontiers—FDH and order-m—coincide.

Yet the most interesting results regarding order-m estimates are those reported in tables 5 and 6, which also report results for other values of m ($m = 50$ and $m = 100$). They display specific results for the efficiency score of each mutual fund, ranked by the order-m, with $m = 150$, in both decreasing (table 5) and increasing order (table 6). Note that the last column in table 5 reports, for comparison purposes, the FDH efficiency score for each observation. The results in this column corroborate the difficulties of FDH to discriminate among efficient units. In contrast, regardless of the value of m considered, the 20 "most efficient" or, more correctly, super-efficient funds (according to $m = 150$) can be ranked.

The varied results obtained for the different values of m might be somewhat puzzling, but the only m parameter for which we actually obtain results that are difficult to reconcile is $m = 25$. Indeed, some of the efficiencies obtained for $m = 25$ could be *negative*, suggesting that this might actually not be the right choice of the m parameter. Of course, the negative efficiencies do not carry any particular economic meaning and therefore should be discarded. The results obtained for $m = 75$ and $m = 150$ are more similar among themselves than compared to the $m = 25$

Table 4: Descriptive statistics for order-m efficiency scores, mutual funds (1999–2007)

Size category (upper limit)[a]	Mean	1st quartile	Median	3rd quartile	Std.dev.
$m = 25$					
9,324,009.54	1.0856	1.000	1.0099	1.0508	0.4042
23,054,509.30	1.0759	1.000	1.0163	1.0449	0.2983
41,828,423.13	1.0612	1.000	1.0102	1.0842	0.1051
101,586,883.90	1.0316	1.000	1.0203	1.0428	0.0652
547,487,533.60	1.0265	1.0094	1.0284	1.0421	0.0317
All funds' categories	1.0562	1 .000	1.0174	1.0525	0.2307
$m = 75$					
9,324,009.54	0.9921	0.9947	1.0000	1.0091	0.0376
23,054,509.30	0.9960	1.000	1.000	1.0044	0.0341
41,828,423.13	1.0057	0.9977	1.0000	1.0096	0.0356
101,586,883.90	0.9982	0.9891	1.0000	1.0054	0.0343
547,487,533.60	0.9942	0.9910	1.0000	1.0045	0.022
All funds' categories	0.9973	0.9925	1.0000	1.0056	0.0332
$m = 150$					
9,324,009.54	0.9839	0.9886	1.0000	1.0000	0.0328
23,054,509.30	0.9880	0.9936	1.0000	1.0000	0.0286
41,828,423.13	0.9943	0.9930	1.0000	1.0003	0.0238
101,586,883.90	0.9893	0.983	1.0000	1.0003	0.0266
547,487,533.60	0.9890	0.9884	1.0000	1.0001	0.0231
All funds' categories	0.9889	0.9887	1.0000	1.0000	0.0272

[a] In €.

Table 5: Order-m efficiencies for selected mutual funds, decreasing order (1999–2007)

Fund	$m = 10$		$m = 25$		$m = 50$		$m = 75$		$m = 100$		$m = 150$		FDH-efficiency
	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	
1	1.3710	16	1.2248	8	1.1508	3	1.0896	4	1.0867	1	1.0497	1	1.0000
2	1.2646	23	1.1643	15	1.0958	8	1.0715	5	1.0664	2	1.0411	2	1.0000
3	1.1866	32	1.1044	22	1.0707	14	1.0488	9	1.0358	7	1.0282	3	1.0000
4	1.5483	10	1.3009	5	1.1478	4	1.0955	3	1.0543	3	1.0239	4	1.0000
5	1.2519	26	1.1341	20	1.0805	12	1.0591	7	1.0402	6	1.0215	5	1.0000
6	1.3529	17	1.1487	19	1.0631	16	1.0418	13	1.0312	10	1.0209	6	1.0000
7	1.3307	18	1.1902	12	1.0934	9	1.0429	12	1.0313	9	1.0201	7	1.0000
8	1.2658	22	1.1494	18	1.0742	13	1.0436	11	1.0358	8	1.0181	8	1.0000
9	1.3889	15	1.1675	13	1.0826	11	1.0575	8	1.0439	5	1.0178	9	1.0000
10	2.2632	3	1.4497	3	1.2066	2	1.1118	2	1.0500	4	1.0133	10	1.0000
11	1.0954	90	1.0482	52	1.0279	33	1.0219	20	1.0160	16	1.0094	11	1.0000
12	1.1813	34	1.0904	29	1.0555	18	1.0270	18	1.0165	15	1.0080	12	1.0000
13	1.1904	31	1.0878	30	1.0482	20	1.0322	15	1.0167	14	1.0080	13	1.0000
14	1.1616	43	1.0810	34	1.0443	23	1.0248	19	1.0154	19	1.0078	14	1.0000
15	1.1308	58	1.0616	41	1.031	30	1.0172	26	1.0132	21	1.0067	15	1.0000
16	1.1815	33	1.0961	26	1.0452	22	1.0304	16	1.0152	20	1.0062	16	1.0000
17	1.2627	24	1.1002	24	1.0332	27	1.0198	23	1.0124	22	1.0054	17	1.0000
18	1.1019	86	1.0455	56	1.0181	47	1.0114	35	1.0068	29	1.0049	18	1.0000
19	1.0585	120	1.0340	74	1.0175	48	1.0118	33	1.0103	24	1.0042	19	1.0000
20	1.1738	37	1.0839	32	1.0328	28	1.0203	22	1.0090	26	1.0037	20	1.0000

Table 6: Order-m efficiencies for selected mutual funds, increasing order (1999–2007)

Fund	$m = 10$		$m = 25$		$m = 50$		$m = 75$		$m = 100$		$m = 150$		FDH-efficiency
	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	Order-m efficiency	Rank	
1	0.9514	197	0.8993	203	0.8834	204	0.8698	205	0.8682	205	0.8634	205	0.8607
2	0.9137	204	0.8874	205	0.8763	205	0.8744	204	0.8740	204	0.8736	204	0.8736
3	0.9721	193	0.9272	198	0.9087	201	0.9014	202	0.8988	203	0.8955	203	0.8947
4	0.8992	205	0.8992	204	0.8992	203	0.8992	203	0.8992	202	0.8992	202	0.8992
5	0.9633	196	0.9194	200	0.9099	199	0.9062	200	0.9054	201	0.9053	201	0.9053
6	0.9303	202	0.9080	202	0.9061	202	0.9061	201	0.9061	200	0.9061	200	0.9061
7	0.9192	203	0.9105	201	0.9098	200	0.9098	199	0.9098	199	0.9098	199	0.9098
8	1.0398	133	0.9577	192	0.9275	197	0.9182	197	0.9174	197	0.9161	198	0.9161
9	0.9447	200	0.9219	199	0.9178	198	0.9170	198	0.9170	198	0.9170	197	0.9170
10	1.0195	146	0.9612	191	0.9417	194	0.9325	194	0.9255	195	0.9228	196	0.9209
11	0.9836	192	0.9524	193	0.9334	195	0.9268	196	0.9247	196	0.9240	195	0.9240
12	0.9380	201	0.9289	197	0.9278	196	0.9278	195	0.9278	194	0.9278	194	0.9278
13	1.1154	72	1.0262	91	0.9838	177	0.9645	183	0.9523	189	0.9383	193	0.9169
14	0.9635	195	0.9479	194	0.9418	193	0.9409	193	0.9404	193	0.9402	192	0.9402
15	0.9453	199	0.9442	196	0.9440	192	0.9440	192	0.9440	192	0.9440	191	0.9440
16	0.9484	198	0.9449	195	0.9449	191	0.9449	191	0.9449	191	0.9449	190	0.9449
17	1.1560	44	1.0540	48	0.9980	152	0.9663	182	0.9600	183	0.9455	189	0.9331
18	1.0933	93	1.0119	111	0.9787	178	0.9700	177	0.9611	181	0.9470	188	0.9397
19	1.1238	66	1.0274	87	0.9943	160	0.9744	174	0.9589	185	0.9472	187	0.9193
20	1.0739	110	0.9882	182	0.9605	189	0.9544	190	0.9515	190	0.9482	186	0.9465

case. Table 6 reports analogous results as table 5 but arrayed in decreasing order, according to $m = 150$. In this case results are almost coincidental for the different m values, especially for $m = 75$ and $m = 150$.

In light of these results, although some readers might be tempted to discard these methods due to the need to select an m parameter somewhat arbitrarily, the order-m technique has the remarkable virtue of ranking the mutual funds with the best performance, as table 5 shows. Although results are not entirely coincidental for all values of m considered, the correlation is very high. In these circumstances, a suitable criterion for selecting funds could consist of selecting only those that are classified in a given percentile according to different values of m, which in practical terms implies making these robust methods even more robust. We will consider a different way to do this based on an analysis of persistence.

5.3. Results for order-α quantile frontiers

Results for the order-α partial frontier are reported in table 7 for $\alpha = .90$, $\alpha = .95$ and $\alpha = .99$. The table reports analogous information to that reported in table 4 for the case of the order-m frontiers. In this case of order-α frontiers, the impact of the α parameter seems *a priori* stronger, yet this only occurs because the range of variation is not exactly equivalent to that chosen for m. Therefore, as indicated in table 7, the average efficiency for $\alpha = .90$ is quite high (148.48%), which implies the presence of super-efficient funds whose efficiencies are remarkably high. As shown by the standard deviation, these are only very specific units which cause a lot of variation. As the α parameter increases, the standard deviation decreases sharply and results converge to FDH.

Tables 8 and 9 provide analogous results to tables 5 and 6 for the order-α case. These tables also report results for other values of α ($\alpha = .50$, $\alpha = .80$ and $\alpha = .975$). One of the main differences one may perceive between results for order-m and order-α is the great impact of α, which makes the order-α to converge to FDH efficiencies much faster. But the same bottom line should apply for order-α results: those funds which perform better regardless of the α parameter considered should be the best candidates for selection by the investor. The analysis in the next section provides further insights on this point.

Table 7: Descriptive statistics for order-α efficiency scores, mutual funds (1999–2007)

Size category (upper limit)[a]	Mean	1st quartile	Median	3rd quartile	Std.dev.
$\alpha = .90$					
9,324,009.54	1.4592	1.0000	1.0594	1.1448	2.0943
23,054,509.30	1.4001	1.0000	1.0697	1.1976	1.6578
41,828,423.13	1.1958	1.0000	1.0959	1.2193	0.2795
101,586,883.90	1.1185	1.0000	1.0819	1.1756	0.1615
547,487,533.60	1.1665	1.0560	1.1039	1.1611	0.2848
Total	1.2664	1.0000	1.0873	1.1825	1.1969
$\alpha = .95$					
9,324,009.54	1.0472	1.0000	1.0000	1.0810	0.1214
23,054,509.30	1.0425	1.0000	1.0000	1.0635	0.0957
41,828,423.13	1.0727	1.0000	1.0000	1.1115	0.1363
101,586,883.90	1.0519	1.0000	1.0039	1.0861	0.0977
547,487,533.60	1.0332	1.0000	1.0327	1.0580	0.0411
Total	1.0496	1.0000	1.0000	1.0755	0.1034
$\alpha = .99$					
9,324,009.54	0.9844	0.9863	1.0000	1.0000	0.0334
23,054,509.30	0.9874	0.9921	1.0000	1.0000	0.0302
41,828,423.13	0.9986	0.9977	1.0000	1.0000	0.0356
101,586,883.90	0.9920	0.9883	1.0000	1.0000	0.0311
547,487,533.60	0.9914	1.0000	1.0000	1.0000	0.0227
Total	0.9908	0.9921	1.0000	1.0000	0.0310

[a] In €.

Table 8: Order-α efficiencies for selected mutual funds, decreasing order (1999–2007)

Fund	$\alpha = .50$		$\alpha = .80$		$\alpha = .90$		$\alpha = .95$		$\alpha = .975$		$\alpha = .99$		FDH-efficiency
	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	
1	1.0407	105	1.0000	46	1.0000	13	1.0000	1	1.0000	1	1.0000	1	1.0000
2	1.3614	25	1.1522	12	1.0966	5	1.0000	2	1.0000	2	1.0000	2	1.0000
3	1.2109	42	1.0000	47	1.0000	14	1.0000	3	1.0000	3	1.0000	3	1.0000
4	1.1321	53	1.0478	24	1.0000	15	1.0000	4	1.0000	4	1.0000	4	0.9664
5	1.0687	80	1.0000	48	1.0000	16	1.0000	5	1.0000	5	1.0000	5	1.0000
6	1.0472	93	1.0000	49	1.0000	17	1.0000	6	1.0000	6	1.0000	6	0.9449
7	1.0472	92	1.0096	40	1.0000	18	1.0000	7	1.0000	7	1.0000	7	0.9278
8	1.0000	168	1.0000	50	1.0000	19	1.0000	8	1.0000	8	1.0000	8	1.0000
9	1.0000	169	1.0000	51	1.0000	20	1.0000	9	1.0000	9	1.0000	9	1.0000
10	1.3778	24	1.0000	52	1.0000	21	1.0000	10	1.0000	10	1.0000	10	1.0000
11	1.0000	170	1.0000	53	1.0000	22	1.0000	11	1.0000	11	1.0000	11	1.0000
12	1.0682	81	1.0000	54	1.0000	23	1.0000	12	1.0000	12	1.0000	12	0.8992
13	6.6868	5	1.0000	55	1.0000	24	1.0000	13	1.0000	13	1.0000	13	1.0000
14	1.0000	171	1.0000	56	1.0000	25	1.0000	14	1.0000	14	1.0000	14	1.0000
15	1.0029	166	1.0000	57	1.0000	26	1.0000	15	1.0000	15	1.0000	15	1.0000
16	1.0000	172	1.0000	58	1.0000	27	1.0000	16	1.0000	16	1.0000	16	1.0000
17	50.8862	2	41.8645	1	11.5030	1	1.0000	17	1.0000	17	1.0000	17	1.0000
18	1.0000	173	1.0000	59	1.0000	28	1.0000	18	1.0000	18	1.0000	18	1.0000
19	1.0000	174	1.0000	60	1.0000	29	1.0000	19	1.0000	19	1.0000	19	1.0000
20	1.0479	90	1.0066	44	1.0000	30	1.0000	20	1.0000	20	1.0000	20	0.9170

Table 9: Order-α efficiencies for selected mutual funds, increasing order (1999–2007)

Fund	$\alpha = .50$		$\alpha = .80$		$\alpha = .90$		$\alpha = .95$		$\alpha = .975$		$\alpha = .99$		FDH-efficiency
	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	Order-α efficiency	Rank	
1	2.2242	11	1.1617	11	0.9645	82	0.8405	186	0.7905	191	0.5627	205	1.0000
2	1.4191	22	1.0238	34	0.9503	119	0.8709	152	0.8190	155	0.5831	204	1.0000
3	1.0695	79	0.9971	89	0.9613	87	0.8810	133	0.8286	144	0.5898	203	1.0000
4	1.0208	137	0.9801	137	0.9609	90	0.8709	151	0.8190	154	0.6038	202	1.0000
5	1.3091	31	1.0227	35	0.9499	124	0.9045	107	0.8571	116	0.6102	201	1.0000
6	1.1281	57	1.0084	42	0.9525	116	0.9139	99	0.8595	112	0.6119	200	1.0000
7	1.0885	68	1.0071	43	0.9801	68	0.8734	147	0.8214	153	0.6187	199	1.0000
8	1.1015	64	1.0387	27	1.0000	49	0.9342	80	0.8786	93	0.6254	198	1.0000
9	1.1635	47	1.0511	23	1.0000	50	0.9296	83	0.8810	90	0.6271	197	1.0000
10	1.0973	65	1.0598	19	1.0276	8	1.0000	37	0.6344	205	0.6344	196	1.0000
11	1.0808	72	1.0000	80	0.9530	113	0.8668	158	0.8214	152	0.6351	195	0.9879
12	2.0498	12	1.1918	8	1.0000	46	0.9523	67	0.7814	197	0.6424	194	0.9942
13	1.0787	73	0.9942	99	0.9448	133	0.8593	168	0.8143	163	0.6536	193	0.9822
14	1.0289	128	0.9799	138	0.9526	115	0.8814	131	0.8143	162	0.6556	192	0.9789
15	1.0445	99	0.9913	104	0.9581	100	0.8840	124	0.8167	159	0.6779	191	0.9439
16	1.0623	83	0.9855	118	0.9631	86	0.9177	96	0.8518	126	0.6806	190	0.9397
17	1.0030	164	0.9566	179	0.9246	167	0.8531	176	0.7881	194	0.6825	189	1.0000
18	1.1947	44	0.9822	130	0.9397	142	0.8557	171	0.7905	190	0.6845	188	1.0000
19	1.1641	46	0.9852	122	0.9302	154	0.8582	169	0.7929	186	0.6848	187	0.9331
20	1.1985	43	0.9852	121	0.9380	145	0.8517	182	0.7929	185	0.6866	186	1.0000

5.4. On the links between partial frontiers and persistence analysis in mutual fund evaluation

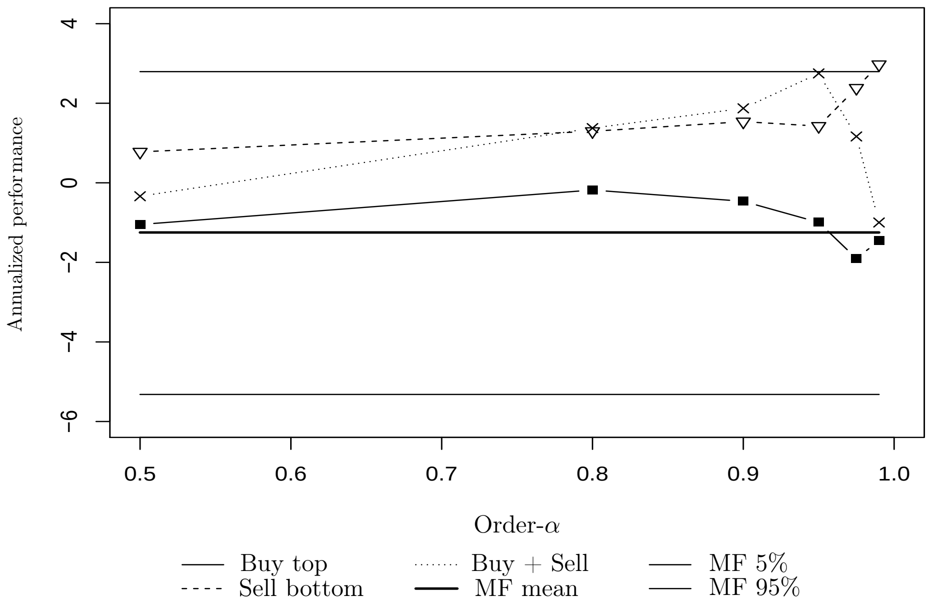
The preceding sections have estimated partial frontiers (order-m and order-α) to evaluate mutual fund efficiency. According to what has been described in section 3, we now analyze the informativeness of combining these methods with a persistence analysis for mutual fund performance evaluation. We consider the performance of these methods to be fair, or appropriate, if they are able to forecast mutual fund efficiencies when they are persistent. Thus, we evaluate the performance of each method, focusing on how the choice of tuning parameters—m in the case of order-m and α in the case of order-α—might influence the results. In order to do this, we consider first the mutual funds' rankings according to their *past* efficiency provided by partial frontiers. Next, based on this information, we construct equally-weighted portfolios that follow several strategies: (i) buying past top-quantile and selling past-bottom quantile (buy + sell); (ii) buying past top quantile (buy); and (iii) selling bottom quantile (sell). The quantiles considered define deciles, quintiles and half of the mutual funds' past efficiency distribution. This leads to a selection of 36 portfolios which follow dynamic strategies, namely, 3 (strategies) $\times$ 2 (efficiency methods, order-m and order-α) $\times$ 6 (parameters for each of them). Finally, we assess the performance of these portfolios using expression (9), corresponding to the multifactor linear model.

We report results in figures 1 and 2. They show that, for any of the selected *trimming* parameters—either for order-m or order-α, i.e., m and α—the strategies based on deciles achieve better performance than those based on quintiles or the median. This indicates that persistence in mutual fund performance is generally focused in the extreme mutual funds, and not in the middle of the distribution of past efficiency. Consequently, only results for strategies based on deciles are presented in figures 1 and 2, specifically the annualized performance, measured by α_p^J in equation (9), of the strategies based on past mutual fund efficiency estimated respectively using order-α (order-m) methods.[5]

We should also consider that the portfolios referred to in the preceding paragraph periodically buy and/or sell funds drawn from our sample. If these funds themselves have good or bad performance, those other portfolios which invest in them will perform similarly. Therefore, results yielded by these portfolios have to be evaluated in comparison with those funds whose efficiency has been measured during the same period. Accordingly, we will be able to measure the value added of adopting strategies which follow recommendations of buy and/or sell funds based on

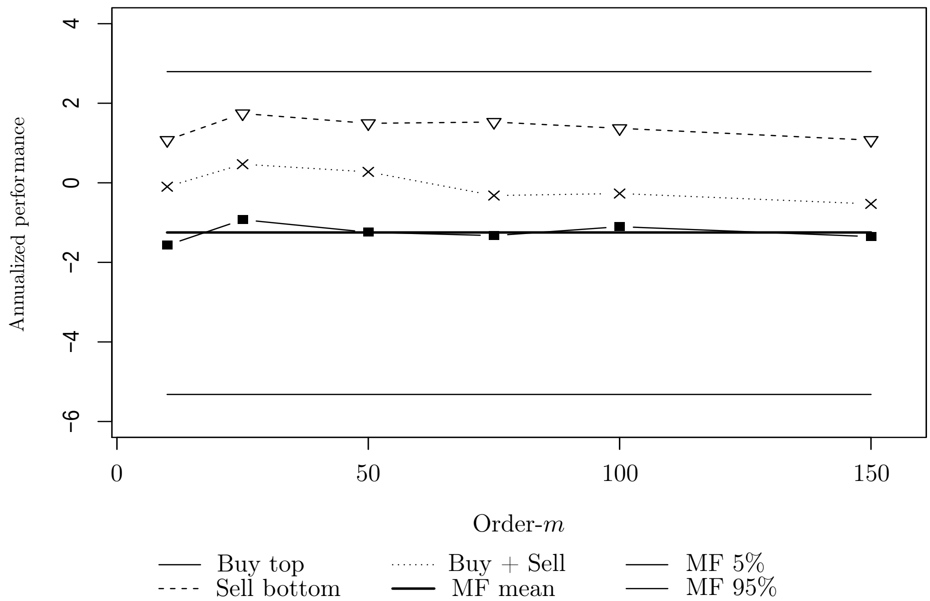
[5] For simplicity, figures 1 and 2 report results for the deciles only.

Figure 1: Persistence, order-α



Notes: Annualized performance of rebalanced portfolios that follow strategies according to past efficiency measured by the order-α method: (i) buy top decile mutual funds; (ii) sell bottom decile; and (iii) buy top decile and sell bottom decile. The mean and the 5% and 95% values of the distribution of the annualized mutual fund performance over the whole sample period are also shown. The horizontal axis represents different levels of the α parameter.

Figure 2: Persistence, order-m



Notes: Annualized performance of rebalanced portfolios that follow strategies according to past efficiency measured by the order-m method: (i) buy top decile mutual funds; (ii) sell bottom decile; and (iii) buy top decile and sell bottom decile. The mean and the 5% and 95% values of the distribution of the annualized mutual fund performance over the whole sample period are also shown. The horizontal axis represents different levels of the m parameter.

the rankings provided by partial frontiers.

An appropriate method to evaluate mutual fund efficiency should provide us with relevant information to obtain enhanced results with respect to the funds in the sample. Therefore, in order to allow for this comparability, we also apply equation (9) to mutual funds in the sample and for the entire sample period. Accordingly, it will be possible to build the distribution of mutual funds' performance to compare portfolios' performances.

In both figures we draw lines for the 5% quantile, 95% quantile, and the mean of the mutual fund performance distribution. This is based on equation (9) and for the entire sample period. In general, portfolios based on past efficiency provide a performance better than, or equal to, the mean of the mutual funds for the entire sample period. More specifically, in both figures the strategy of buying top past efficiency deciles achieves the worst performance of the three strategies analyzed, and is close to the mean of the mutual funds' performance. The best performance result is achieved by following the strategy of selling the mutual funds from the last decile based on past efficiency. This evidence is similar to findings by Carhart (1997), Lynch and Musto (2003), Loon (2011) and, for the case of Spanish mutual funds, Menéndez and Álvarez (2000). This implies that the worst mutual funds are more persistent over time than better funds. Therefore, it is interesting to stress that the results yielded by partial frontiers, which maintain the advantages of both DEA and FDH (i.e., higher flexibility to handle several inputs and outputs) are consistent with the evidence found in the financial literature, even when we focus on those studies that have used other samples of funds, periods and, especially, have applied very different methodologies to measure mutual fund efficiency.

A comparison of figures 1 and 2 reveals that the order-m results are less sensitive to the (trimming) parameter variation than those yielded by order-α, and the performance of the different strategies is more stable, although the range of variation is not directly comparable. Thus, for any m parameter, the best strategy is to sell the bottom decile mutual funds according to past efficiency. However, performance based on past efficiency according to order-α is more sensitive to the α parameter selected. This is especially the case for the buy-sell strategy, because order-α shows less power to provide a robust ranking. In other words, in some cases many mutual funds obtain a score of 1 at the top of the ranking and order-α has problems in identifying those funds in the top decile. Nevertheless, the best performance is achieved by following a strategy based on selling the bottom decile mutual funds according to past efficiency measured by higher α parameters in the order-α method.

6. Conclusions

The mutual fund industry is immersed in a process of continuous expansion and changes and, therefore, its analysis is gaining importance over time. A financial investor has to consider not only the actual number of competitors (i.e. the number of funds in the same objective category) but also the fact that it could vary up and down. Like other financial industries it is subject to expansion, acceleration and contraction cycles which greatly affect the performance of firms managing the assets.

The interest comes from academics and investment industry participants alike. Since the traditional methodologies were initially proposed, many studies have been developed around the evaluation of mutual funds. In more recent years the importance of using nonparametric approaches for mutual fund performance evaluation has been stressed because of the key benefits they offer. These tools provide a single value of efficiency and have the great advantage of allowing one to include a high number of inputs and outputs in the model specified. Although these techniques, basically DEA and its non-convex sibling, FDH, are not free from disadvantages, the literature applying them to evaluate the performance of mutual funds is becoming voluminous.

Some of the disadvantages of both DEA and FDH are the sensitivity to both outliers and the curse of dimensionality. This has been recognized by the literature (Dyson *et al.*, 2001). Yet in recent times some methods have been proposed in order to overcome these pitfalls. The order-m (Cazals *et al.*, 2002) and order-α (Daouia and Simar, 2007) partial frontiers are more robust both to the presence of outliers and the curse of dimensionality. Although they require selecting some parameters which may be difficult, their advantages outweigh their disadvantages. In the specific case of mutual funds' performance evaluation, they have the great ability of ranking all mutual funds.

Yet applications of these techniques are still scarce. Only Daraio and Simar (2005, 2006, 2007b) have considered order-m techniques, although their contributions are theoretical. We extend their applications by considering not only order-m but also order-α partial frontiers. We also have a much tighter focus on the application. Specifically, we measure the performance of a sample of Spanish mutual funds for the 1998–2007 period. Therefore, we uncover a long and recent period, just before the recent international economic and financial crisis hit most developed countries hard. Applying both order-m and order-α methods to our sample of mutual funds enables us to provide a full ranking. In contrast to FDH, which cannot discriminate among efficient funds, both order-m and order-α efficiency scores provide a full ranking of the mutual funds in our sample. Although some readers might be puzzled by the fact that results vary for different m and α values, most funds rank very well regardless of the value of m or α considered.

We combine these methods and results with the literature on mutual fund performance persistence. This facilitates obtaining an algorithm to guide the choice of m and α parameters which is a relevant decision intrinsic to order-m and order-α , respectively. Specifically, using the multifactor linear model, we will be able to conclude whether the performance of either order-m or order-α is satisfactory if they are able to forecast mutual funds' efficiencies when they are persistent. According to our methods, the best performance is achieved when the strategy followed is based on selling the bottom decile mutual funds according to past efficiency measured by higher α parameters in the order-α method.

References

Andersen, P. and Petersen, N. C. (1993). A procedure for ranking efficient units in Data Envelopment Analysis. *Management Science*, 39(10):1261–1264.

Annaert, J., Van den Broeck, J., and Vander Vennet, R. (2003). Determinants of mutual fund underperformance: a Bayesian stochastic frontier approach. *European Journal of Operational Research*, 151(3):617–632.

Aragon, Y., Daouia, A., and Thomas-Agnan, C. (2005). Nonparametric frontier estimation: A conditional quantile-based approach. *Econometric Theory*, 21(2):358–389.

Banker, R. D., Charnes, A., and Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in Data Envelopment Analysis. *Management Science*, 30:1078–1092.

Banker, R. D. and Maindiratta, A. (1986). Piecewise loglinear estimation of efficient production surfaces. *Management Science*, 32(1):126–135.

Banker, R. D. and Morey, R. C. (1986). The use of categorical variables in Data Envelopment Analysis. *Management Science*, 32:1613–1627.

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

Bollen, N. P. B. and Busse, J. A. (2005). Short-term persistence in mutual fund performance. *Review of Financial Studies*, 18(2):569–597.

Briec, W. and Kerstens, K. (2009). Multi-horizon Markowitz portfolio performance appraisals: a general approach. *Omega*, 37(1):50–62.

Briec, W., Kerstens, K., and Jokung, O. (2007). Mean-variance-skewness portfolio performance gauging: A general shortage function and dual approach. *Management science*, 53(1):135–149.

Briec, W., Kerstens, K., and Lesourd, J. B. (2001). Markowitz portfolio selection, performance gauging and duality: Generalising Luenberger's shortage function. Document de Travail 01A25, GREQAM, Marseille.

Briec, W., Kerstens, K., and Lesourd, J. B. (2004). Nonparametric tests of portfolio investment efficiency: A shortage function generalization. *Journal of Optimization Theory and Applications*, 120(1):1–27.

Brown, S. J. and Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance*, pages 679–698.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Cazals, C., Florens, J.-P., and Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106:1–25.

Cesari, R. and Panetta, F. (2002). The performance of Italian equity funds. *Journal of Banking & Finance*, 26(1):99–126.

Chang, K. P. (2004). Evaluating mutual fund performance: an application of minimum convex input requirement set approach. *Computers and Operations Research*, 31(6):929–940.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6):429–444.

Choi, Y. K. and Murthi, B. P. S. (2001). Relative performance evaluation of mutual funds: A non-parametric approach. *Journal of Business Finance & Accounting*, 28(7 & 8):853–876.

Cremers, K. J. and Petajisto, A. (2009). How active is your fund manager? A new measure that predicts performance. *Review of Financial Studies*, 22(9):3329–3365.

Daouia, A. and Simar, L. (2007). Nonparametric efficiency analsyis: A multivariate conditional quantile approach. *Journal of Econometrics*, 140:375–400.

Daraio, C. and Simar, L. (2005). Introducing environmental variables in nonparametric frontier models: a probabilistic approach. *Journal of Productivity Analysis*, 24:93–121.

Daraio, C. and Simar, L. (2006). A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research*, 175(1):516–542.

Daraio, C. and Simar, L. (2007a). *Advanced Robust and Nonparametric Methods in Efficiency Analysis. Methodology and Applications*. Studies in Productivity and Efficiency. Springer, New York.

Daraio, C. and Simar, L. (2007b). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of Productivity Analysis*, 28(1):13–32.

Droms, W. G. (2006). Hot Hands, Cold Hands: Does Past Performance Predict Future Returns? *Journal of Financial Planning*, 19(5):60–69.

Dyson, R. G., Allen, R., Camanho, A. S., Podinovski, V. V., Sarrico, C. S., and Shale, E. A. (2001). Pitfalls and protocols in DEA. *European Journal of Operational Research*, 132(2):260–273.

Eling, M. (2006). Performance measurement of hedge funds using data envelopment analysis. *Financial Markets and Portfolio Management*, 20(4):442–471.

Eling, M. and Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*, 31(9):2632–2647.

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

Farrell, M. J. (1957). The measurement of productive efficiency. *Journal of the Royal Statistical Society*, Ser.A,120:253–281.

Glawischnig, M. and Sommersguter-Reichmann, M. (2010). Assessing the performance of alternative investments using non-parametric efficiency measurement approaches: Is it convincing? *Journal of Banking & Finance*, 34(2):295–303.

Gregoriou, G. N. and Zhu, J. (2005). *Evaluating Hedge Fund and CTA performance: Data Envelopment Analysis Approach*. Wiley, Hoboken, NJ.

Grinblatt, M. and Titman, S. (1992). The persistence of mutual fund performance. *The Journal of Finance*, 47(5):1977–1984.

Grinblatt, M. and Titman, S. (1995). Performance evaluation. In Jarrow, R., Maksimovic, V., and Ziemba, W. T., editors, *Handbooks in OR & MS: Finance*, volume 9, pages 581–609. Elsevier, Amsterdam.

Hendricks, D., Patel, J., and Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974-1988. *The Journal of Finance*, 48(1):93–130.

Ippolito, R. A. (1993). On studies of mutual fund performance, 1962–1991. *Financial Analysts Journal*, 49(1):42–50.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance*, 23:389–416.

Joro, T. and Na, P. (2002). Data Envelopment Analysis in mutual fund evaluation: a critical review. Research Report 02-2, Department of Finance and Management Science, School of Business, University of Alberta, Edmonton, Alberta.

Joro, T. and Na, P. (2006). Portfolio performance evaluation in a mean–variance–skewness framework. *European Journal of Operational Research*, 175(1):446–461.

Kamakura, W. A. (1988). A Note on "The Use of Categorical Variables in Data Envelopment Analysis". *Management Science*, 34(10):1273–1276.

Koenker, R. (2001). Quantile regression. *Journal of Economic Perspectives*, 15(4):143–156.

Kuosmanen, T. (2007). Performance measurement and best-practice benchmarking of mutual funds: combining stochastic dominance criteria with data envelopment analysis. *Journal of Productivity Analysis*, 28(1):71–86.

Kuosmanen, T., Cherchye, L., and Sipiläinen, T. (2006). The law of one price in data envelopment analysis: Restricting weight flexibility across firms. *European Journal of Operational Research*, 170(3):735–757.

Loon, Y. C. (2011). Model uncertainty, performance persistence and flows. *Review of Quantitative Finance and Accounting*, 36(2):153–205.

Lovell, C. A. K. and Kumbhakar, S. C. (2000). *Stochastic Frontier Analysis*. Cambridge University Press, Cambridge.

Lozano, S. and Gutiérrez, E. (2007). Tsd-consistent performance assessment of mutual funds. *Journal of the Operational Research Society*, 59(10):1352–1362.

Lynch, A. W. and Musto, D. K. (2003). How investors interpret past fund returns. *The Journal of Finance*, 58(5):2033–2058.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1):77–91.

Matallín-Sáez, J. C. (2006). Seasonality, market timing and performance amongst benchmarks and mutual fund evaluation. *Journal of Business Finance & Accounting*, 33(9-10):1484–1507.

McMullen, P. R. and Strong, R. A. (1998). Selection of mutual funds using Data Envelopment Analysis. *Journal of Business and Economic Studies*, 4(1):1–14.

Menéndez, S. and Álvarez, S. (2000). La rentabilidad y persistencia de los resultados de los fondos de inversión españoles de renta variable. *Revista Española de Financiación y Contabilidad*, 103:15–36.

Morey, M. R. and Morey, R. C. (1999). Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking. *Omega*, 27(2):241–258.

Murillo-Zamorano, L. R. (2004). Economic efficiency and frontier techniques. *Journal of Economic Surveys*, 18(1):33–77.

Murthi, B. P. S., Choi, Y. K., and Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: a nonparametric measurement. *European Journal of Operational Research*, 98:408–418.

Nguyen-Thi-Thanh, H. (2006). On the use of Data Envelopment Analysis in hedge fund selection. Working paper, Université d'Orléans.

Park, B. U., Simar, L., and Weiner, C. (2000). The FDH estimator for productivity efficiency scores. *Econometric Theory*, 16(6):855–877.

Pástor, L. and Stambaugh, R. (2002). Mutual fund performance and seemingly unrelated assets. *Journal of Financial Economics*, 63(3):315–349.

Pätäri, E. J. (2009). Do hot hands warm the mutual fund investor? The myth of performance persistence phenomenon. *International Research Journal of Finance and Economics*, 34:117–139.

Sampaio de Sousa, M. C. and Schwengber, S. B. (2005). Efficiency estimates for judicial services in Brazil: Nonparametric FDH and the expected order-m efficiency scores for Rio Grande do Sul courts. In *Anais do XXXIII Encontro Nacional de Economia*. ANPEC-Associação Nacional dos Centros de Pósgraduação em Economia.

Sengupta, J. K. (1991). Maximum probability dominance and portfolio theory. *Journal of Optimization Theory and Applications*, 71(2):341–357.

Sengupta, J. K. (2000). *Dynamic and Stochastic Efficiency Analysis: Economics of Data Envelopment Analysis*. World Scientific Publishers, Singapore.

Sengupta, J. K. and Park, H. S. (1993). Portfolio efficiency tests based on stochastic dominance and co-integration. *International Journal of Systems Science*, 24(11):2135–2158.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business*, 39(1):119–138.

Simar, L. (2003). Detecting outliers in frontier models: A simple approach. *Journal of Productivity Analysis*, 20(3):391–424.

Simar, L. and Wilson, P. W. (2008). Statistical inference in nonparametric frontier models: Recent developments and perspectives. In Fried, H., Lovell, C. A. K., and Schmidt, S. S., editors, *The Measurement of Productive Efficiency*, chapter 4, pages 421–521. Oxford University Press, Oxford, 2nd edition.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review*, 43(1):63–75.

Tulkens, H. (1993). On FDH efficiency analysis: Some methodological issues and applications to retail banking, courts, and urban transit. *Journal of Productivity Analysis*, 4:183–210.

Wheelock, D. C. and Wilson, P. W. (2009). Robust nonparametric quantile estimation of efficiency and productivity change in U.S. commercial banking, 1985–2004. *Journal of Business and Economic Statistics*, 27(3):354–368.

Wilkens, K. and Zhu, J. (2001). Portfolio evaluation and benchmark selection. *The Journal of Alternative Investments*, 4(1):9–19.

Wilson, P. W. (1993). Detecting outliers in deterministic nonparametric frontier models with multiple outputs. *Journal of Business and Economic Statistics*, 11(3):319–23.

Wilson, P. W. (1995). Detecting influential observations in Data Envelopment Analysis. *Journal of Productivity Analysis*, 6(1):27–45.

Wilson, P. W. (2008). FEAR: A software package for frontier efficiency analysis with R. *Socio-Economic Planning Sciences*, 42(4):247–254.

serie ec

Ivie

Guardia Civil, 22 - Esc. 2, 1º
46020 Valencia - Spain
Phone: +34 963 190 050
Fax: +34 963 190 055

Website: http://www.ivie.es
E-mail: publicaciones@ivie.es