An analysis of the dynamics of efficiency of mutual funds

Jorge Galán, Sofia B. Ramos and Helena Veiga

Abstract

This paper studies the efficiency of a sample of mutual funds that invest in the United States. Estimating a production function using Bayesian stochastic frontier analysis, we find evidence that the underlying technology presents economies of scale both at the fund and firm level. We also find evidence that informational asymmetries affect efficiency. Funds that invest domestically are likely to be more efficient than foreign funds investing in the US. Moreover, an inspection at the distribution process shows that funds sold directly to investors rather than by financial intermediaries are more efficient. The level of inefficiency persistence is overall high. Persistency of inefficiency is particularly higher for ethical funds, funds oriented to large firms and lower in funds oriented to growth firms. The analysis done in two separate periods also shows that the efficiency of the funds changes. In particular, funds oriented to non-ethical, small and growth firms become more efficient over the period. Finally, funds' efficiency decreases during global financial crisis, but at the end of the sample period some funds recover and their efficiency levels are higher than those registered before the financial crisis. Our results have implications for investors' decisions in mutual funds.

Keywords: Economies of Scale; Efficiency; Mutual Funds; Persistence; Stochastic Frontier Analysis

Acknowledgements: The authors acknowledge financial support from Fundação para a Ciência e Tecnologia PEst-OE/EGE/UI0315/2011 and from the Spanish Ministry of Education and Science, research project ECO2012-32401.
An Analysis of the Dynamics of Efficiency of Mutual Funds

Jorge Galan*  Sofia B. Ramos†  Helena Veiga‡

ABSTRACT

This paper studies the efficiency of a sample of mutual funds that invest in the United States. Estimating a production function using Bayesian stochastic frontier analysis, we find evidence that the underlying technology presents economies of scale both at the fund and firm level. We also find evidence that informational asymmetries affect efficiency. Funds that invest domestically are likely to be more efficient than foreign funds investing in the US. Moreover, an inspection at the distribution process shows that funds sold directly to investors rather than by financial intermediaries are more efficient. The level of inefficiency persistence is overall high. Persistency of inefficiency is particularly higher for ethical funds, funds oriented to large firms and lower in funds oriented to growth firms. The analysis done in two separate periods also shows that the efficiency of the funds changes. In particular, funds oriented to non-ethical, small and growth firms become more efficient over the period. Finally, funds’ efficiency decreases during global financial crisis, but at the end of the sample period some funds recover and their efficiency levels are higher than those registered before the financial crisis. Our results have implications for investors’ decisions in mutual funds.

JEL classification: C11; C23; C51; G11; G14; G15; G24

Keywords: Economies of Scale; Efficiency; Mutual Funds; Persistence; Stochastic Frontier Analysis

*Bank of Spain, Spain.
†NEOMA Business School, France. Business Research Unit/UNIDE, Portugal. Corresponding author: sofia.ramos@neoma-bs.fr.
‡Universidad Carlos III de Madrid, Department of Statistics and Instituto Flores de Lemus, Spain. Business Research Unit/UNIDE, Portugal.
I. Introduction

The performance of the mutual fund industry is a widely debated issue both in academia and in the industry. Early research has documented that risk adjusted performance tends to be negative and the large majority of funds underperform the benchmark (e.g., Daniel et al., 1997; Elton et al., 1993; Ferreira et al., 2013; Grinblatt and Titman, 1989; Indro et al., 1999), which is usually attributed to the lack of skill of the portfolio manager. Previous studies have failed to address whether there are more fundamental reasons that could explain that.

This paper takes a fresh view on the analyses of efficiency of portfolio management of equity mutual funds. We analyze the portfolio management process by estimating the relationship between the expected return of portfolios and risk and identifying the impact of several fund characteristics on efficiency.

In order do that, we use Stochastic Frontier Analysis (SFA) introduced in Aigner et al. (1977) and Meeusen and van den Broeck (1977). This methodology assumes an efficient frontier, which in this context represents those funds who maximize the expected return of their portfolios given its risk, and allows identifying the factors that explain the deviations from this efficient frontier.

Recently, SFA has been studied from a Bayesian point of view due to some of the attractive features of this approach, such as, formal specification of uncertainty, easy incorporation of prior ideas and restrictions, and the computation of distributions of efficiency for each individual firm. Since the introduction of the Bayesian approach to SFA in van den Broeck et al. (1994) there has been an increasing number of theoretical studies and applications of SFA from this perspective and it is currently a very influential approach (see, for instance, the papers by Feng and Zhang, 2012, 2014; Fenn et al., 2008; Lensink and Meesters, 2014).\footnote{A survey on these methodologies can be seen in Murillo-Zamorano (2004) and Greene (2008).} They can overcome problems of performance measurement with traditional performance measures that arise due to large sampling errors, for instance, funds seem to outperform their benchmarks, even when the managers do not have skills to outperform (see Annaert et al., 2003; Kothari and Warner, 2001).

We use a sample of mutual funds that invest in the United States. We assume that the
efficiency is affected by the following factors: First, the nature of the returns to scale relation (i.e. the existence of increasing or decreasing economies of scale). Along the arguments that support increasing economies of scale are that fixed expenses such as managerial and administrative costs that are distributed over a large base of assets under management or large funds that have more investment opportunities available. However, decreasing economies of scale can happen because when large funds trade and they impact the market, it may create adverse effects. Moreover, large funds have difficulty in finding good investment opportunities, or organizational diseconomies due to the large number of securities to deal.

We also investigate the role of economies of scale at family level. Larger families are more capable of affecting the performance of their funds. Guedj and Papastaikoudi (2005) assume that firm expertise depends on total assets under management of the company, as larger companies are likely to be better equipped with resources (e.g. fund managers, research teams, databases, software or administrative staff).

The estimation supports the existence of increasing economies of scale. Larger size leads to higher efficiency and also funds from large families are likely to be more efficient. Fund flows relate positively with inefficiency and also inefficiency is positively related with volatility of flows since managers have to keep large amounts in cash that erodes performance or to buy or sell securities frequently that also erodes performance.

We also consider information as an efficiency determinant. The models of Gehrig (1993) or Brennan and Cao (1997) have posit that foreign portfolio managers are at informational disadvantage to local portfolio managers, therefore we investigate whether it relates with inefficiency. The results point that informational asymmetries are likely to be a source of inefficiency as we find that funds that invest domestically are likely to be more efficient than foreign funds investing in the United States.

By following a certain investment style, the portfolio managers commit to certain stocks and restrain their investment opportunity set. A consequence of restricting the investment universe is that diversification benefits decrease, therefore efficiency might decrease. Moreover, additional
inefficiencies might arise because of the type of asset selected. Confirming our hypothesis, we find that ethical funds are likely to be more inefficient as well as funds that invest in large companies. On the contrary, we find that funds oriented to growth firms are likely to be more efficient.

The way the fund is delivered to the investor also matters in the management process. This can arise because the effort of the portfolio managers is likely to be higher for some types of investors. For instance, less informed investors might choose a certain channel of distribution and these investors are less likely to exert less pressure on fund managers. The results show that funds sold directly rather than using financial intermediaries are more efficient.

Finally, we analyze the persistence of inefficiency by modeling a dynamic specification for the inefficiency that captures the proportion of inefficiency that is transmitted from one period to the next. We estimate fund-specific inefficiency persistence, which allows us to identify important differences in persistence among funds with different characteristics. We observe that funds with more persistence inefficiency do not become more efficient and even might be less efficient in the future. Persistency of inefficiency is particularly higher for ethical funds, funds oriented to large firms and lower in funds oriented to growth firms. The high level of inefficiency persistence suggests that adjustment costs on the portfolio can be high preventing managers from making instant adjustments. Moreover, regulation, information failures and other management rigidities may cause funds to remain partly inefficient in the short-run.

The analysis done in two separate periods, 2003 and 2013, also shows that the inefficiency persistence keeps constant over the period of analysis but the efficiency of the funds change considerably. In particular, non-ethical funds and funds oriented to small and growth firms become more efficient over the period. These funds are those with less inefficiency persistence.

Finally, during the recent financial crisis the funds’ efficiency decreases drastically. Nevertheless, at the end of the sample some funds recover and they efficiency levels are even higher than those registered before the financial crisis.

Our work contributes to the understanding of the efficiency of the mutual fund industry confirming that factors, such as economies of scale, information, style, and incentives are relevant
inefficiency drivers. In addition, the level of inefficiency seems quite persistent suggesting that adjustment costs are high or other barriers prevent the portfolio management process from being more efficient. Moreover, our paper contributes to the burgeoning literature that uses parametric frontier methods to measure mutual fund performance by considering a dynamic stochastic frontier model.

The paper has the following structure. Section II introduces the methodology. In particular, we define the frontier and propose a dynamic stochastic frontier model to mutual funds that includes inefficiency persistence and observed heterogeneity. Section III describes the data used in the paper. Section IV presents and discusses the results. The final remarks are presented in Section V.

II. Performance Methodology

In the finance literature several measures of fund performance have been developed since the seminal papers of Treynor (1965), Sharpe (1966) and Jensen (1968). Our main benchmark model for evaluating performance is the three-factor model of Fama and French (1993) as in Fama and French (2010) that improves the Capital Asset Pricing Model (CAPM) by including additional factors such as size and book-to-market and matches the categorization if investment style of US funds (see e.g. Morningstar style box).

In our case, for a fund $i$ at month $t$ the alpha corresponds to:

$$\alpha_{it} = R_{it} - E(R_{it}) = R_{it} - \beta_{0it}E(RM_t) + \beta_{1it}E(SMB_t) + \beta_{2it}E(HML_t),$$

where $R_{it}$ is the excess return in US dollars of fund $i$ in month $t$ from 1-month interbank rate, $E(.)$ denotes the expectations operator, $RM_t$ is the excess return in US dollars on the market in month $t$; $SMB_t$ (small minus big) is the average return on the small-capitalization portfolio minus the average return on the large-capitalization portfolio in month $t$; $HML_t$ (high minus low) is the difference in return between the portfolio with high book-to-market stocks and the
portfolio with low book-to-market stocks in month $t$. Moreover, $\beta_{0it}$ is the beta of the excess return on the market portfolio for fund $i$ at time $t$, $\beta_{1it}$ is the beta of the average return on the small-capitalization portfolio minus the average return on the large-capitalization portfolio for fund $i$ at time $t$ and $\beta_{2it}$ is the beta of the difference in return between the portfolio with high book-to-market stocks and the portfolio with low book-to-market stocks for fund $i$ at time $t$. The second equality follows from assuming the following three-factor model of Fama and French (1993) to hold:

$$R_{it} = \alpha_{it} + \beta_{0it}RM_t + \beta_{1it}SMB_t + \beta_{2it}HML_t + \varepsilon_{it}. \tag{1}$$

This benchmark model improves average CAPM pricing errors by including size and book-to-market factors and allow us to compare the efficiency controlling for the style of the fund.

In a sample, $\alpha_{it}$ can be estimated using a panel regression as

$$R_{it} = \alpha_0 + \beta_0 R_{Mt} + \beta_1 S_{MBt} + \beta_2 H_{MLt} + \alpha_{it}, \tag{2}$$

where the beta coefficients are estimated from equation (1) using a rolling window of 36 months.$^2$

A. Stochastic frontier methodology

Frontier stochastic analysis mainly used to assess technical and economic efficiency of firms within a sector can be also useful to understand the efficiency of portfolio management. It is designed to capture the deviation of the funds’ performance from what an efficient portfolio can achieve. As Annaert et al. (2003), we assume that mutual funds cannot earn systematically positive abnormal returns. This assumption is in accordance with the assumption of efficient market hypothesis and with the evidence of the empirical literature. However, empirically we can find positive significant excess returns due to sampling noise. A way of increasing power is to augment model (2) with a composed error that consists of an idiosyncratic error and a non-negative component.

$^2$Note that our measure of performance is already adjusted to risk, so we do not consider it as an input.
that measures inefficiency, such as:

\[ R_{it} = \gamma_0 + \gamma_1 \beta_{0it} + \gamma_2 \beta_{1it} + \gamma_3 \beta_{2it} + v_{it} - u_{it}, \]  

(3)

where \( v_{it} \) is the idiosyncratic error assumed to follow a normal distribution, and \( u_{it} \) is the inefficiency component. The error \( u_{it} \) comes from the assumption that a managed fund cannot systematically outperform the market portfolio. Model (3) specifies the stochastic frontier. Earlier studies have presented different choices regarding the output such as total gross returns (Daraio and Simar, 2006), annualized 3-years returns (Tsolas, 2014) or risk-adjusted return from a one-factor model (Annaert et al., 2003).

SFA has the advantage of allowing inferences on the parameters and considering idiosyncratic errors. It also allows dealing easier with panel data structures and to model the evolution of efficiency over time. Certainly, two different approaches have been used in the SFA literature for this purpose. The most common approach is to use deterministic specifications of time (see Kumbhakar, 1990; Battese and Coelli, 1992, for some proposals). These models have the problem of imposing arbitrary restrictions on the short-run efficiency and they are not able to model firm-level dynamic behavior. An alternative approach is to specify an autoregressive structure that recognizes the dynamic behavior of the inefficiency (see Ahn and Sickles, 2000; Tsionas, 2006). In particular, Regulation, transaction costs, information failures and other rigidities may prevent managers from making instant adjustments towards optimal conditions and might cause inefficiency persistence. Funds’ managers may find it optimal to remain partly inefficient in the short-run.

**B. A dynamic stochastic frontier model**

In this Section, we propose to model observed heterogeneity and performance persistence in mutual funds with SFA. We apply the dynamic specifications proposed by Galán and Pollitt
(2014) and Galán et al. (2015) in order to capture unobserved sources of heterogeneity\(^3\), i.e., the proposed models allow modeling fund-specific inefficiency persistence through a random autoregressive coefficient and they extend the dynamic model in Tsionas (2006). Finally, the general implemented model is given by the following equations:

\[
\begin{align*}
R_{it} &= \hat{\beta}_{it} \gamma + v_{it} - u_{it}, & v_{it} &\sim N(0, \sigma^2_v) \\
\log u_{it} &= \omega + z_{it} \delta + \rho_i \log u_{i,t-1} + \xi_{it}, & \xi_{it} &\sim N(0, \sigma^2_{\xi}), \quad t = 2 \ldots T \\
\log u_{i1} &= \omega + z_{i1} \delta + \xi_{i1}, & \xi_{i1} &\sim N \left(0, \frac{\sigma^2_{\xi}}{1 - \rho^2_i}\right), \quad t = 1.
\end{align*}
\]

Equation (4) is equivalent to equation (3) and it represents the stochastic frontier, where \(R_{it}\) is the excess return for fund \(i\) at time \(t\), \(\hat{\beta}_{it}\) is a row vector, \(\gamma\) is a vector of parameters including a constant, \(v_{it}\) is the idiosyncratic error assumed to follow a normal distribution, and \(u_{it}\) is the inefficiency component. The dynamic specification for the inefficiency is represented by (5), where \(\omega\) is a constant term, \(z_{it}\) is a row vector of observed heterogeneity variables, \(\delta\) is a vector of parameters, \(\rho_i\) is the fund-specific persistence parameter that captures the proportion of inefficiency that is transmitted from one period to the next for every fund, and \(\xi_{it}\) is a white noise process with constant variance \(\sigma^2_{\xi}\), which may capture unobserved random shocks in the dynamic component. Finally, equation (6) represents the specification of the inefficiency in the first period and is intended to initialize a stationary dynamic process.

In order to avoid divergences of \(\log u_{it}\) to positive or negative infinity, which would lead to efficiencies equal to zero or to one, stationarity is imposed by requiring \(|\rho_i| < 1\). In general, if a firm has a value of \(\rho_i\) close to 1 it would suggest that this fund presents high transaction or management costs, which translates into a high proportion of inefficiency being transmitted from one period to the next. On the other hand, if this value is close to 0, a low proportion of inefficiency is persistent in time, implying that the fund moves quick towards more optimal conditions.

\(^3\)Galán et al. (2015) also study the effects of including observed variables in and out of the dynamic component of the inefficiency.
The general model in (5) and (6) allows to evaluate different specifications by imposing restrictions over some parameters. If \( \rho_i = \rho \) is imposed, homogeneous persistence is assumed for all companies in the sector. If \( \rho = 0 \) the model reduces to a static model where the inefficiency follows a log-normal distribution with firm specific mean. Finally, if \( \delta = 0 \), no observed inefficiency heterogeneity is modeled.

Using the results for the posterior inefficiencies \( u_{it} \), the efficiency of individual funds in each period, which is a measure between 0 and 1, is calculated as:

\[
Eff_{it} = \exp(-u_{it}).
\]  

(7)

**C. Bayesian Inference**

The inference of the dynamic model presented in equations (4)–(6) is performed through Bayesian methods. The Bayesian approach to SFA was introduced by van den Broeck et al. (1994) and presents as its main advantages the formal incorporation of parameter uncertainty, the derivation of posterior distributions of inefficiencies for every fund, and the easy modeling of random parameters through hierarchical structures.

We assume non-informative but proper prior distributions for all the parameters throughout. For parameters in \( \gamma \) we assume a normal prior distribution such that \( \gamma \sim N(0, \Lambda_{\gamma}^{-1}) \) where \( \Lambda_{\gamma} \) is a precision diagonal matrix with priors set to 0.001 for all parameters. The variance of the idiosyncratic error component is assumed to follow an inverse gamma distribution \( \sigma_v^2 \sim IG(a, b) \) with priors set to 0.01 and 100 for the shape and scale parameters.

As defined in (5) and (6), the inefficiency component follows a log-normal distribution where

\[
u_{it}|u_{i,t-1}, \omega, z_{it}, \delta, \rho_i, \sigma_\epsilon^2 \sim LN(\omega + z_{it} \delta + \rho_i \log u_{i,t-1}, \sigma_\epsilon^2) \] for \( t = 2...T \); and

\[
u_{i1}|\omega, z_{i1}, \delta, \rho_i, \sigma_\epsilon^2 \sim LN\left(\frac{\omega + z_{i1} \delta}{1-\rho_i}, \frac{\sigma_\epsilon^2}{1-\rho_i}\right) \] for \( t = 1 \). For these inefficiency parameters, we have that \( \omega \sim N(\mu_\omega, \lambda_\omega^{-1}) \) with priors set to -1.5 and 1 for the mean and precision parameters, respectively; \( \delta \sim N(0, \Lambda_\delta^{-1}) \) where \( \Lambda_\delta^{-1} \) is a diagonal matrix of precisions with priors set to 0.1 for every precision parameter; and, for the fund-specific persistence parameters, we define a hierarchical structure with \( \rho_i = 2k_i - 1 \),
where \( k_i \sim \beta(k, 1-k) \). This hyperparameter is distributed \( k \sim \beta(r, s) \) with priors set to 0.5 for shape parameters. The variance of the inefficiency component is assumed to follow an inverse gamma distribution where \( \sigma^2_\zeta \sim G(n, d) \) with priors set to 10 and 0.01 for the shape and scale parameters, respectively.\(^4\)

Given the intractability of the joint posterior distribution in these models, numerical integration methods such as Markov Chain Monte Carlo (MCMC) have to be used (see Koop et al., 1995, for the implementation of the Gibbs Sampling algorithm with data augmentation). In our case, we carry out the implementation of the proposed model using the WinBUGS package (see Griffin and Steel, 2007, for a general procedure in applications to SFA). The MCMC algorithm in all the estimated models involves 30,000 iterations with 10,000 discarded in a burn-in phase and a thinning equal to 5 to remove autocorrelations.

Finally, we perform sensitivity analysis using different values for prior parameters in the distributions of \( \omega, k \) and \( \sigma^2_\zeta \) and posterior results are found to converge to approximately the same values. For the persistence parameter \( \rho \) we also studied the sensitivity to the use of a truncated normal distribution and posterior results were found to be robust to the use of this alternative.

C.1. Model selection

We use two different model election criteria widely used under the Bayesian approach to SFA. The first one is a robust version of the Deviance Information Criterion (DIC) called \( DIC_3 \), as developed in Richardson (2002) and Celeux et al. (2006). \( DIC \) is a within-sample measure of fit introduced by Spiegelhalter et al. (2002) and defined as:

\[
DIC = 2\overline{D(\theta)} - D(\bar{\theta})
\]

with \( D(\theta) = -2\log f(y|\theta) \), where \( D(\theta) \) defines the deviance of a model with parameters \( \theta \) and data \( y \). The alternative \( DIC_3 \) uses an estimator of the density \( f(y|\theta) \) instead of the posterior mean \( \bar{\theta} \) and has been found to be more stable in models with random effects, mixtures and with data

---

\(^4\)These priors have been used before in Tsionas (2006) and Galán and Pollitt (2014) and centers the efficiency prior distributions at 0.8.
augmentation (see Li et al., 2012). The formulation of the $DIC_3$ is the following:

$$DIC_3 = -4E_\theta[\log f(y|\theta)|y] + 2 \log \widehat{f}(y).$$  \hfill (8)

The second criterion is the Log Predictive Score (LPS). This is a predictive performance criterion that assesses the out-of-sample behavior of the models. This criterion was first introduced by Good (1952) and is intended to examine model performance by comparing its predictive distribution with out-of-sample observations. In order to do this, the sample is split into a training and a prediction set. In our specific case, the training set contains all the observations except those for the last period of every fund, and the prediction set consists of those observations for the last observed period. The formulation of the $LPS$ is the following:

$$LPS = -\frac{1}{k} \sum_{i=1}^{k} \log f(y_{i,t_i}|\text{previous data}),$$  \hfill (9)

where $y_{i,t_i}$ are the observations in the predictive set for the $k$ funds in the sample and $t_i$ represents the penultimate time point with observed data for fund $i$.

Applications of both criteria to SFA models can be found in several papers using the Bayesian approach (see Galán and Pollitt, 2014; Griffin and Steel, 2004; Ferreira and Steel, 2007, for some examples).

### III. Data

Our dataset of funds is extracted from Lipper. The sample is compiled of open-end equity funds that invest in the United States from several countries. The estimation is done quarterly time series for the nine-year period from January 2003 to December 2013.\footnote{This database has been used in the paper of Ferreira et al. (2012, 2013) among others studies.} We consider only primary funds and exclude different asset classes.

For a fund to be in our sample it must report information on net asset value and monthly returns. Moreover, we require that it has information for all the variables used in the models.
and to have data for at least 40 consecutive months. The latter is important given the dynamic specification that we use in our models. The final data set contains 1,124 funds. After imposing some filters on explanatory variables we obtain a total of 50,586 fund-quarters.

A. Fund Features

In this subsection we present the variables of funds that we consider as inputs.

**Economies of scale** The returns to scale of the underlying technology is a natural input to consider in a production function. There has been an intense debate in the literature about the nature of economies of scale in the mutual fund industry. Among the arguments that support increasing economies of scale are that fixed expenses such as managerial and administrative costs are distributed over a large base of assets under management, more opportunities available for large funds or fixed transaction costs decrease with size.

Elton et al. (2012) find that large US funds perform better. Reuter and Zitzewitz (2010) find little impact of fund growth on fund performance, while Pastor et al. (2014) finds that once industry size is controlled, there is constant returns to scale in fund management in the United States.⁶

On the contrary, decreasing economies of scale can happen because when large funds trade they impact the market, generating adverse effects such as the price impact associated with large transactions (Perold and Salomon, 1991; Beckers and Vaughan, 2001). Edelen et al. (2007) advocate that trading costs are the primary source of diseconomies of scale and find evidence that large funds do not underperform per se, they underperform to the extent they incur in trading costs. Chen et al. (2004) argue that the lagged fund size and performance relationship is due to transactions costs associated with liquidity or price impact. They find that the adverse effect of size and performance is stronger among “small cap funds”. To the extent that small cap stocks

---

⁶For non US mutual funds both Otten and Bams (2002) and Ferreira et al. (2013) find a positive relation between risk-adjusted performance and fund size suggesting the presence of economies of scale for European funds.
are less liquid, they argue that their evidence provides support for the hypothesis that fund size erodes performance because of liquidity. Yan (2008) also finds an inverse relation between fund size and fund performance, which is stronger among funds that hold less liquid portfolios. He concludes that the inverse relation between fund size and fund performance is more pronounced among growth and high turnover funds that tend to have high demands for immediacy.

Another reason for decreasing economies of scale is that large funds have difficulty in finding good investment opportunities that fit their scale and they have to accept inferior investments. Chen et al. (2004) argue that small funds can easily put all the money in its best ideas, but because of lack of liquidity a large fund has to invest in “not-so-good-ideas” thereby eroding performance. Pollet and Wilson (2008) study how funds respond to asset growth. They find that funds are reluctant to diversify in response to growth but instead tend to acquire even larger ownership shares in the companies they already own. However, they find that small-cap funds diversify as they growth, which is associated with better performance. They argue that the cause of diminishing returns to scale for mutual funds is the inability to scale an investment strategy related to liquidity constraints as the fund grows.

Other studies refer organizational diseconomies for explaining decreasing returns to scale. As funds grow this is accompanied by a large increase either in the variety of securities in portfolio or in the number of accounts, and other authors argue that economies of scale are offset due to the complexity of the number of securities to deal with and the large number of accounts (Amel et al., 2004; Daraio and Simar, 2006).

Other authors highlighted the existence of an optimal fund size. For Indro et al. (1999) mutual funds must attain a minimum fund size in order to achieve sufficient returns to justify their costs of acquiring and trading information. Furthermore there are diminishing marginal returns to information acquisition and trading, and the marginal gains become negative when the mutual fund exceeds its optimal fund size. Bodson et al. (2011) finds a quadratic relation between the size of the fund and performance. Grinblatt and Titman (1994) find mixed evidence that fund returns decline with fund size.
Following the above literature fund size is measured by the logarithm of total net asset values in millions of USD \((\ln t_{na})\). To capture non linear effects we include the square term of total assets \((\ln t_{na}^2)\).

**Economies of scale at family level** We follow Guedj and Papastaikoudi (2005) that argue that larger families are more capable of affecting the performance of their funds. They assume that firm expertise depends on total assets under management of the company, as larger companies are likely to be better equipped with resources. Resources can be fund managers, research teams, databases, software or administrative staff. Certain resources as research teams, databases or administrative stuff can be shared between several funds. Therefore being part of a large company can economize certain fixed costs creating increasing returns to scale.

The works of Chen et al. (2004), Elton et al. (2012) and Ferreira et al. (2013) show that controlling for fund size, belonging to larger families increases the funds performance. The authors attribute this to economies associated with trading commissions and lending fees at family level.

Following previous work, we use the logarithm of the sum of total net assets of the fund firm \((\ln t_{na\_family})\).

We also investigate if fund flows can also affect fund efficiency.

Fund flows are computed as the according to the standard definition in the literature (see e.g. Chevalier and Ellison, 1997; Sirri and Tufano, 1998), as the change (in percentage) in total net assets due exclusively to new external money:

\[
Flow_{it} = \frac{NAV_i - NAV_{i\_t-1} \cdot (1 + r_{it})}{NAV_{i\_t-1}},
\]

thus \(Flow_{it}\) is the percentage growth in TNA in a month \(t\) net of internal growth (assuming reinvestment of dividends and distributions), where \(NAV_i\) is the total net asset value of fund \(i\) at the end of month \(t\) and \(r_{it}\) is fund is raw return in month \(t\). This definition assumes flows take
place at the end of the month as we have no information regarding the timing of new investment. To ensure that extreme values do not drive our results, we winsorize fund flows at the bottom and top one-percent level of the distribution.

A manager that has to deal with volatile flows has to keep large amounts in cash or to buy or sell securities frequently that erodes performance. Therefore we expect a positive relation between inefficiency and volatility of flows. Volatility of flows is computed based on the last 12 months flows ($sd_{flows}$).

**Informational advantages** Superior information can be a source of abnormal performance and we consider it an input in the production process. Several works assume that foreign investors are at informational disadvantage to local investors (see e.g. Gehrig, 1993; Brennan and Cao, 1997). Also empirical evidence supports that geographical proximity is a source of informational advantages. Shukla and van Inwegen (1995) and Clare et al. (2013) find that US mutual fund managers perform better than other portfolio managers when they invest in the United States.

We test whether informational advantage of investing locally is important in explaining inefficiency. We create a dummy variable that assumes the value of one if it is a domestic fund ($domestic$), i.e. if the portfolio manager invests in local stocks.\(^7\)

**The investment style of the fund** By following a certain investment style the portfolio manager commit to certain stocks and certain inefficiencies might arise. A first obvious consequence of restricting the investment universe is that diversification benefits decrease, which is likely to increase inefficiency.

We consider several styles of funds. Ethical funds take into account responsible, environmental and ethical goals in their investment goals. Their investment universe is therefore restricted, which implies that the diversification of the portfolio is limited which must lead to higher risk or lower return compared to a well-diversified portfolio with no limitations in its investment universe.

\(^7\)We cannot disregard that larger portfolio management companies have larger resources and are better equipped to make better investment decisions and produce more efficient funds. Therefore our variable $ln_{family}$ can also be related with informational advantages at family level.
On the other hand, ethical funds mostly include companies with good corporate governance and therefore they are likely to perform better than do conventional funds, which cover all sorts of companies. The conclusions from studies that compare the performance of ethical versus non-ethical funds find mixed results. Bauer et al. (2005), Bello (2005) and Statman (2000) do not find statistically significant differences between the performance of conventional funds and ethical funds, while Renneboog et al. (2008) find that ethical funds tend to underperform the benchmark. We use a dummy variable to indicate if the fund is classified as ethical (ethical).

A well-know classification of investment styles are growth versus value investing and it is widely used to classify mutual funds. In a similar way we also indicate if the fund invests in growth or value firms (growth). Daraio and Simar (2006) finds only evidence of economies of scale for the following US mutual funds categories: Balanced, Growth and Growth Income.

Another well-know style is large versus small capitalization firms. One of the reasons that studies pointed for decreasing economies of scale is the investment in small capitalization stocks that are less liquid. Therefore, we create a dummy variable that indicates whether the fund invests in large capitalization firms firms (large).

**The Distribution process**  The studies that address the performance of mutual and distribution channels are scarce, however the way the fund is distributed also might impact the efficiency of the production process. The impact might be direct on the costs of distribution or indirect, because the type of investors in each distribution channel might be different and might exert different pressure on the effort of the portfolio manager. Recently Guercio and Reuter (2014) show that investing in direct-sold US equity mutual funds are more performance-sensitive (when compared to funds sold through brokers) which creates greater incentives for fund managers to engage in more active investment strategies and generate alpha.

Bergstresser et al. (2009), studying the US mutual fund industry from 1996 to 2004, find that broker-sold funds deliver lower risk-adjusted returns relative to direct-sold funds. We also use dummy variables to indicate the type of distribution of the fund: Bank retail (bank), dealer
(dealer) or direct (direct). We identify the direct channel because the direct channel is less costly for portfolio management firms as they do not need to rely on intermediaries. We also use dummies to identify distribution by broker or by bank that are an important source of product information for uninformed investors. Following Bergstresser et al. (2009), we expect that direct sold funds are more efficient.\(^8\)

**Institutional** Our sample also comprises a sample of funds classified as institutional. This segment is described as more sophisticated therefore more likely to choose efficient funds. We use a dummy variable to indicate if the fund is classified as institutional (institutional).

**IV. Empirical Results**

In this section we present the results of the estimation of the model.

**A. All sample**

Table II presents the results of the estimation for actively managed funds. The coefficient of \(\ln \text{tna}_\text{family}\) is negative indicating that more efficient funds are likely to be managed by large families. This result is consistent with the evidence of Chen et al. (2004) for US funds and for Ferreira et al. (2013) for a sample of international funds and supports the explanation that being part of large company can economize certain fixed costs such as research teams, databases or administrative staff.

The coefficients of \(\ln \text{tna}\) and \(\ln \text{tna}^2\) show that the relation between size and efficiency is non-linear, funds gain efficiency by becoming large. We extend the estimation with additional variables such as fund flows and standard deviation of flows. Both coefficients are positive

\(^8\)Many studies have used expense ratios and loads as inputs. As they are commonly referred as "costs". They are costs for investors as they are charged to investors. However they do not correspond to costs in the production function of portfolio managers. They are revenues for firms, so firms try to maximize them. We consider inputs from the perspective of portfolio management company. We opt not to use age of the fund as an input, since earlier studies do not find that age seems to be important and it is usually correlated with the size of the fund (see Annaert et al., 2003).
implying that both variables relate positively to inefficiency. As expected a higher volatility of fund flows contributes to inefficiency as fund managers have to keep more cash that in turn erodes performance.

The coefficient on *domestic* is negative confirming our expectation that informational advantages are important to efficiency. Informational advantages of geographically close portfolio managers lead to more efficient funds than their foreigner peers, which is consistent with earlier evidence (Shukla and van Inwegen, 1995). Overall, it supports the importance of information on efficiency.

Next, we can find the coefficient of variables related with the style of the fund. The coefficients show that ethical funds, and funds that invest in large capitalization firms are less efficient. The coefficient is large for funds that invest in large capitalization firms. Funds that are more oriented to growth firms are more efficient consistent with Daraio and Simar (2006) that also find that growth funds are more efficient. The coefficient on *institutional* is positive indicating that funds for institutional investors are less efficient.

The results show that funds distributed by banks and dealers are less efficient, while funds sold directly to investors are more efficient. The result on directly sold funds is consistent with Guercio and Reuter (2014) that show that fund sold directly are less likely to underperform and also with the result of Bergstresser et al. (2009) that find that fund sold by dealers underperform direct sold funds.

Table III provides more details on how economies of scale change with style. To analyze that we interact the coefficients of $\ln{t\text{na}}$ and $\ln{t\text{na}}^2$ with the style of the fund. Earlier results have shown that as funds grow they become more efficient. The interaction coefficients show that this is stronger for domestic and funds oriented to growth firms, but not for ethical funds and funds that invest in large firms.
B. Persistence of Inefficiency

In this section we analyse whether inefficiency persists over time. Traditionally studies measure performance persistence using correlation between returns, frequency tables or relating abnormal performance of one period with past period. They address the issue by analyzing whether “winners” in one period persist in being “winners” in the next period. Hendricks et al. (1993) and Brown and Goetzmann (1995) find that winning funds (“winners”) over a reference period are more likely to be in the better performing group in the subsequent period than in the worse performing group (“losers”), but Carhart (1997) only finds persistence for losers, that is said on funds that underperform the benchmark. Using SFA, Annaert et al. (2003) find that poor performers tend to be less efficient in a subsequent period.

The persistence parameter $\rho$ in the model measures the proportion of inefficiency that is transmitted from one period to the next due to costs of adjustment and restrictions in their portfolios. The models allow that each fund has its specific $\rho_i$, recognizing heterogeneity in the adjustment process of inefficiency among funds with different characteristics. Tables II–III present the posterior mean of $\rho_i$. Its estimates range from 0.6532 to 0.7163, which represents a quite high inefficiency persistence and they are different from zero in all models. Figure 1 shows the densities of the posterior inefficiency persistence by fund style. The top panel of the figure displays the posterior densities by fund styles and their complementaries. We observe that funds with investment styles oriented to large, ethical and non-growth firms present more inefficiency persistence. On the other hand, the bottom panel of the figure presents the densities of the posterior inefficiency persistence of large, ethical and growth funds. We observe, in general, that the least inefficiency persistent funds are funds oriented to growth and those with the most inefficiency persistence are funds oriented to growth firms. Figure 2 plots the inefficiency persistence versus efficiency by fund style in two periods 2003 and 2013. We observe that the inefficiency persistence keeps constant over the period of analysis but the efficiency of the funds change considerably. In particular, non-ethical, small and growth funds become more efficient over the period. These funds are those with less inefficiency persistence. Those with the most
inefficiency persistence register the same efficiency or worsen their performance over the period. They are ethical, large and non-growth funds.

Finally, Figure 3 shows that around 2008 the funds’ efficiency decreases drastically. These year corresponds to the beginning of the financial crises. Nevertheless, in 2013 funds oriented to growth firms and non large and non ethical segment recover and they efficiency levels are higher than those register in 2003.

V. Conclusion

In this paper we analyze the efficiency of the portfolio management of mutual fund using stochastic frontier models. This methodology allows us to measure the deviation of the funds performance from what an efficient portfolio can achieve and to identify key inefficiency determinants. In particular, we propose a dynamic stochastic frontier model that is able to capture fund-specific persistence of inefficiency and to model scale technology, information, investment style of the fund and the distribution channel as inefficiency drivers.

The results show that larger funds are less inefficient than small ones which supports increasing economies of scale in the technology. Moreover, funds from larger mutual fund companies are less inefficient supporting the existence of economies of scale at family level. The results point that informational asymmetries are likely to be a source of inefficiency as we find that funds that invest domestically are likely to be more efficient than funds foreign funds investing in the United States. The results also show that some investment styles are more inefficient such as ethical funds and funds that invest in large firms, while funds oriented to growth firms are more efficient. We find evidence that the distribution process also matters. Funds sold directly to investors than by financial intermediaries are more efficient which might due to investors in the direct channel be more sensitive to performance.

Results also show that a lot of inefficiency persists to the other period, which might be due to low flexibility in the management process. Moreover, the least inefficiency persistent funds are
that oriented to growth firms and those with the most inefficiency persistence are that oriented to large cap firms. The analysis done in two separate periods, 2003 and 2013, also shows that the inefficiency persistence keeps constant over the period of analysis but the efficiency of the funds change considerably. In particular, non-ethical funds and funds oriented to small and growth firms become more efficient over the period. These funds are those with less inefficiency persistence.

Finally, at the beginning of the recent financial crisis the funds’ efficiency decreases drastically. Nevertheless, at the end of the sample some funds recover and they efficiency levels are higher than those registered before the financial crisis.

References


Figures and Tables

Table I

Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Returns</td>
<td>0.0057</td>
<td>0.0347</td>
<td>-0.1848</td>
<td>0.1523</td>
</tr>
<tr>
<td>Beta market</td>
<td>0.6430</td>
<td>0.4731</td>
<td>-0.2582</td>
<td>2.4492</td>
</tr>
<tr>
<td>Beta smb</td>
<td>0.0055</td>
<td>0.1772</td>
<td>-1.8952</td>
<td>1.5539</td>
</tr>
<tr>
<td>Beta hml</td>
<td>-0.0186</td>
<td>0.2205</td>
<td>-2.3859</td>
<td>1.3579</td>
</tr>
<tr>
<td>TNA family</td>
<td>16,006.3</td>
<td>50,991.9</td>
<td>0.1000</td>
<td>367,207.9</td>
</tr>
<tr>
<td>TNA</td>
<td>1,197.9</td>
<td>5,204.2</td>
<td>0.1000</td>
<td>95,976.9</td>
</tr>
<tr>
<td>Flows</td>
<td>-0.0044</td>
<td>0.0367</td>
<td>-0.2281</td>
<td>0.2718</td>
</tr>
<tr>
<td>S.D flows</td>
<td>0.0155</td>
<td>0.0238</td>
<td>0.0000</td>
<td>0.2485</td>
</tr>
</tbody>
</table>
### Table II

Posterior results for model 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1</th>
<th>Z-Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95%PI</td>
</tr>
<tr>
<td>Frontier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_0 ) (\text{beta}_{mkt})</td>
<td>0.1326</td>
<td>[0.1310, 0.1337]</td>
</tr>
<tr>
<td>( \gamma_1 (\text{beta}_{smb}) )</td>
<td>0.0393</td>
<td>[0.0372, 0.0413]</td>
</tr>
<tr>
<td>( \gamma_2 (\text{beta}_{hml}) )</td>
<td>0.0326</td>
<td>[0.0307, 0.0347]</td>
</tr>
<tr>
<td>( \gamma_3 (\text{beta}_{km}) )</td>
<td>0.0036</td>
<td>[0.0023, 0.0052]</td>
</tr>
<tr>
<td>Inefficiency</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \omega )</td>
<td>-0.4362</td>
<td>[-0.4382, -0.4345]</td>
</tr>
<tr>
<td>( \rho_1 )</td>
<td>0.6574</td>
<td>[0.5592, 0.7024]</td>
</tr>
<tr>
<td>( \delta_1 (\ln \text{tnafamily}) )</td>
<td>-0.0179</td>
<td>[-0.0183, -0.0174]</td>
</tr>
<tr>
<td>( \delta_2 (\ln \text{tna}) )</td>
<td>0.0195</td>
<td>[0.0189, 0.0204]</td>
</tr>
<tr>
<td>( \delta_3 (\ln \text{tna}^2) )</td>
<td>-0.0027</td>
<td>[-0.0028, -0.0026]</td>
</tr>
<tr>
<td>( \delta_4 (\text{flows}) )</td>
<td>0.3723</td>
<td>[0.3391, 0.3980]</td>
</tr>
<tr>
<td>( \delta_5 (\text{sd flows}) )</td>
<td>0.4093</td>
<td>[0.3897, 0.4320]</td>
</tr>
<tr>
<td>( \delta_6 (\text{Domestic}) )</td>
<td>-0.2666</td>
<td>[-0.2681, -0.2649]</td>
</tr>
<tr>
<td>( \delta_7 (\text{Ethical}) )</td>
<td>0.0388</td>
<td>[0.0337, 0.0454]</td>
</tr>
<tr>
<td>( \delta_8 (\text{Large}) )</td>
<td>0.1535</td>
<td>[0.1528, 0.1542]</td>
</tr>
<tr>
<td>( \delta_{10} (\text{Growth}) )</td>
<td>-0.1197</td>
<td>[-0.1211, -0.1179]</td>
</tr>
<tr>
<td>( \delta_{11} (\text{Institutional}) )</td>
<td>0.0255</td>
<td>[0.0234, 0.0281]</td>
</tr>
<tr>
<td>( \delta_{12} (\text{BankRetail}) )</td>
<td>0.0239</td>
<td>[0.0198, 0.0286]</td>
</tr>
<tr>
<td>( \delta_{13} (\text{Dealer}) )</td>
<td>0.0285</td>
<td>[0.0247, 0.0301]</td>
</tr>
<tr>
<td>( \delta_{14} (\text{Direct}) )</td>
<td>-0.0247</td>
<td>[-0.0260, -0.0236]</td>
</tr>
<tr>
<td>( \sigma^2_\theta )</td>
<td>0.0000</td>
<td>[0.0000, 0.0000]</td>
</tr>
<tr>
<td>( \sigma^2_\omega )</td>
<td>0.0148</td>
<td>[0.0145, 0.0152]</td>
</tr>
<tr>
<td>Efficiency</td>
<td>0.8594</td>
<td>[0.7921, 0.9326]</td>
</tr>
</tbody>
</table>

\( DIC \) = -605.28

\( LPS \) = -19.37

\(^1\) Corresponds to the average posterior distribution for all funds
Table III

Posterior results for Model 1 with interactions

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Model 1A</th>
<th>Model 1B</th>
<th>Model 1C</th>
<th>Model 1D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Domestic</td>
<td>Growth</td>
<td>Large</td>
<td>Ethical</td>
</tr>
<tr>
<td></td>
<td>Mean     95%PI</td>
<td>Mean     95%PI</td>
<td>Mean     95%PI</td>
<td>Mean     95%PI</td>
</tr>
<tr>
<td>🍊μ</td>
<td>0.1322 [0.1305,0.1339]</td>
<td>0.1320 [0.1301,0.1337]</td>
<td>0.1286 [0.1203,0.1365]</td>
<td>0.1038 [0.1017,1.1062]</td>
</tr>
<tr>
<td>🍊β1(β_mkt)</td>
<td>0.0405 [0.0385,0.0426]</td>
<td>0.0396 [0.0375,0.0413]</td>
<td>0.0403 [0.0385,0.0447]</td>
<td>0.0319 [0.0284,0.0347]</td>
</tr>
<tr>
<td>🍊β2(β_mkt)</td>
<td>0.0334 [0.0301,0.0367]</td>
<td>0.0327 [0.0314,0.0338]</td>
<td>0.0341 [0.0301,0.0379]</td>
<td>0.0339 [0.0306,0.0366]</td>
</tr>
<tr>
<td>🍊β3(β_mkt)</td>
<td>0.0033 [0.0017,0.0048]</td>
<td>0.0037 [0.0017,0.0056]</td>
<td>0.0037 [0.0017,0.0055]</td>
<td>0.0036 [0.0028,0.0044]</td>
</tr>
<tr>
<td>Efficiency</td>
<td>-0.5345 [-0.5849,-0.4809]</td>
<td>-0.4989 [-0.5476,-0.4519]</td>
<td>-0.3961 [-0.4501,-0.3595]</td>
<td>-0.2682 [-0.2698,-0.2671]</td>
</tr>
<tr>
<td>🍊θ1</td>
<td>0.0405 [0.0385,0.0426]</td>
<td>0.0396 [0.0375,0.0413]</td>
<td>0.0403 [0.0385,0.0447]</td>
<td>0.0319 [0.0284,0.0347]</td>
</tr>
<tr>
<td>🍊θ2</td>
<td>0.0334 [0.0301,0.0367]</td>
<td>0.0327 [0.0314,0.0338]</td>
<td>0.0341 [0.0301,0.0379]</td>
<td>0.0339 [0.0306,0.0366]</td>
</tr>
<tr>
<td>🍊θ3</td>
<td>0.0033 [0.0017,0.0048]</td>
<td>0.0037 [0.0017,0.0056]</td>
<td>0.0037 [0.0017,0.0055]</td>
<td>0.0036 [0.0028,0.0044]</td>
</tr>
<tr>
<td>🍊σ (flows)</td>
<td>0.2931 [0.2150,0.3729]</td>
<td>0.2963 [0.2182,0.3796]</td>
<td>0.4787 [0.4030,0.5425]</td>
<td>0.6405 [0.5162,0.6601]</td>
</tr>
<tr>
<td>🍊σ (sd flows)</td>
<td>0.7462 [0.6468,0.8391]</td>
<td>0.5553 [0.3902,0.3916]</td>
<td>0.2507 [0.1920,0.3294]</td>
<td>0.1959 [0.1356,0.1808]</td>
</tr>
<tr>
<td>🍊σ (Domestic)</td>
<td>-0.2455 [-0.2725,-0.2122]</td>
<td>-0.2502 [-0.2714,-0.2385]</td>
<td>-0.2895 [-0.3261,-0.2410]</td>
<td>-0.4476 [-0.4499,-0.4468]</td>
</tr>
<tr>
<td>🍊σ (Growth)</td>
<td>0.0349 [0.0204,0.0474]</td>
<td>0.0231 [0.0204,0.0263]</td>
<td>0.0357 [0.0199,0.0502]</td>
<td>-0.2447 [-0.2576,-0.2296]</td>
</tr>
<tr>
<td>🍊σ (Large)</td>
<td>0.1158 [0.1012,0.1397]</td>
<td>0.1369 [0.1283,0.1521]</td>
<td>0.3118 [0.2673,0.3701]</td>
<td>0.0743 [0.0718,0.0761]</td>
</tr>
<tr>
<td>🍊σ (Ethical)</td>
<td>0.0159 [0.0151,0.0167]</td>
<td>-0.0460 [-0.0649,-0.0301]</td>
<td>-0.0293 [-0.0319,-0.0264]</td>
<td>-0.0401 [-0.0477,-0.0333]</td>
</tr>
<tr>
<td>🍊σ (Institutional)</td>
<td>0.0255 [-0.0247,-0.0262]</td>
<td>0.0252 [-0.0247,-0.0262]</td>
<td>0.0205 [0.0200,0.0209]</td>
<td>0.0249 [0.0301,0.0534]</td>
</tr>
<tr>
<td>🍊σ (BankRetail)</td>
<td>0.0293 [0.0278,0.0326]</td>
<td>0.0297 [0.0282,0.0312]</td>
<td>0.0294 [0.0283,0.0306]</td>
<td>0.0403 [0.0407,0.0454]</td>
</tr>
<tr>
<td>🍊σ (Dealer)</td>
<td>0.0235 [0.0201,0.0263]</td>
<td>0.0253 [0.0201,0.0265]</td>
<td>0.0249 [0.0200,0.0257]</td>
<td>0.0579 [0.0314,0.0722]</td>
</tr>
<tr>
<td>🍊σ (Direct)</td>
<td>0.0111 [0.0107,0.0115]</td>
<td>0.0111 [0.0107,0.0115]</td>
<td>0.0111 [0.0107,0.0115]</td>
<td>0.0111 [0.0107,0.0115]</td>
</tr>
</tbody>
</table>

1 Corresponds to the average posterior distribution for all funds

DIC 3 -1154.82 -1153.11 -1150.72 -1150.63
LPS -109.21 -109.08 -108.25 -108.10
Figure 1. Densities of the posterior inefficiency persistence by fund style

Figure 2. Inefficiency persistence versus Efficiency in 2003 and 2013
Figure 3. Evolution of efficiency over time by fund style