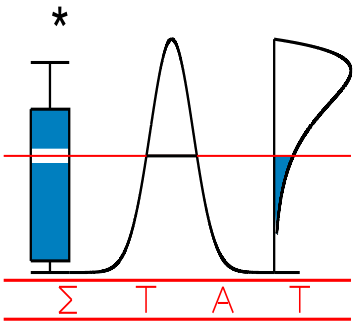
T E C H N I C A L

R E P O R T

0413

A ROBUST NONPARAMETRIC APPROACH TO EVALUATE AND EXPLAIN THE PERFORMANCE OF MUTUAL FUNDS

DARAIO, C. and L. SIMAR



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

A Robust Nonparametric Approach to Evaluate and Explain the Performance of Mutual Funds*

Cinzia Daraio
IIT-CNR and Scuola Superiore S. Anna, Pisa (Italy)
E-mail: cinzia@sssup.it

Léopold Simar
Institut de Statistique, Université Catholique de Louvain (Belgium)
E-mail: simar@stat.ucl.ac.be

March 19, 2004

Abstract

The topic of the measurement of mutual funds' performance is receiving an increasing interest both from an applied and a theoretical perspective. Beside the traditional financial literature, a growing body of works has started to apply the tools of frontier analysis for benchmarking comparisons in portfolio analysis. Our paper contributes to this literature proposing a robust nonparametric approach for analysing mutual funds. It is based on the concept of order$-m$ frontier (Cazals, Florens and Simar, 2002) and on a probabilistic approach (Daraio and Simar, 2003) to find out the factors explaining mutual funds' performance. The usefulness of this approach is illustrated by using US mutual funds data, grouped for category by objective. Economies of scale, slacks and market risks are investigated. A comparison of traditional, nonparametric and robust performance measures is also offered.

Keywords: economies of scale, mutual funds, nonparametric frontier, portfolio analysis, robust estimation.
JEL Classification: C14, C15, C50, C61, G11

*This paper has been presented during a seminar at the Dipartimento di Scienze Economiche, University of Verona, the 23 February 2004. We acknowledge the stimulating comments of seminar's participants. We would like also to thank Carlo Bianchi for his helpful suggestions on a previous version of this paper. The usual disclaimers apply. Research support from "Projet d'Actions de Recherche Concertées" (No. 98/03–217) and from the "Interuniversity Attraction Pole", Phase V (No. P5/24) from the Belgian Government (Belgian Science Policy) is acknowledged.

1 Introduction

The literature on mutual funds performance evaluation is rich both from a methodological and an empirical point of view.

Treynor (1965) proposes to adjust the excess return of a portfolio (with respect to the risk free return) by the portfolio's β, using the CAPM (Capital Asset Pricing Model) introduced by Markowitz (1952, 1959) and developed by Lintner (1965). Similarly, Jensen (1968)'s alpha (α) is defined as the difference between the actual excess portfolio return and the (estimated) expected excess benchmark return. The benchmark could be based on either the CAPM or on the APT (Arbitrage Pricing Theory) model developed by Ross (1976).

Some empirical applications (see e.g. Elton et al. 1993, Choi 1995) have shown that the Jensen's alpha is sensitive to the choice of the benchmark model employed for comparison. It has been argued (see e.g. Admati and Ross 1985) that the estimation of Jensen's alpha may be biased due to *market timing*, which is the ability of fund managers to systematically change the target risk of the fund. When portfolio managers change the target beta for the fund by moving money among different investments, estimation bias will be introduced into the benchmark model because it assumes a constant beta coefficient over the period considered.

The Sharpe (1966) index is defined as the ratio of the excess return of the portfolio (with respect to the risk free return) to the standard deviation of its return. It measures the risk premium earned per unit of risk taken. With respect to Jensen's alpha, the Sharpe index avoids the problem of the specification of the benchmark model. The Sharpe index is more robust to the market index because it uses standard deviation as a risk measure, but it does not totally eliminate the market index. In fact, the final Sharpe index of a portfolio is compared to that-one of the market index. However, even this index does not take into account the *transaction costs*, i.e. the expenses associated with the purchase and sale of assets.

Since the pioneering works by Treynor, Sharpe and Jensen, a lot of performance measures have been introduced and empirically applied for evaluating the performance of mutual funds.[1]

In recent years, there is a growing body of works that applies efficiency and productivity techniques[2] for evaluating the performance of mutual funds. The problem of estimating

[1]For a nice summary see Cesari and Panetta (2002) which report also an application on Italian Equity Funds. General surveys can be found in Shukla and Trzcinca (1992), Ippolito (1993) and Grinblatt and Titman (1995).

[2]Starting from the first empirical application by Farrell (1957), a huge literature has been developed. For a recent review on nonparametric approach models, see Cooper, Seiford and Zhu (2004); for an updated description of parametric models, see Kumbhakar and Lovell (2000).

monotone concave boundaries naturally occurs in portfolio management, as well as in the production setting. In Capital Assets Pricing Models (CAPM) the objective is to analyze the performance of investment portfolios. Risk (volatility or variance) and average return on a portfolio are analogous to inputs and outputs in model of production. In Capital Assets Pricing Models, the boundary of the attainable set of portfolios gives a benchmark relative to which the efficiency of a portfolio can be measured.

Works which apply the parametric approach in frontier analysis to mutual funds include Briec and Lesourd (2000), where an application of the stochastic parametric approach is provided, and Annaert, van den Broeck and Vennet (2003) which apply the stochastic bayesian approach (van den Broeck, Koop, Osiewalski and Steel, 1994).

Among the nonparametric approach, we can distinguish between a theoretical view (e.g. Sengupta, 1991; Sengupta and Park, 1993; Briec, Kerstens and Lesourd, 2001) and a more applied perspective (e.g. Murthi, Choi and Desai, 1997; Morey and Morey, 1999; Sengupta, 2000).

Sengupta (1991) and Sengupta and Park (1993) provide links between CAPM and nonparametric estimation of frontiers from a theoretical point of view. Briec, Kerstens and Lesourd (2001) analyze the relation between the hypothesis of the basic Markowitz model and efficiency analysis theory, developing a dual framework for assessing the degree of satisfaction of investor's preference.

The main object of this paper is to introduce a robust nonparametric method to evaluate and explain the performance of mutual funds.

We base on Murthi, Choi and Desai (1997) study the choice of most US mutual funds variables to analyse in this paper. Moreover, Sengupta (2000) uses market risks of mutual funds (the percentage of fund's movements that can be explained by movements in its benchmark index) as an input in his analysis underlying that the effect of market risks is conducive for mutual funds performance. Applying a probabilistic approach (Daraio and Simar, 2003) we use this variable (market risks) as external-environmental variable, to investigate its effect on our data, i.e. if it is detrimental or favourable to the performance of mutual funds in the period under consideration.

In particular, we illustrate the power of this approach, analyzing the existence of economies of scale and the effect of market risks on US mutual funds for the year June 2001- May 2002.

After that we compare the robust efficiency indicators with nonparametric indicators (DEA, FDH) and traditional indicators of mutual funds' performance (i.e. Sharpe index and Jensen's α).[3]

[3]We use the Sharpe index and Jensen's α as provided by Morningstar, due to the lack of data on time series, necessary to estimate the performance measures in the traditional way. We notice that in computing

The paper is organised as follows. In the next section we outline the main ideas of our approach and describe the advantages of adopting a robust nonparametric approach based on the order$-m$ frontier introduced by Cazals, Florens and Simar (2002) and generalized by Daraio and Simar (2003). In this framework, a decomposition of mutual funds efficiency is proposed and its advantage for interpretation highlighted.

Section 3 describes the data and analyses the dynamics of transaction costs and funds' size. In section 4 we provide empirical evidence on returns to scale, analysis of slacks, economies of scale and the effect of market risks on the US mutual funds industry. Section 5 makes a comparison of traditional, nonparametric and robust measures of mutual fund performance. Section 6 concludes the paper.

2 A robust nonparametric approach

2.1 The methodology

In portfolio management the objective is to analyze the performance of investment portfolios. Risk (volatility or variance) and average return of a portfolio are analogous to inputs and outputs in model of production; the boundary of the attainable set of portfolios gives a benchmark relative to which the efficiency of a portfolio can be measured.

More generally, in an activity analysis framework (see e.g. Debreu, 1951; Shephard, 1970), the management of the mutual funds is characterized by a set of inputs $x \in I\!R_+^p$ used to produce a set of outputs $y \in I\!R_+^q$. The set of technically feasible combinations of (x, y) is defined as:

$$\Psi = \{(x, y) \in I\!R_+^{p+q} \mid x \text{ can produce } y\}. \tag{2.1}$$

In this setting, the Farrell measure of input-oriented[4] efficiency for a fund operating at the level (x, y) can be defined as:

$$\theta(x, y) = \inf\{\theta \mid (\theta x, y) \in \Psi\}, \tag{2.2}$$

where $\theta(x, y) \leq 1$ is the proportionate reduction of inputs a fund working at the level (x, y) should perform to achieve efficiency. The efficient frontier corresponds to those funds where $\theta(x, y) = 1$.

In efficiency analysis, the nonparametric approach is based on envelopment techniques, whose main estimators are Data Envelopment Analysis (DEA, see Charnes, Cooper and

α, Morningstar deducts the current return of risk-free asset from the total return of both the fund and the benchmark index. Thus the α figures shown by Morningstar may be lower than those usually computed.

[4]To save place, in this paper we only present the input oriented case. For the treatment of the output oriented case and further details, see Daraio and Simar (2003).

Rhodes, 1978) and Free Disposal Hull (FDH, see Deprins, Simar and Tulkens, 1984). These estimators rely on the idea that the attainable set is defined by the set of minimum volume containing all the observations. The DEA estimator relies on the free disposability[5] and on the convexity of the set Ψ, whereas the FDH relies only on the free disposability assumption.

The FDH estimator of Ψ, based on a sample of n observations (x_i, y_i), is the free disposal closure of the reference set $\{(x_i, y_i)|i = 1, ..., n\}$. It can be defined as:

$$\widehat{\Psi}_{FDH} = \Big\{(x, y) \in I\!R_+^{p+q} \mid y \leq y_i, x \geq x_i, i = 1, ..., n\Big\}. \tag{2.3}$$

The DEA estimator of Ψ, is the convex closure of $\widehat{\Psi}_{FDH}$:

$$\begin{aligned}\widehat{\Psi}_{DEA} \;=\; \Big\{(x, y) \in I\!R_+^{p+q} \mid y \leq \sum_{i=1}^{n} \gamma_i \, y_i \;;\; x \geq \sum_{i=1}^{n} \gamma_i \, x_i, \\ \text{for } (\gamma_1, ..., \gamma_n) \quad \text{s.t.} \sum_{i=1}^{n} \gamma_i = 1; \gamma_i \geq 0, i = 1, ..., n\Big\},\end{aligned} \tag{2.4}$$

The estimated FDH efficiency score of a fund (x, y) is given by:

$$\widehat{\theta}_{FDH} = \min\{\theta \mid (\theta x, y) \in \widehat{\Psi}_{FDH}\}. \tag{2.5}$$

Similarly, the estimated DEA efficiency score of a fund (x, y) is given by:

$$\widehat{\theta}_{DEA} = \min\{\theta \mid (\theta x, y) \in \widehat{\Psi}_{DEA}\}. \tag{2.6}$$

One of the main drawbacks of DEA/FDH nonparametric estimators is their sensibility to extreme values and outliers. In this context, Cazals, Florens and Simar (2002) propose a nonparametric estimator of the frontier, more robust to extreme values and outliers. It is based on the concept of the *expected minimum input function* of order$-m$. Extending these ideas to the full multivariate case, Daraio and Simar (2003) define the concept of *expected* order-m input efficiency score. This robust approach is based on a probabilistic formulation of the model. The production process is described by the joint probability measure of (X, Y) on $I\!R_+^p \times I\!R_+^q$.

In this formulation, the support of (X, Y) is the attainable set Ψ and the Farrell input efficiency[6] can be characterized, under the free disposability assumption, in terms of the conditional distribution function of X, given than $Y \geq y$, as follows:

$$\theta(x, y) = \inf\{\theta \mid F_X(\theta x \mid y) > 0\}, \tag{2.7}$$

[5] A set Ψ is free disposal if $(x, y) \in \Psi$ implies $(x', y') \in \Psi$ for any $x' \geq x$ and $y' \leq y$).

[6] Here again, we only describe the input-oriented case, for the output oriented case, for more technical details and illustrations, see Daraio and Simar (2003), Cazals, Florens and Simar (2002) and Daraio (2003).

where $F_X(x \mid y) = \text{Prob}(X \leq x \mid Y \geq y)$ (Note the non-standard condition $Y \geq y$ defining this conditional distribution, in place of the more common $Y = y$ used in regression models).

The FDH estimator can be written as:

$$\widehat{\theta}_n(x, y) = \inf\{\theta \mid \widehat{F}_{X,n}(\theta x \mid y) > 0\}, \tag{2.8}$$

where $\widehat{F}_{X,n}(x \mid y)$ is the natural nonparametric estimator of $F_X(x \mid y)$ given by the empirical conditional distribution function of X. Namely

$$\widehat{F}_{X,n}(x \mid y) = \frac{\sum_{i=1}^{n} \mathbb{1}(X_i \leq x, Y_i \geq y)}{\sum_{i=1}^{n} \mathbb{1}(Y_i \geq y)}, \tag{2.9}$$

where $\mathbb{1}(\cdot)$ is the indicator function.

The order-m input efficiency can be defined as in Daraio and Simar (2003). For a given level of outputs y in the interior of the support of Y, consider m i.i.d. random variables $X_i, i = 1, \ldots, m$ generated by the conditional p-variate distribution function $F_X(x \mid y)$ and define the set:

$$\Psi_m(y) = \{(x, y') \in \mathbb{R}_+^{p+q} \mid x \geq X_i, y' \geq y, i = 1, \ldots, m\}. \tag{2.10}$$

Then, for any x, we may define

$$\tilde{\theta}_m(x, y) = \inf\{\theta \mid (\theta x, y) \in \Psi_m(y)\}. \tag{2.11}$$

Note that $\tilde{\theta}_m(x, y)$ may be computed by the following formula:

$$\tilde{\theta}_m(x, y) = \min_{i=1,\ldots,m} \left\{ \max_{j=1,\ldots,p} \left(\frac{X_i^j}{x^j} \right) \right\}. \tag{2.12}$$

$\tilde{\theta}_m(x, y)$ is a random variable since the X_i are random variables generated by $F_X(x \mid y)$.

For any $x \in \mathbb{R}_+^p$, the (expected) order-m input efficiency measure is defined for all y in the interior of the support of Y as:

$$\theta_m(x, y) = E(\tilde{\theta}_m(x, y) \mid Y \geq y). \tag{2.13}$$

So, in place of looking for the lower boundary of the support of $F_X(x \mid y)$, as was typically the case for the full-frontier and for the efficiency score $\theta(x, y)$, the order-m efficiency score can be viewed as the expectation of the minimal input efficiency score of the unit (x, y), when compared to m units randomly drawn from the population of units producing more outputs than the level y. This is certainly a less extreme benchmark for the unit (x, y) than the "absolute" minimal achievable level of inputs: it is compared to a set of m peers (potential

competitors) producing more than its level y and we take as benchmark, the expectation of the minimal achievable input in place of the absolute minimal achievable input.

Then for any $x \in I\!R_+^p$, the expected minimum level of inputs of order-m is defined as $x_m^\partial(y) = \theta_m(x,y)\, x$ which can be compared with the full-frontier $x^\partial(y) = \theta(x,y)\, x$. Note that the order-$m$ efficiency can be computed as:

$$\begin{aligned} \theta_m(x,y) &= \int_0^\infty (1 - F_X(ux \mid y))^m du \\ &= \theta(x,y) + \int_{\theta(x,y)}^\infty (1 - F_X(ux \mid y))^m du, \end{aligned} \tag{2.14}$$

where we see that $\lim_{m\to\infty} \theta_m(x,y) = \theta(x,y)$.

A nonparametric estimator of $\theta_m(x,y)$ is then obtained by pluging the empirical version of $F_X(x \mid y)$ in (2.14). We have

$$\begin{aligned} \hat{\theta}_{m,n}(x,y) &= \int_0^\infty (1 - \widehat{F}_{X,n}(ux \mid y))^m du, \\ &= \hat{\theta}_n(x,y) + \int_{\hat{\theta}_n(x,y)}^\infty (1 - \widehat{F}_{X,n}(ux \mid y))^m du \end{aligned} \tag{2.15}$$

This leads to an estimator of the frontier, which for finite m, does not envelop all the observed data points and so, is less sensitive to extreme points or to outliers. As shown by (2.15), as m increases and for fixed n, $\hat{\theta}_{m,n}(x,y) \to \hat{\theta}_n(x,y)$.

Another useful analysis is the exploration of the factors which might influence or explain the performance differentials.

In the literature, mainly two approaches have been developed[7]: a *one-stage* approach, and a *two-stage* approach; but both are flawed by restrictive assumptions on the data generating process and/or on the role of these *external* factors on the production process.

Based on the probabilistic formulation presented above, Daraio and Simar (2003) propose a general full nonparametric approach that overcomes most drawbacks of previous approaches. The probabilistic formulation allows an easy introduction of additional information provided by external- environmental variables $Z \in \mathbb{R}^r$. Hence, the joint distribution on (X,Y) conditional on $Z = z$ defines the production process if the external factor $Z = z$.

The efficiency measure under the condition $Z = z$ can be defined as:

$$\theta(x,y \mid z) = \inf\{\theta \mid F_X(\theta x \mid y,z) > 0\}, \tag{2.16}$$

where $F_X(x \mid y,z) = \text{Prob}(X \le x \mid Y \ge y, Z = z)$. It has to be noted that the conditioning here is on $Y \ge y$ and $Z = z$ because there are no *a priori* assumptions on Z (if it acts on the production process as a favorable input or as an undesired output).

[7] See Daraio and Simar (2003) and the references cited there.

The *conditional FDH input efficiency measure* is the estimator of $\theta(x, y \mid z)$ defined as:

$$\widehat{\theta}_n(x, y \mid z) = \inf\{\theta \mid \widehat{F}_{X,n}(\theta x \mid y, z) > 0\}, \tag{2.17}$$

A nonparametric estimator of $F_X(x|y, z)$ requires some smoothing in z. We could chose for instance, the following kernel estimator:

$$\widehat{F}_{X,n}(x \mid y, z) = \frac{\sum_{i=1}^{n} \mathbb{1}(x_i \leq x, y_i \geq y) K((z - z_i)/h_n)}{\sum_{i=1}^{n} \mathbb{1}(y_i \geq y) K((z - z_i)/h_n)}, \tag{2.18}$$

where $K(\cdot)$ is the kernel and h_n is the bandwidth of appropriate size. Daraio and Simar (2003) propose a simple data-driven method for the choice of the bandwidth.[8]

In a similar way, *mutatis mutandis*, Daraio and Simar (2003) introduce also the conditional order$-m$ measures of efficiency with their nonparametric estimators. For the input-oriented case, the order$-m$ measure of efficiency is defined as:

$$\theta_m(x, y|z) = \int_0^\infty (1 - F_X(ux \mid y, z))^m du, \tag{2.19}$$

and its nonparametric estimator is obtained as follows:

$$\hat{\theta}_{m,n}(x, y|z) = \int_0^\infty (1 - \widehat{F}_{X,n}(ux \mid y, z))^m du. \tag{2.20}$$

Again, when $m \to \infty$, we recover the full frontier conditional measures, but for finite m, $\hat{\theta}_{m,n}(x, y|z)$ provides a more robust estimator of the frontier, robust to extremes or outliers.

The procedure for evaluating the effect of Z on the production process is based on the comparison of the conditional FDH measure $\hat{\theta}_n(x, y \mid z)$ with the unconditional FDH measure $\hat{\theta}_n(x, y)$. Accordingly, the same comparison is done for the robust order$-m$ efficiency measures. In particular, the ratios $Q^z = \hat{\theta}_n(x, y \mid z)/\hat{\theta}_n(x, y)$ (and their robust version Q^z_m) are useful to investigate on the effects of Z on performance: if $Q^z = 1$, then the conditional and unconditional efficiency measures are equal: this means that Z does not affect the performance of the analysed fund; if Q^z is much higher than 1, this means that the fund has been *highly* influenced by Z.

When Z is univariate, the scatterplot of these ratios against Z and its smoothed nonparametric regression line is also very helpful in describing the effect of these external-environmental variables on the production process. In the input oriented case (used in this paper), we have:

[8] In a *first step*, a bandwidth h is selected which optimizes the estimation of the density of Z based on the likelihood cross validation criterion, using a k-NN (Nearest Neighborhood) method (see e.g. Silverman, 1986). This allows to obtain bandwidths which are localized, insuring we have always the same number of observations Z_i in the local neighbor of the point of interest z when estimating the density of Z. In a *second step*, taking into account for the *dimensionality* of x and y, and the sparsity of points in larger dimensional spaces. the local bandwidths h_{Z_i} are expanded by a factor $1 + n^{-1/(p+q)}$, increasing with $(p+q)$ but decreasing with n. For more details, see Daraio (2003).

- An *increasing* smoothed nonparametric regression line denotes a Z that is *unfavorable* to the production process. In this case, in fact, the external- environmental variable acts like an "extra" undesired output to be produced asking for the use of more inputs in production activity, hence Z has a "negative" effect on the production process. In this case, the ratios Q^z (and of their robust version Q_m^z) will increase, on average, with Z.

- A *decreasing* smoothed nonparametric regression line indicates a Z that is *favorable* to the production process. In this case, the environmental variable plays the role of a "substitutive" input in the production process, giving the opportunity to "save" inputs in the activity of production; in this case, Z has a "positive" effect on the production process. By construction $\hat{\theta}_n(x, y \mid z)$ is higher or equal to $\hat{\theta}_n(x, y)$, but here it will be much larger for small values of Z (less substitutive inputs) than for large values of Z. Hence, here Q^z (and Q_m^z) will decrease, on average, when Z increases.

This approach, with its robust counterpart, offers a rigorous methodology to identify the factors that might influence mutual funds efficiency by measuring their *global* effect on the performance. It gives also the possibility to analyse the effect of these variables on each individual mutual fund. The performance of a fund (x, y), measured by the *Conditional Efficiency* index (CE), $\widehat{\theta}_n(x, y \mid z)$, can be decomposed in three main indicators:

1. An *unconditional* efficiency score (UE), $\widehat{\theta}(x, y)$ that represents the *internal or managerial efficiency*;

2. An *externality index* (EI) defined as $\widehat{E}\big(Q^z | Z = z\big)$. It is the *expected value* of the ratios Q^z given the value of z owned by the fund. It is given by the nonparametric fitted value of Q^z obtained by some appropriate nonparametric regression of Q^z on Z:

$$\widehat{E}\big(Q^z | Z = z\big) = \frac{\sum_{i=1}^{n} Q^{z_i} K((z - z_i)/h)}{\sum_{i=1}^{n} K((z - z_i)/h)}, \tag{2.21}$$

   where $K(\cdot)$ is the Kernel and h an appropriate bandwidth.

3. An *individual index* (II) defined as $Q^z / \widehat{E}\big(Q^z | Z = z\big)$. It compares the value of Q^z with the value we would expect for it, given its value of Z. It represents the fund expected *intensity* in catching the opportunities or threats by the environment (external factor).

In formulae:

$$CE(x, y) = UE(x, y) * EI(x, y) * II(x, y); \tag{2.22}$$

where, for the input oriented case:

$$\widehat{\theta}_n(x,y|z) = \widehat{\theta}_n(x,y) * \widehat{E}\big(Q^z|Z=z\big) * \frac{Q^z}{\widehat{E}\big(Q^z|Z=z\big)}. \tag{2.23}$$

Accordingly, the same decomposition can be done in the case of conditional order$-m$ efficiency score.

This decomposition is particularly useful in the interpretation of the funds performance. It offers the possibility for analyzing *individual* and *localized* effects of external factors and interpret them together with their global influence on the portfolio management. In section 4.2 we apply this decomposition to US mutual fund performance, considering the market risks as the external factor that conditions the efficiency of portfolio management.

2.2 Application to portfolio analysis

The methodology described above is promising for portfolio performance evaluation. The nonparametric approach in efficiency analysis, offers several advantages in this framework:

- absence of specification of the functional form for the input-output relationship;
- measurement of the efficiency with respect to the efficient frontier which measures the best performance that can be practically achieved whereas traditional methods estimate efficiency relative to the average performance, and hence the possibility to calculate an efficiency index for each mutual fund, and not only statistical averages;
- appropriate benchmark to be used for comparison: non requirement of any theoretical models (CAPM or APM) as benchmarks. Instead the measurement of how well a fund performs relative to the best set of funds within the declared objective category is done (*endogenous benchmarking*);
- possibility of addressing the problem of *endogeneity* of transaction costs, i.e. to include expense ratio, turnover and loads as well as the return simultaneously in the analysis (*multi-inputs multi-outputs indicators*);
- possibility to analyze the relative importance among the inputs: i.e. the marginal contribution of each input in affecting returns trough the *analysis of slacks*;
- no need to assume the normality of return distribution.

The nonparametric methodology we propose in this paper with its robust version (based on order-m efficiency) adds some new advantages.

- As the robust indicators are based on estimators that do not envelop all funds, they are more robust to outliers and extreme values which can strongly influence the nonparametric estimation of portfolio efficiency. The level of robustness can be set by means of m (tuning parameter).

- The robust nonparametric indicators avoid the *curse* of dimensionality, typical of nonparametric estimators. The order-m indicators are $\sqrt{n}$-consistent estimators whereas the DEA are only $n^{2/(p+q+1)}$-consistent estimators ($n^{1/(p+q)}$ for the FDH). This indicates for the DEA/FDH the necessity of increasing the number of observations when the dimension of the input-output space increases to achieve the same level of statistical precision;

- The order-m indicators allow to compare samples with different size, in an indirect way, avoiding the *sample size bias*[9], of which nonparametric indicators (DEA/FDH) suffer. In this case, m plays an important role. The benchmark, in fact, is not made against the most efficient units in the group, but against an appropriate measure drawn from a large number of random samples of size m within the group. In this way size-dependent effects are eliminated.

- The possibility of explaining efficiency, considering the *conditional influence* of *external factors* Z on the full frontier and on its robust counterpart.

- Using our approach, it is possible to decompose the performance of the fund (x, y), as measured by the *Conditional Efficiency* index in three main indicators: an indicator of the *internal* or managerial efficiency, an *externality index*, and finally, an *individual* index.

- Moreover, we can evaluate the effect of external/environmental Z variables on the performance of mutual funds in different economic *scenarios*, contemplating various numbers of *potential competitors*, using the parameter m in its dual meaning.

The possibility of constructing different "economic scenarios", of monitoring how the performance of mutual funds can be affected by progressively increasing the number of *potential* competitors, and of taking into account the influence of external/environmental variables, are particularly relevant for mutual funds analysis. The industry of mutual funds is an industry in expansion and subject to rapidly changes. Hence, to make choices more

[9]Zhang and Bartels (1998) show formally how DEA efficiency scores are affected by sample size. They demonstrate that comparing measures of structural (i.e. average) inefficiency between samples of different sizes leads to biased results.

accurate, 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, due to the expansion, acceleration and contraction of the industry.

3 Data description

We apply our methodology analyzing US mutual funds data. We use a cross-section data set, collected by the reputed Morningstar and updated at May 2002. Among this universe we selected 6 categories of Mutual Funds which were used in earlier studies (Grinblatt and Titman, 1989, 1993; Murthi et al., 1997).

The categories of funds analyzed are: Asset-Allocation (AA), Aggressive-Growth (AG), Balanced (B), Equity-Income (EI), Growth (G), Growth-Income (GI).

This classification of mutual funds follows the codes on portfolio *objectives* listed by Weisenberger Financial Services for all funds traded in the New York Stock Exchange market.

According to the description given by Morningstar, *Asset Allocation* funds are funds that seek to pursue their dual goals (income and capital appreciation) by using a flexible combination of stocks, bonds, and cash.

Aggressive Growth(AG) are funds that seek rapid growth of capital and that may invest in emerging market growth companies without specifying a market capitalization range.

Balanced(B) funds seek both income and capital appreciation by investing in a generally fixed combination of stocks and bonds. These funds generally hold a minimum of 25% of their assets in fixed-income securities at all times.

Equity-Income (EI) are funds that are expected to pursue current income by investing at least 65% of assets in dividend-paying equity securities.

Growth (G) funds pursue capital appreciation by investing primarily in equity securities.

Growth-Income (GI) funds seek growth of capital and current income as near-equal objectives. Investments are typically selected for both appreciation potential and dividend paying ability.

In the first step of the analysis we selected for each category the total number of funds available. In a second step, we dropped the funds with missing values. Table 1 shows the number of observations before missing value exclusion (first line) and after missing values exclusion (second line).

Table 1: *Data considered in the analysis.*

Step	*Asset Allocation*	*Aggressive Growth*	*Balanced*	*Equity Income*	*Growth*	*Growth Income*
1	471	247	503	254	3,359	1,017
2	284	144	405	197	1,888	763
3	239	129	342	176	1,653	627

In the further step we applied the Simar's (2003) approach to detect outliers. However, we are aware of the importance of outliers for the analysis and we left further investigations to future work. The last line of Table 1 reports the number of funds for each category finally used to make the computations.

3.1 Description of variables

The selection of variables has been done by taking the same variables chosen in earlier studies (Murthi et al., 1997; Sengupta, 2000) applying the nonparametric approach. This choice is aimed at comparing the results of our approach with those of previous methodologies. In the following we describe the variables used in the analysis[10].

The traditional output in this framework is the *Total Return* (the annual return at the 05-31-2002, expressed in percentage terms). The traditional input is the *Risk*, the standard deviation of Return. Other inputs include the *Transaction costs*, made by Expense Ratio, Loads and Turnover Ratio.

- *Expense Ratio* is the percentage of fund assets paid for operating expenses and management fees, including 12b-1 fees (the annual charge deducted from fund assets to pay for distribution and marketing costs), administrative fees, and all other asset-based costs incurred by the fund except brokerage costs. Sales charges are not included in the expense ratio.

- *Loads* have been obtained by summing the Front-End Load and the Deferred Load of each fund. The Front-End Load is the initial sales charge which consists in a one-time deduction from an investment made into the fund. The sales charge serves as a commission for the broker who sold the fund. The Deferred Loads are also known as back-end sales charges and are imposed when investors redeem shares.

- *Turnover ratio* is a measure of the fund's trading activity which is computed by taking the lesser of purchases or sales and dividing by average monthly net assets. It gives

[10]The definition of these variables follows strictly the way Morningstar calculate them. By the way, we notice that Morningstar data are widespread used in the studies on US mutual funds.

an indication of trading activity: funds with higher turnover (implying more trading activity) incur greater brokerage fees for affecting the trades. It is also an indication of management strategy: a low turnover figure would indicate a "buy-and-hold strategy"; high turnover would indicate an investment strategy involving "considerable buying and selling" of securities.

- *Market risks* reflects the percentage of a fund's movements that can be explained by movements in its benchmark index.[11]

Sengupta (2000) uses market risks of mutual funds as an input in his analysis, underlying that the effect of market risks is favorable for mutual fund performance. In our analysis we use market risks as external-environmental variable, to investigate its effects on our data: i.e. if it is detrimental or favorable to the performance of mutual funds in the period under consideration.

Summing up, in section 4 we apply an input oriented framework, considering as output the Total return, and as inputs: risk, expense ratio, and turnover ratio[12]. We consider also fund size as external-environmental variable to analyse the existence of economies of scale in US mutual funds industry.

3.2 The dynamic of transaction costs and funds size

The connection between transaction costs and portfolio performance has been highlighted in the literature (see Grossman and Stiglitz, 1980; Ippolito 1989, Elton, Gruber, Das and Hlavka, 1993): if the collection and usage of information is costly, informed investors should obtain higher returns relative to the uninformed investors; *i.e.*, if portfolio managers have superior ability, they may be able to expropriate the economic rents by charging higher fees. Thus, one would expect a relation between the transaction costs and the return.

In Table 2, we report the mean characteristics of transaction costs and fund size for the six categories of mutual funds, fund size is measured by the NAV (Net Asset Value) in million of US dollars. From a comparison with other studies[13] (Murthi et al., 1997; Grinblatt and Titman, 1989; Ippolito, 1989) we find that the Net Asset Value (NAV) is lower than in 1993 as it passed from an average value of 544.98 to 471.56. After a large expansion of the mutual fund industry (Grinblatt and Titman, 1989, report an average NAV of the funds for

[11]Morningstar compares all equity funds to the S&P 500 index and all fixed-income funds to the Lehman Brothers Aggregate Bond Index.

[12]We do not consider the loads in our analysis because they remain stable across categories and a lot of funds show a level of loads equal to zero.

[13]We are aware that these comparisons could be affected by *survivor bias*; they are intended to trace just a "rough" picture of average tendencies.

about 113.8 million of US dollars in 1974; this value more than quadrupled in the following 20 years to a value of 544.98 million of US dollars in 1993 -Murthi et al., 1997 study-) it seems that the industry is facing a contraction cycle. We remark a huge standard deviation of Fund size across category.

On the contrary, turnover has drastically increased to 109.51 % in 2002 from 86.8% in 1993 (Murthi et al., 1997). In 1982-84 the average turnover was of 63.4% (Ippolito, 1989). The highest turnover, on average, is (as expected) in the Aggressive Growth category, that has also the highest risk, on average (See Table 3). Anyway, also the Growth, and (surprisingly) Asset Allocation and Balanced category, show high average turnover, indicating that there has been a lot of trading activity in these funds. Expense Ratio, has been very stable at around 1.4 in 2002, which is similar to the expense ratio for previous periods. The average load charged is of 2.3% and appears to be constant across categories.

Table 3 compares the average values of risk, return and market risks among categories. Note that most returns where negative in the considered period, hence we shift them to get all positive returns by adding 100. This transformation does not affect our input oriented analysis.

Table 2: *Mean characteristics of Transaction Costs and Fund Size*

		Asset Allocation	*Aggressive Growth*	*Balanced*	*Equity Income*	*Growth*	*Growth Income*
Transaction Cost							
Expense Ratio							
1.43	*Mean*	1.381	1.681	1.359	1.342	1.503	1.297
	St. Dev.	0.544	1.281	0.557	0.474	0.569	0.546
	Skewness	0.014	8.260	0.394	0.250	1.062	0.175
	Curtosis	-0.499	83.089	0.098	-0.615	6.064	-0.545
Loads							
2.34	*Mean*	2.237	2.399	2.388	2.369	2.356	2.272
	St. Dev.	2.452	2.428	2.406	2.420	2.450	2.445
	Skewness	0.379	0.260	0.234	0.254	0.310	0.371
	Curtosis	-1.715	-1.751	-1.784	-1.777	-1.692	-1.670
Turnover							
109.51	*Mean*	104.456	155.194	114.497	73.631	117.852	91.434
	St. Dev.	116.400	98.778	98.790	45.102	133.389	192.822
	Skewness	4.833	1.743	2.801	1.079	5.743	12.650
	Curtosis	31.538	5.167	11.289	0.996	52.329	183.647
Fund Size							
471.56	*Mean*	381.654	446.784	266.859	408.501	535.148	790.384
	St. Dev.	1710.784	1163.777	842.975	1059.919	2643.502	3863.678
	Skewness	9.071	4.540	7.447	6.434	16.438	13.477
	Curtosis	95.971	24.402	68.176	53.988	368.252	214.152
N. Obs.		239	129	342	176	1,653	627

Note: Fund size is measured by the NAV (Net Asset Value) in million of US dollars.

***

Table 3: *Mean characteristics of Risk, Return and Market Risks*

		Asset Allocation	*Aggressive Growth*	*Balanced*	*Equity Income*	*Growth*	*Growth Income*
Risk							
	Mean	12.831	34.978	10.493	13.806	23.908	16.487
	St. Dev.	8.599	8.850	3.456	2.553	9.792	3.868
	Skewness	2.719	0.796	4.996	-0.527	2.581	4.505
	Curtosis	8.701	4.732	53.114	3.460	14.845	47.267
Return							
	Mean	93.779	81.833	94.235	93.053	85.505	88.991
	St. Dev.	8.287	9.803	4.445	5.897	11.122	7.370
	Skewness	-0.968	-0.473	-0.136	0.032	0.589	0.519
	Curtosis	2.999	2.074	0.152	-0.069	1.385	1.529
Market risks							
	Mean	65.381	47.225	75.196	59.188	65.961	75.234
	St. Dev.	26.314	15.680	17.704	18.041	21.066	20.598
	Skewness	-0.544	0.565	-0.979	0.224	-0.414	-0.612
	Curtosis	-0.817	1.913	0.292	-0.500	-0.775	-0.547
N. Obs.		239	129	342	176	1653	627

Note: Most returns where negative in the considered period, hence we shift them to get all positive returns by adding 100. This transformation does not affect our input oriented analysis.

***

4 Empirical results

4.1 Returns to scale and analysis of slacks

Due to the nature of mutual funds, the structure of Constant Returns to Scale (CRS) may be assumed. This is what has been done in the literature (see Murthi et al., 1997). Anyway, we have applied the Simar and Wilson (2002) bootstrap-based procedure to test the nature of returns to scale of the different US mutual funds categories.

For the following mutual funds categories we cannot reject the CRS assumption: Aggressive Growth, Equity Income and Growth; instead the remaining categories: Balanced, Asset Allocation and Growth Income present Variable Returns to Scale (VRS).[14] Therefore, we compute the DEA efficiency score (one output -return- three inputs -risk, turnover, and expense ratio-) of each mutual fund with respect to its category by objective, according to its convenient returns to scale assumption.

[14]For these latter US mutual funds categories we performed a test on Non Increasing Returns to Scale (NIRS), concluding the existence of VRS.

After that, we examine the slacks of the input variables to identify the source of mutual funds inefficiency. This is what has been done in previous works (see e.g. Murthi et al, 1997). Table 4 shows the mean of the absolute slacks (first line), the mean value of the input (second line) and the relative mean slacks (third line); the latter are calculated by dividing the absolute mean slack of the input for the mean value of the input considered.

Using the relative slacks, we can compare the *marginal impact* of inputs on portfolio return across different categories of funds. The slacks measure whether the portfolio managers expend resources inefficiently. In opposition to previous results (Murthi et al., 1997), we fail to confirm that the risk (measured by standard deviation of return) has virtually no slacks throughout all investment categories, i.e. we *fail* to confirm that all mutual funds categories are on average, *mean-variance* efficient. We found that 3 out of 6 categories of US mutual funds analysed are *mean-variance* efficient: Balanced, Growth and Growth Income.

Table 4 shows that Aggressive Growth funds have the highest relative risk slacks, meaning that the marginal impact of risk on the portfolio return is the highest among US mutual funds categories. The Growth Income funds are the most inefficient in turnover and have the smallest slacks in risks.

The same data on US mutual funds have been analysed in Daraio (2003), where a bootstrap exercise has been carried out and the US mutual funds efficiency distributions have been explored.

Table 4: *Absolute and relative slacks.*

		Asset Allocation	*Aggressive Growth*	*Balanced*	*Equity Income*	*Growth*	*Growth Income*
Risk	MEAN slacks	0.50	1.91	0.02	0.08	0.00	0.01
	ST DEV	2.51	4.02	0.24	0.40	0.01	0.21
	MIN	0.00	0.00	0.00	0.00	0.00	0.00
	MAX	24.75	32.58	3.23	3.83	0.37	4.70
	Mean Input	12.83	34.98	10.49	13.81	23.91	16.49
	Relative Slacks	0.04	0.05	0.00	0.01	0.00	0.00
Turnover	MEAN slacks	6.80	10.78	1.68	1.01	11.49	29.53
	ST DEV	38.10	46.51	14.44	3.25	38.34	81.83
	MIN	0.00	0.00	0.00	0.00	0.00	0.00
	MAX	503.23	402.86	214.40	21.34	609.52	1346.44
	Mean Input	104.46	155.19	114.50	73.63	117.85	91.43
	Relative Slacks	0.07	0.07	0.01	0.01	0.10	0.32
Expense Ratio	MEAN slacks	0.02	0.06	0.03	0.06	0.00	0.07
	ST DEV	0.10	0.42	0.11	0.11	0.10	0.20
	MIN	0.00	0.00	0.00	0.00	0.00	0.00
	MAX	0.79	4.45	1.17	0.62	2.43	1.08
	Mean Input	1.38	1.68	1.36	1.34	1.50	1.30
	Relative Slacks	0.01	0.04	0.02	0.04	0.00	0.06
N. Obs.		239	129	342	176	1653	627

4.2 Robust measures of performance

The methodology described in section 2 provides several information that can be useful to interpret how the management of assets portfolio has been carried out in the period under consideration. In this section we report some results in order to show their usefulness.[15]

Table 5 shows the averages of FDH efficiency scores (i.e. $\widehat{\theta}_n(x,y)$, third column) and of the conditional FDH efficiency scores (i.e. $\widehat{\theta}_n(x,y|z)$, sixth column) for US mutual funds categories by objective. N is the average number of observations dominating (x,y) and N_z is the average number of dominating points given the value of Z. h is the average of the selected local bandwidths. Z is the average value of market risks, Q^z is the geometric mean of the individuals ratios Q^{z_i}; and II is the geometric mean of the Individual index indicators II_i.

[15]We do not report all of them to save space.

Table 5: *FDH, Conditional FDH and other indicators - average values for US mutual funds category by objective.*

Category	N	$\widehat{\theta}(x,y)$	h	N_z	$\widehat{\theta}(x,y\|z)$	Z	Q^z	II
AA	29.042	0.53185	0.035905	2.2636	0.82126	0.65374	1.641	0.935
AG	9.2093	0.60827	0.082307	1.814	0.88249	0.47213	1.498	0.972
B	35.661	0.65824	0.037568	3.4386	0.85749	0.75181	1.332	0.976
EI	16.489	0.61383	0.051212	2.2727	0.87685	0.592	1.499	0.934
G	189.12	0.57205	0.029438	12.675	0.74346	0.65966	1.320	0.983
GI	63.423	0.60857	0.023881	3.8708	0.87344	0.75236	1.458	0.980

Note that for the indicators Q^z and II we report the geometric mean.

***

Table 6: *Order$-m$, Conditional Order$-m$ and other indicators - average values for US mutual funds category by objective.*

Category	$\widehat{\theta}_m(x,y)$	$\widehat{\theta}_m(x,y\|z)$	Z	Q^z_m	II_m	m
AA	0.67057	0.8268	0.65374	1.376	0.946	75
AG	0.687	0.89724	0.47213	1.335	0.981	75
B	0.73542	0.87036	0.75181	1.205	0.980	75
EI	0.76168	0.88489	0.592	1.245	0.943	50
G	0.76778	0.79322	0.65966	1.076	0.984	50
GI	0.8328	0.8853	0.75236	1.111	0.974	45

Note that for the indicators Q^z_m and II_m we report the geometric mean.

***

In Table 6 the average values of the robust unconditional ($\widehat{\theta}_m(x,y)$, second column) and conditional to Z ($\widehat{\theta}_m(x,y|z)$, third column) measures for mutual funds categories are reported. Q^z_m is the geometric mean of the individual ratios $Q^{z_i}_m$ ratios and II_m is the geometric mean of the individual ratios $II_{i,m}$.

For the order$-m$ measures, we choose the value of m for each mutual funds category by objective robust at 10%, *i.e.* the value of m that leaves below the order$-m$ frontier a percentage of observations of around 10%. The chosen value of m are reported in the last column of Table 6.

To compute the efficiency measures (full and of order$-m$) conditional to market risks we use a triangular kernel (we obtain very similar results by using other kernels with compact support). The optimal value for k (fixing the number of nearest neighbors useful for estimating the conditional measures) is 15 for Asset Allocation, 21 for Aggressive Growth, 32 for Balanced, 20 for Equity Income, 99 for Growth and 31 for Growth Income.

Analysing Table 5 we see the global effect of the market risks factor Z on the full efficiency measures. For instance, for the Aggressive Growth (AG) category, we observe an increase of

0.274, from 0.608 to 0.882.

In Table 6 we can find the average order$-m$ efficiency scores (second column), and the mean effect of market risks on it (third column). For the AG category, we have a mean effect of 0.210, going from 0.687 to 0.897. So, it appears that the effect is more important for the full FDH frontier (0.274) than for the order$-m$ frontier (0.210), as expected, since the former measures are more sensitive to extremes values.

By comparing the *robust* mean effects of market risk across mutual funds categories (see Table 6), we find that the Aggressive Growth is the most influenced category, while the Growth and Growth Income are the categories less influenced by market risks in the period 2001-2002. This finding is confirmed also by the inspection of the bottom panel of Figures 2, 5 and 6, where the smoothed nonparametric regressions of $Q_m^{z_i}$ on Z are reported.

Table 7 shows, as an example, the analytical results available for some selected individual mutual funds, regarding the full frontier efficiency scores. It presents the category of the fund, the fund number, the number of observations dominating the fund (N), its FDH score, the selected local bandwidth h, the number of dominating points given the value of Z (N_z), the conditional FDH score, the value of market risk (Z), the value of Q^z, the Externality Index (EI), the Individual Index (II), and finally the R value, given by the ratio of Q^z on the geometric mean of all Q^{z_i}.

Table 7: *Example of empirical results for single US Mutual Funds across categories. FDH, Conditional FDH and other indicators.*

Cat.	Fund	N	$\widehat{\theta}(x,y)$	h	N_z	$\widehat{\theta}(x,y\|z)$	Z	Q^z	EI	II	R
AA	61	0	1.00	0.04	0	1.00	0.48	1.00	1.45	0.69	0.61
AG	21	5	0.53	0.03	1	1.00	0.40	1.89	1.54	1.23	1.26
B	258	9	0.74	0.02	0	1.00	0.83	1.35	1.32	1.02	1.01
EI	24	66	0.31	0.01	6	0.89	0.44	2.88	2.00	1.44	1.92
G	1224	45	0.68	0.06	4	0.97	0.99	1.43	1.32	1.08	1.09
GI	169	119	0.56	0.03	11	0.56	0.78	1.00	1.11	0.90	0.69

Table 8 gives the results for the same selected funds, for to order$-m$ measures. It reports the category of the fund, the fund's order-m score, its value of Z, its value of Q_m^z, its Externality Index of order$-m$ (EI_m), its Individual Index or order$-m$ (II_m), and finally its R_m value, given by the ratio of Q_m^z on the geometric mean of all Q_m^z.

Let us give an example of interpretation of the unconditional order$-m$ efficiency measure for the fund AG no. 21, $\widehat{\theta}_{m,n}(x,y) = 0.74$. A value of 0.74 indicates that the mutual fund AG no. 21 has a level of risk-transaction costs higher than 26% with respect to the expected minimum risk-transaction costs level of m other funds drawn from the population of mutual funds of the same objective category having a level of return $\geq y$.

The economic meaning of conditional to Z order$-m$ index of the fund, $\widehat{\theta}_{m,n}^z(x,y)$, comes

straightforward. The only difference is that the m mutual funds are drawn from the population of AG funds staying in the neighborhood of z, the market risks of AG no. 21 fund having a return at least equal to y.

Table 8: *Example of numerical results for single US Mutual Funds across categories. Order−m, Conditional Order−m and other indicators.*

Cat	Fund	$\widehat{\theta}_m(x,y)$	$\widehat{\theta}_m(x,y\|z)$	Z	Q_m^z	EI_m	II_m	R_m
AA	61	1.01	1.00	0.48	0.99	1.21	0.82	0.72
AG	21	0.74	1.00	0.40	1.35	1.37	0.99	1.01
B	258	0.80	1.00	0.83	1.26	1.17	1.08	1.04
EI	24	0.52	0.89	0.44	1.70	1.59	1.07	1.37
G	1224	0.83	0.98	0.99	1.19	1.06	1.12	1.10
GI	169	0.74	0.64	0.78	0.86	0.87	0.99	0.78

4.3 Evaluating the influence of market risks

The industry of mutual funds is an industry subjected to rapidly changes. It is subject to expansion, acceleration and contraction cycles that affect the performance of the asset manager firms.

It is interesting to see how market risks affected the performance of US mutual funds industry in the year 2001-2002. In order to do that we apply the methodology described above and we analyze the effect of market risks on the performance of mutual funds categories by objective. The results are reported in Figures 1 to 6.

By inspecting the bottom panel of Figures 1 to 6, it appears that most US funds show a *discontinuous* pattern with respect to market risks.

This result can be related to the confusion and negative performance of the stock exchanges all over the world, after the September, 11th 2001 terrorist attack to US.

Only the Aggressive Growth US mutual fund category presents a globally decreasing movement of the smoothed line (see Figure 2), confirming that the influence of market risks on its performance has been positive. With our data, it appears that market risks acts like an input in the AG funds production process. This evidence is consistent with the assumption made by Sengupta (2000), but here, it came out as a result of the analysis and not from some *a priori* postulate.

4.4 Economies of scale

In a production process, economies of scale are present if average costs per unit of output decline as the volume of output increases. The usual source of scale economies is the spreading of the firms fixed costs over a larger volume of output.

In the case of mutual fund companies, when the scale of activity expands, a less than proportional increase in costs may be recorded both in the area of portfolio management (information technology and security turnover) and in shareholder servicing (record, keeping and distribution).

In order to assess the effects of scale on US mutual funds performance we applied the same methodology, using mutual fund size as external-environmental variable. We use as proxy of size the Net Assets Value (NAV) of mutual funds, as done in previous studies (see e.g. Murthi et al, 1997, which found that the efficiency of US mutual funds in 1993 was not related to the funds size[16]).

Figures 7 to 12 show the results. The smoothed nonparametric regression of Q^z (Q_m^z) on Z (in this case the NAV) is offered in the top (bottom) panel of each figure. It seems that there are no economies of scale for the following US mutual funds categories: Balanced (see Figure 9), Growth (see Figure 11) and Growth Income (see Figure 12), in the considered period. The other categories of funds experimented some effects of scale among smaller firms, while they do not show any influence of size starting from a NAV of 1000 (million of US dollars).

The evidence found in this paper suggests that probably for most US mutual funds analysed (except some movements for smaller firms of AA, AG and EI categories), the asset growth is accompained by a large increase either in the variety of securities in portfolio or in the number of accounts. Hence, operational economies of scale have been offset by the complexity induced by learning to deal with a larger number of securities and/or customers.

5 A comparison of performance measures

Several measures of productivity can be used in a comprehensive way in order to make a rigorous comparative productivity analysis. It is important to emphasize that each productivity indicator (i.e. simple traditional measures, DEA efficiency score, FDH and order$-m$ measure) must be correctly interpreted, taking into account its economic meaning.

Table 9 shows the results of a comparison among traditional, nonparametric and robust nonparametric indices of mutual funds performance evaluation.

[16]For a review of the international evidence see Amel, Barnes, Panetta and Salleo (2002).

Table 9: *A comparison of robust nonparametric index, nonparametric indices and other traditional indices for US Mutual Fund category by objective.*

Performance Indicators	*Descriptive Statistics*	*Asset Allocation*	*Aggressive Growth*	*Balanced*	*Equity Income*	*Growth*	*Growth Income*
Robust and Nonparametric							
Order$-m$							
	Mean	0.671	0.687	0.735	0.762	0.768	0.833
	St. Dev.	0.725	0.197	0.208	0.729	0.448	0.454
	Skewness	9.275	1.045	2.916	11.264	9.306	7.728
	Curtosis	98.614	1.166	24.530	140.755	115.420	87.210
	value of m	75	75	75	50	50	45
FDH							
	Mean	0.532	0.608	0.658	0.614	0.572	0.609
	St. Dev.	0.247	0.207	0.185	0.255	0.181	0.174
	Skewness	0.738	0.838	0.266	0.234	0.346	0.978
	Curtosis	-0.745	-0.507	-0.484	-1.300	-0.062	0.380
DEA							
	Mean	0.450	0.468	0.536	0.461	0.414	0.530
	St. Dev.	0.223	0.145	0.171	0.198	0.149	0.150
	Skewness	1.174	1.548	1.179	1.143	0.923	1.478
	Curtosis	0.445	3.465	1.436	0.707	2.043	2.563
Traditional							
Jensen α							
	Mean	1.800	10.667	0.769	1.883	5.629	1.560
	St. Dev.	5.219	8.677	3.764	3.752	8.782	5.235
	Skewness	2.116	0.733	2.249	0.580	0.627	1.484
	Curtosis	6.519	0.419	15.958	0.159	0.738	3.723
Sharpe index							
	Mean	-0.358	-0.241	-0.500	-0.374	-0.355	-0.517
	St. Dev.	0.443	0.316	0.400	0.342	0.466	0.399
	Skewness	0.814	0.422	0.346	0.424	0.328	0.888
	Curtosis	0.400	-0.066	1.305	0.017	0.114	1.019
N. Obs.		239	129	342	176	1653	627

By inspecting Table 9 several interesting features of performance measures emerge. On average, we find a sort of "order" among the nonparametric indicators that is explicable coming back to their economic meaning. The DEA efficiency score is a performance indicator obtained by comparing each mutual fund with the best performers of its objective group; i.e. the benchmarking is, in a certain sense, *severe* for each fund and *sensitive* to extremely performing funds. The same nature of DEA is shared by the FDH estimator. Nevertheless, the difference between DEA and FDH is important: the DEA rests on the hypothesis of convexity of the attainable set, whereas the FDH does not. If the convexity hypothesis is questionable, the DEA might be a wrong measure (i.e. statistically inconsistent).

Different with respect to DEA and FDH is the robust nonparametric order$-m$ efficiency score. It reflects a *more realistic* benchmark for the fact that in the construction of it we

make a comparison of the performance of each fund (in terms of its use of inputs, i.e. risk and transaction costs) not with respect to 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, *i.e.* a return, equal or higher than that of the analyzed fund.

The values of traditional indicators Alpha and Sharpe ratio are different with respect to the nonparametric indicators. This is due to the different construction of the indicators and mainly to the fact that they consider only the risk-return dimension of mutual funds performance and make reference to external benchmark portfolio; on the contrary the nonparametric indicators include transaction costs into the performance comparison and rely on an *endogenous* benchmark. For that reason, the nonparametric indicators give a more complete view of funds performance and their ranking and can be useful to make in-depth mutual funds comparative analysis.

To highlight the relation existing among several productivity indicators, we compared the simple traditional indicators (i.e. Alpha and Sharpe ratio) with the nonparametric (i.e. DEA scores) and robust nonparametric indicators (order$-m$ measures), using the Pearson, Spearman and Kendall's tau-b measures of correlation.

From the correlation analysis[17] two general findings emerge:

- Sharpe ratio and Jensen's α indicators are positively correlated to each other across all mutual fund categories. This empirically confirms the result by Shukla and Trzcinca (1992) showing that the measures by Sharpe (1966) and Jensen (1968) are highly correlated in terms of funds ranking.

- While indicators based on nonparametric and robust approaches (DEA, FDH, Order$-m$) are highly positively correlated, they are weakly correlated with the traditional indicators (Sharpe ratio and Jensen's α). This is not a surprise due to the different perspectives taken by the two approaches.

These findings confirm the importance of using several indicators and the *superiority* of nonparametric and robust nonparametric measures in the comparative analysis of mutual funds management, since they are based on more information and use an endogenous benchmark.

6 Conclusions

In this paper, the general methodology proposed by Daraio and Simar (2003), extending previous results from Cazals, Florens and Simar (2002), is suggested to evaluate the per-

[17]The results are not reproduced to save place. All the results are available upon request.

formance of mutual funds and analyse the possible factors influencing it. The methodology includes techniques which are robust to extreme values and outliers. We study US mutual funds industry, by comparing the performance of six categories of funds namely according to their objective as Asset Allocation, Aggressive Growth, Balanced, Equity Income, Growth and Growth Income.

After a description of the dynamics of transaction costs and fund size, we empirically investigate on the slacks, on the existence of economies of scale and on the effects of market risks in US mutual funds industry.

From the analysis of slacks, previous results (see Murthi et al. 1997) on *mean-variance* efficiency of US mutual funds, on average, have not been confirmed for the categories: Asset Allocation, Aggressive Growth and Equity income.

For the analysed period, most US mutual funds did not exploit the economies of scale deriving from the portfolio management and shareholder services on a larger number of securities/customers. It seems that US funds experimented the complexity induced by learning to deal with an increased number of securities/customers that offset the operational economies of scale.

In particular, we show the power of this approach, by describing the effect of market risks on US mutual funds. Spanning the terrorist attach of the 11th September 2001 against US, the analysed year offers a first picture of that terrible event's effects on the return and performance of US mutual funds, and more generally on the international financial markets. The uncertainty and confusion following that unpleasant incident, roughly illustrated by the scatterplots of our analysis, have contributed to the profound collapse of financial markets determined by the crises of confidence of investors that led the speculative bubble to explode in the financial markets of the most advanced countries.

The robust approach proposed here can be a *useful tool* for practitioners and authorities:

- It can be used by *financial analysts* to monitor the performance of mutual funds industry either at an aggregate level (for a comparison across categories) or at a sectorial level (analyzing single category).

- *Financial investors* could get more information on their investments by the robust indicators, and could use the potential competitors scenario to compare mutual fund holding politics.

- The *regulator authority* of financial markets could be interested in monitoring the performance of mutual fund industry along with its efficiency, and the influence of market risks on the structure and performance of the business.

This methodology is based on robust nonparametric indices with clear interpretability and economic meaning. We have shown that the robust nonparametric approach can be useful to complete the analysis carried out through the nonparametric approach. The rapidity of its computational algorithms as well as the easy interpretability of results make this approach a promising tool to apply in Portfolio Analysis.

The rapidly changes in portfolio dynamics ask for new and more *rigorous* quantitative methods to accumulate empirical evidence and derive sound policy implications. The robust nonparametric approach we propose in this paper could very well accomplish this task.

References

[1] Admati, A. and Ross, S. (1985), "Measuring investment performance in a rational expectations equilibrium model", *Journal of Business*, 58, 1-26.

[2] Amel, D., Barnes C., Panetta F. and Salleo C. (2002), "Consolidation and efficiency in the financial sector:a review of the international evidence", Tema di discussione, Banca d'Italia n. 464.

[3] Annaert J., van den Broeck, J. and Vennet R.V. (2003), "Determinants of mutual funds underperformance: a bayesian stochastic frontier approach", *European Journal of Operational Research*, 151, pp. 617632.

[4] Briec, W., Kerstens, K., and Lesourd J.B., (2001), "Markowitz Portfolio Selection, Performance Gauging and Duality: Generalising Luenberger's Shortage Function", GREQAM, Document de Travail no.01A25.

[5] Briec, W., and Lesourd J.B., (2000), "The Efficiency of Investment Fund Management: An Applied Stochastic Frontier Model, in: C.L. Dunis (ed.), *Advances in Quantitative Asset Management*, Kluwer, Boston, 41-59.

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

[7] Cesari, R. Panetta, F. (2002), "The performance of the Italian equity funds", *Journal of Banking and Finance*, 26, 99-126.

[8] Charnes, A., Cooper, W.W., and Rhodes, E. (1978), "Measuring the Efficiency of Decision Making Units", *European Journal of Operational Research*, 2, 429-444.

[9] Choi, Y. (1995), "The sensitivity in tests of the efficiency of a portfolio and portfolio performance measurement", *The Quarterly Review of Economics and Finance*, 35, 187-206.

[10] Cooper, W.W., Seiford L.M., and Zhu J. (eds.) (2004), *Handbook on Data Envelopment Analysis*, Kluwer Academic Publishers, Boston.

[11] Daraio, C. (2003), *Comparative Efficiency and Productivity Analysis based on Nonparametric and Robust Nonparametric Methods. Methodology and Applications*, Ph.D. Dissertation, Sant'Anna School of Advanced Studies, Pisa.

[12] Daraio C. and Simar, L. (2003), "Introducing environmental variables in nonparametric frontier estimation: a probabilistic approach", Discussion Paper no. 0313, Institut de Statistique, UCL, Belgium, *forthcoming* on *The Journal of Productivity Analysis*.

[13] Debreu, G. (1951), "The Coefficient of Resource Utilization", *Econometrica*, 19, 273-292.

[14] Deprins, D., Simar L. and Tulkens H. (1984), "Measuring labor-efficiency in post offices", in Marchand, M., Pestieau, P. and Tulkens, H. (eds.) *The Performance of public enterprises - Concepts and Measurement*, Amsterdam, North-Holland, 243-267.

[15] Elton, E., Gruber,M., Das S. and Hlavka M. (1993), "Efficiency with costly information: A reinterpretation of evidence for managed portfolios", *Review of Financial Studies*, 6, 1-22.

[16] Farrell, M.J. (1957), "The measurement of the Productive Efficiency", *Journal of the Royal Statistical Society*, Series A, CXX, Part 3, 253-290.

[17] Grinblatt, M. and Titman S. (1989), "Mutual Funds Performance: An Analysis On Quarterly Portfolio Holdings", *Journal of Business* 62, 393-416.

[18] Grinblatt, M. and Titman S. (1995), "Performance Evaluation", in R.Jarrow, V. Maksimovic, W.T. Ziemba (eds.), *Handbooks in OR & MS: Finance*, vol 9, Amsterdam, Elsevier, 581-609.

[19] Grossman, S. and Stiglitz J. (1980), "On the impossibility of Informationally Efficient Markets", *American Economic Review*, Vol.70, No.3, 393-408.

[20] Indro, D.C., Jiang C.X., Hu M. Y., and Lee W.Y. (1999), "Mutual fund performance: does fund size matter?", *Financial Analysts Journal* (3), 74-87.

[21] Ippolito, R.A. (1989), "Efficiency with Costly information: a study of Mutual Fund Performance, 1965-1984", *Quarterly Journal of Economics*, Vol. 104, No.1, 1-23.

[22] Ippolito, R.A. (1993), "On Studies of Mutual Fund Performance, 1962-1991", *Financial Analysts Journal*, 49 (1), 42-50.

[23] Jensen, M. (1968), "The Performance of Mutual Funds in the Period 1945-1964", *Journal of Finance*, 23, 389-416.

[24] Kumbhakar S. C., Lovell C.A.K. (2000), *Stochastic Frontier Analysis*, Cambridge University Press, UK.

[25] Lintner, J. (1965), "The Valuation of Risky Assets: The Selection of Risky Investments in Stock Portfolios and Capital Budgets", *Review of Economics and Statistics*, 47,13-37.

[26] Markowitz, H. (1952), "Portfolio Selection", *Journal of Finance*, 77-91.

[27] Markowitz, H. (1959), *Portfolio Selection. Efficiency diversification of investments*, Cowles Foundation for research in Economics at Yale University, John Wiley and sons, Inc.

[28] Morey, M.R., Morey, R.C. (1999), "Mutual fund performance appraisals: a multi-horizon perspective with endogenous benchmarking", *Omega, Int. J. Mgmt Sci.*, 27, 241-258.

[29] Murthi, B., Choi, Y. and Desai, P. (1997), "Efficiency of Mutual Funds and Portfolio Performance Measurement: a Nonparametric Measurement", *European Journal of Operational Research*, 98, 408-418.

[30] Ross, S.A. (1976), "The Arbitrage Theory of Capital Asset Pricing", *Journal of Economic Theory*, 13, 3, 341-360.

[31] Sengupta J.K.(1991), "Maximum Probability Dominance and Portfolio Theory", *Journal of Optimization Theory and Applications*, 71, 341-357.

[32] Sengupta, J.K., and Park, H.S. (1993), "Portfolio Efficiency tests Based on Stochastic Dominance and Cointegration", *International Journal of Systems Science*, 24, 2135-2158.

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

[34] Sharpe, W. (1966), "Mutual Fund Performance", *Journal of Business*, 39, 119-138.

[35] Shephard, R.W. (1970). *Theory of Cost and Production Function*. Princeton, NJ: Princeton University Press.

[36] Shukla, R. and Trzcinca C. (1992), "Performance Measurement of Managed Portfolios", *Financial Markets, Institutions and Instruments*, 1, 1-59.

[37] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

[38] Simar, L. (2003), "Detecting outliers in frontier models: a simple approach" , *Journal of Productivity Analysis*, 20, 391-424.

[39] Simar, L. and Wilson, P.W. (2002), "Nonparametric tests of returns to scale", *European Journal of Operational Research*, 139, 115-132.

[40] Treynor, J.L. (1965), "How to Rate Management of Investment funds", *Harvard Business Review*, 43, 63-75.

[41] van den Broeck J., Koop G., Osiewalski J., and Steel M.F.J. (1994), "Stochastic frontier models: a Bayesian perspective", *Journal of Econometrics*, 61, 273-303.

[42] Zhang, Y., Bartels R. (1998), "The Effect of sample size on mean efficiency in DEA with application to electricity distribution in Australia, Sweden and New Zealand", *Journal of Productivity Analysis*, 9, 187-204.

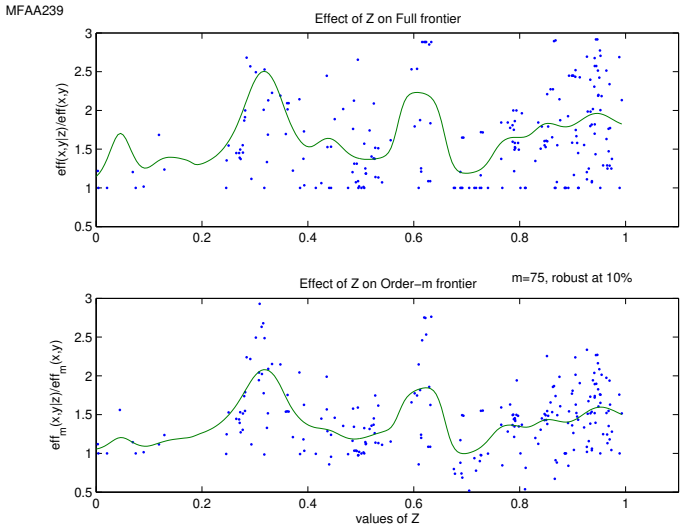


Figure 1: *Evaluating the influence of market risks on the performance of Asset Allocation (AA239) Mutual Funds.*

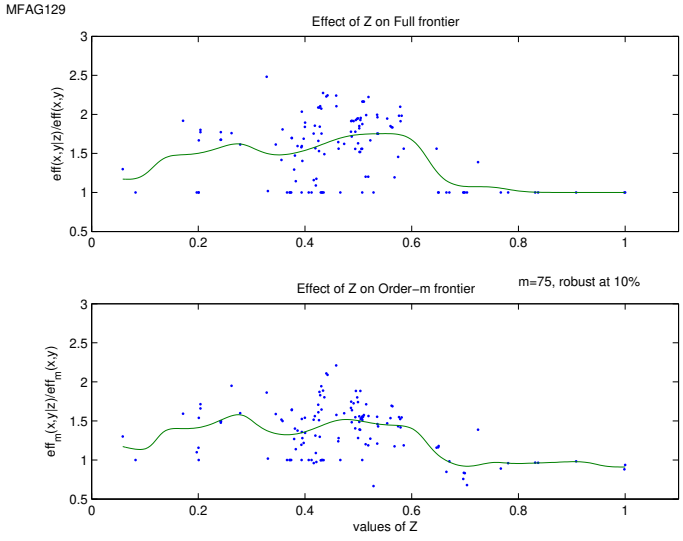


Figure 2: *Evaluating the influence of market risks on the performance of AG129 Mutual Funds.*

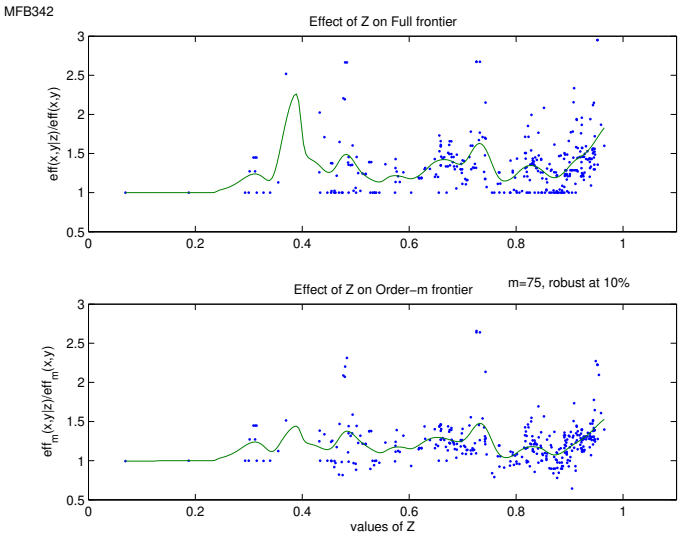


Figure 3: *Evaluating the influence of market risks on the performance of B342 Mutual Funds.*

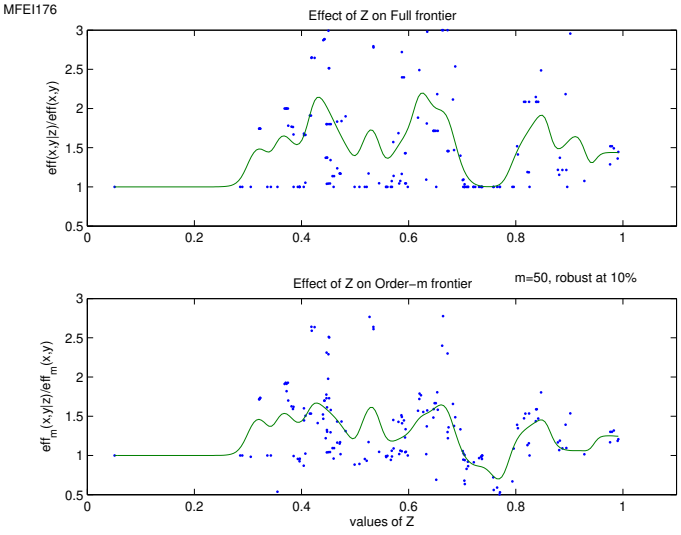


Figure 4: *Evaluating the influence of market risks on the performance of EI176 Mutual Funds.*

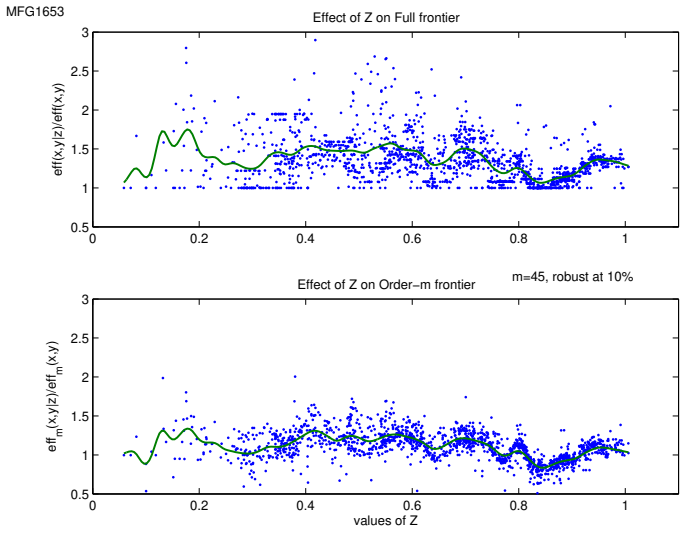


Figure 5: *Evaluating the influence of market risks on the performance of G1653 Mutual Funds.*

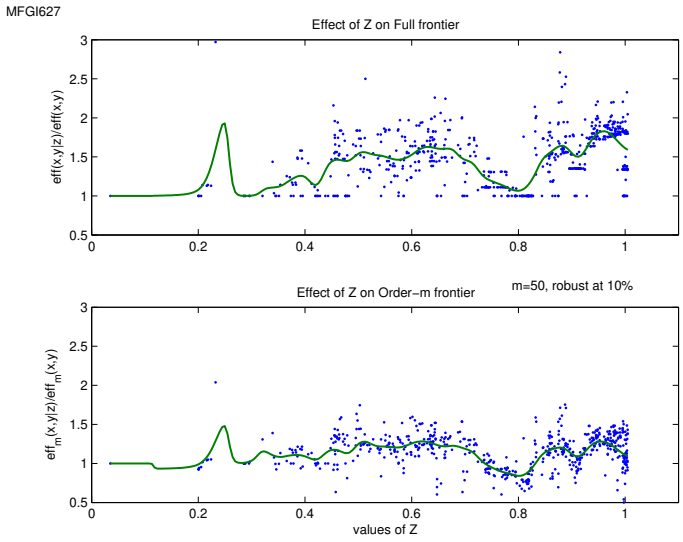


Figure 6: *Evaluating the influence of market risks on the performance of GI627 Mutual Funds.*

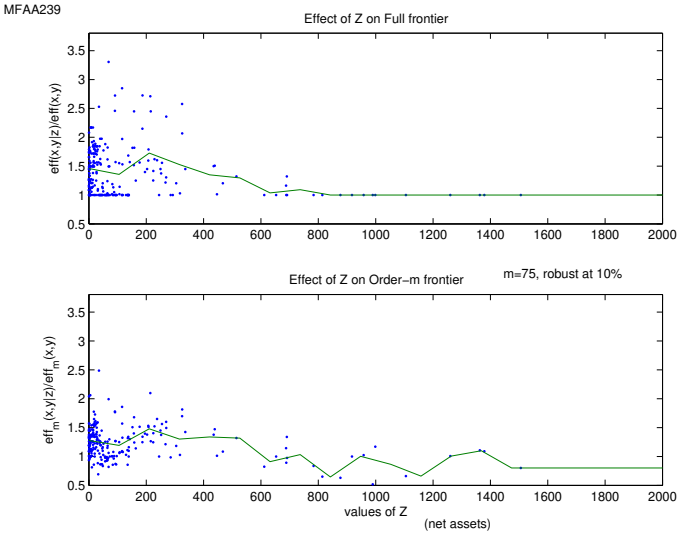


Figure 7: *Evaluating the influence of fund size (net assets) on the performance of AA239 Mutual Funds.*

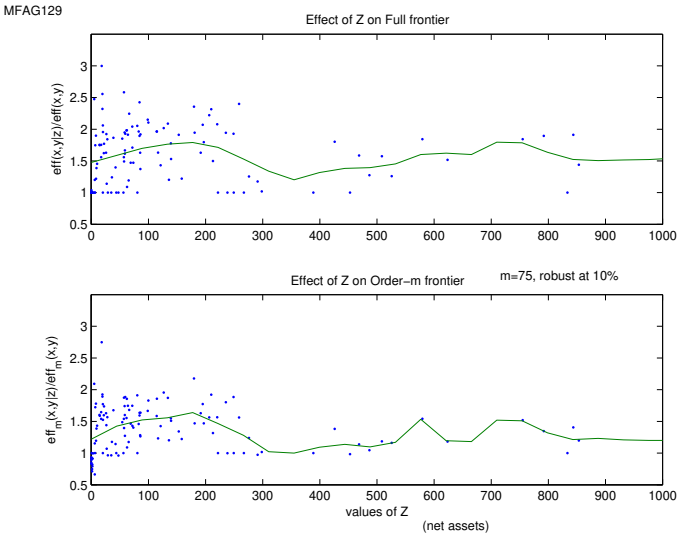


Figure 8: *Evaluating the influence of fund size (net assets) on the performance of AG129 Mutual Funds.*

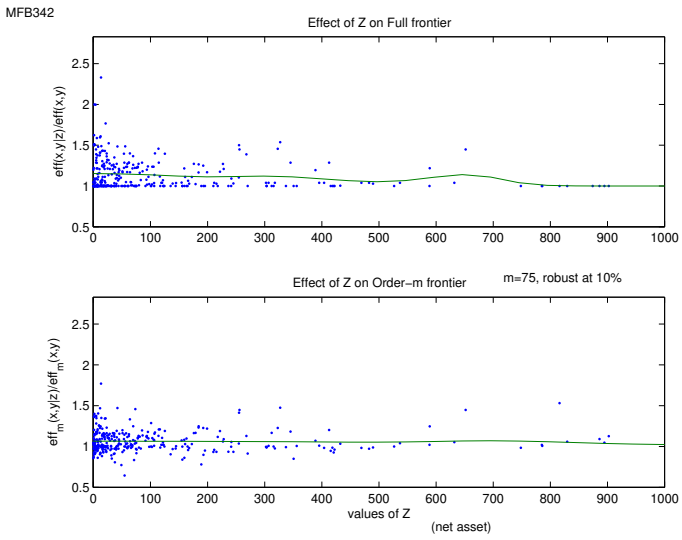


Figure 9: *Evaluating the influence of fund size (net assets) on the performance of B342 Mutual Funds.*

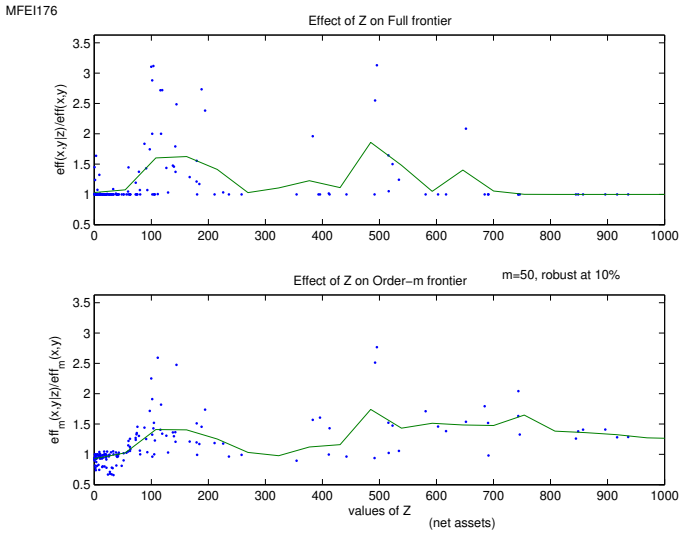


Figure 10: *Evaluating the influence of fund size (net assets) on the performance of EI176 Mutual Funds.*

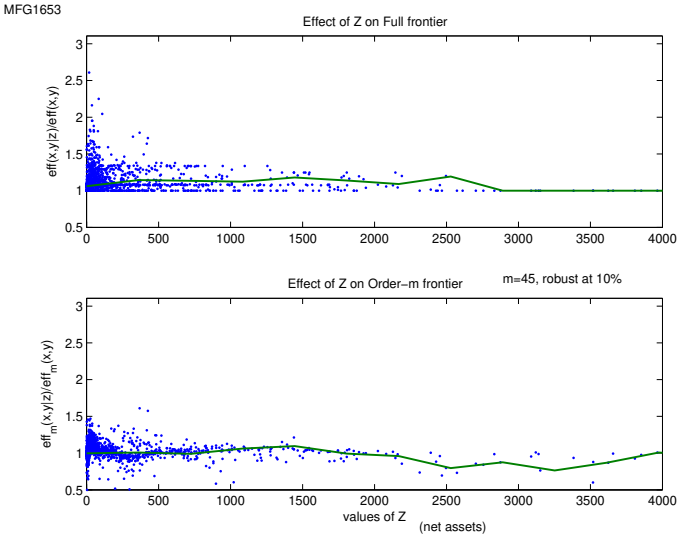


Figure 11: *Evaluating the influence of fund size (net assets) on the performance of G1653 Mutual Funds.*

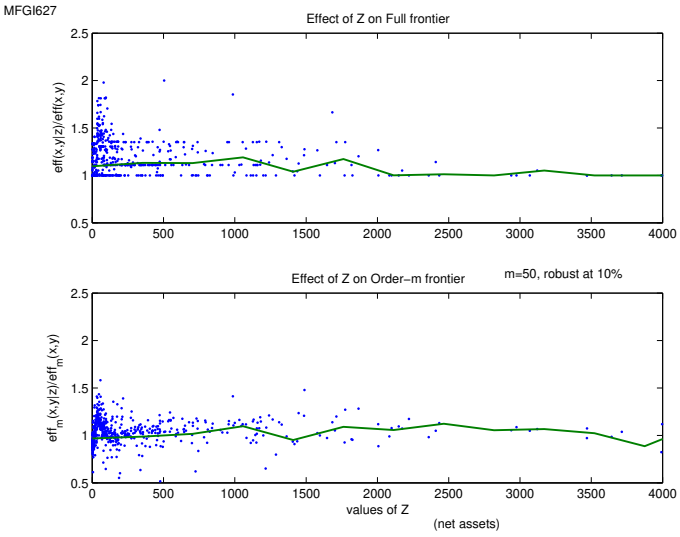


Figure 12: *Evaluating the influence of fund size (net assets) on the performance of GI627 Mutual Funds.*