On the Comparative Performance of Socially Responsible and Islamic Mutual Funds[*]

Omneya Abdelsalam

Aston University

Meryem Duygun Fethi

University of Leicester

Juan Carlos Matallín

Universitat Jaume I

Emili Tortosa-Ausina

Universitat Jaume I and Ivie

May 20, 2013

Abstract

This is the first study to provide comprehensive analyses of the relative performance of both Socially Responsible Investment (SRI) and Islamic mutual funds. The analysis proceeds in two stages. In the first, the performance of the two categories of funds is measured using partial frontier methods. In the second stage, we use quantile regression techniques. By combining two variants of the Free Disposal Hull (FDH) methods (order-m and order-α) in the first stage of analysis and quantile regression in the second stage, we provide detailed analyses of the impact of different covariates across methods and across different quantiles. In spite of the differences in the screening criteria and portfolio management of both types of funds, variation in the performance is only found for some of the quantiles of the conditional distribution of mutual fund performance. We established that for the most inefficient funds the superior performance of SRI funds is significant. In contrast, for the best mutual funds this evidence vanished and even Islamic funds perform better than SRI. These results show the benefits of performing the analysis using quantile regression.

Keywords: socially responsible funds, Islamic funds, performance, quantile regression

JEL Classification: G11, G23

Communications to: Emili Tortosa-Ausina, Departament d'Economia, Universitat Jaume I, Campus del Riu Sec, 12071 Castelló de la Plana, Spain. Tel.: +34 964387168, fax: +34 964728591, e-mail: tortosa@uji.es.

---

1. Introduction

Over the past few years, the growth of some types of mutual funds has been remarkable. In particular among these, one may highlight the socially responsible investment (SRI) mutual funds which have paralleled the growth in the business ethics literature. SRI is thus labelled for characteristics that are not exclusive to this type of fund; for example, Islamic mutual funds also have ethical restrictions on investments.

SRI mutual funds have demonstrated an important growth pattern since the beginning of the Seventies, when the first one was created.[1] The ethical screening of corporate behaviour has become fashionable in the wake of recent reports of environmental and accounting scandals. However, it has also been argued that imposing ethical constraints on equity investment can adversely affect portfolio performance (Bauer et al., 2007). The SRI mutual funds industry refers to the practice of directing investment funds using techniques that combine investors' financial objectives with their commitments to social concerns; for example, social justice, economic development, peace or a healthy environment (Hiagh and Hazelton, 2004).

Paralleling the trend for SRI funds is the large-scale growth in Islamic finance over the last twenty years. This growth has been influenced by the global integration of financial markets and the introduction of innovative Islamic finance products. The Islamic mutual funds market is one of the fastest growing sectors within the Islamic finance industry. Islamic mutual funds embody a unique "filtration" process to select appropriate shares. The process ensures that the mode of operation and capital structure of each business which the funds invested in are Shariah[2] compliant. In addition, invested funds must be free of Riba (usury) or Gharar (speculation). Although both SRI and Islamic investment instruments employ screening strategies in their investment process, there are distinct differences between them (Forte and Miglietta, 2007). Table 1 illustrates the differences between both types of funds.

Addressing the differences between SRI and Islamic mutual funds, when selecting investments for their portfolio (asset allocation), SRI mutual funds can freely choose between debt-bearing investments and equity-bearing investments, as long as the stocks selected strictly adhere to social, moral or environmental beliefs. A large part of investment from SRI funds is directed toward small companies with lower dividend yields (Fowler and Hope, 2007). On the opposite side, Islamic mutual funds undergo a more rigorous screening process in an attempt to select portfolios that meet both qualitative and quantitative criteria set by Shariah guidelines. The filtration process eliminates companies engaged in prohibited activities under Shariah, such as companies dealing in tobacco, alcohol, weapons, biotechnology for human cloning, arms, etc. and companies whose capital structures rely heavily on debt financing to avoid dealing with interest. Islamic mutual funds exclude investments in fixed income instruments, such as corporate bonds, certificates of deposit (CDs), preferred stocks, warrants and some derivatives (e.g. options). Furthermore, Shariah provides structural guidance regarding the governance and management structure of Islamic mutual funds. Hence, it is plausible to identify Islamic mutual funds as reflecting wider pools of investment that involve ethical and socially responsible mutual funds. Examining SRI mutual funds provides the closest alternative for evaluating the performance of Islamic mutual funds. However, we expect that the above identifiable differences are likely to result in significant discrepancies in both investment behaviour

---

[1] According to the International Investment Funds Association (IIFA), `http://www.iifa.ca/`.
[2] Islamic Law.

and fund performance.

In spite of the astounding growth of Islamic finance there are no studies that have compared Islamic mutual funds with SRI. The main objective of the paper is to explore the impact of different portfolio restrictions, expenses and value added criteria on the performance of both types of funds. In order to undertake this comparison we proceed in two stages. In the first stage, we measure the performance of both SRI and Islamic mutual funds using nonparametric frontier methods. These methods originated in an attempt to evaluate the efficiency of DMUs (Decision Making Units) that produce multiple outputs with multiple inputs. The most popular among these methods is Data Envelopment Analysis (see Charnes et al., 1978). As indicated by Glawischnig and Sommersguter-Reichmann (2010a), these are increasingly used for the performance assessment of capital markets, traditional and alternative investment (particularly hedge funds) and managed future funds, but DEA estimators are sensitive to extremes and/or outliers, which can undermine the estimation of performance. Two main methods have been proposed in the literature, namely order-m (Cazals et al., 2002) and order-α (Daouia and Simar, 2007). With both, the underlying idea is to estimate a partial frontier inside the cloud of data points but close to the lower (or upper) boundary. This boundary is simultaneously sensitive to the magnitude of extreme valuable observations as well as robust to their influence (Daouia et al., 2012). Despite their advantages, neither order-m nor order-α estimators have been considered when evaluating the performance of either Islamic or SRI funds.

Our results illustrate how the average efficiency of socially responsible SRI funds is slightly higher than that of Islamic funds. However, in general the differences are not statistically significant. In this regard, the second stage of the analysis provides a more insightful view, as it takes into account the fact that, if performance differences existed between SRI and Islamic funds, they could partly depend on the degree of efficiency of the funds—i.e. the relevance of the differences could vary for the best and worst performers. Once the performance of each type of fund has been measured, significant differences can be assessed in a number of ways.

In the second stage we employ a regression procedure to analyse the impact of other covariates on the performance of mutual funds. Some researchers Simar and Wilson (2011) expressed concerns about combining certain regression procedures (Ordinary Least Squares or Tobit) in the second stage when analysing determinants of efficiencies obtained in the first stage using nonparametric frontier techniques (such as DEA or its non-convex counterpart, Free Disposal Hull). Accordingly, we use quantile regression for a variety of reasons. First, it is more robust to departures from normality. Second, it provides information on differential impacts across the conditional distribution efficiencies—i.e. conclusions are not confined to analysing the average impact of a given covariate on mutual funds' performance. Finally, we contribute to the literature by combining order-m and order-α methods in the first stage, with quantile regression in the second stage. This procedure facilitates analyses of not only the impact of different covariates across methods (order-m and order-α) but also across different quantiles.

Indeed, we find evidence of how the results vary depending on the conditional distribution of efficiencies. In the set of covariates which might affect the funds' efficiencies, apart from their investment objective (SRI or Islamic), we also considered their geographical focus and their age. For the best mutual funds these determinants show very little explanatory ability and, therefore, the value added by managers would be the relevant factor. However, the analysis illustrates that for the worst funds the analyzed determinants are actually relevant. Specifically, only for the most inefficient funds is the superior performance of SRI funds found to be significant. In other cases there is no difference and, for some parts of the distribution

(especially for the best mutual funds), Islamic funds perform better. In total, we find an asymmetry regarding the determinants of efficiency for the best and for the worst funds. In our opinion, these findings would represent an interesting contribution that opens opportunity for further research in this area given the non-conclusive literature on the potential differences between conventional and SRI funds.

The paper is organized as follows: in section 2 we briefly review the literature. Section 3 presents the methodology used to assess the financial performance of mutual funds and to evaluate their comparative performance. In section 4 we describe the sample data employed for the empirical analysis and this is discussed further in section 5. Finally, section 6 highlights the most important conclusions.

2. Background

2.1. Socially Responsible Investment (Ethical Mutual Funds)

Prior literature suggests that investors are attracted to SRI vehicles from a desire to match their investment policies with their values. SRI involves a fund implementing "ethical screens" to ensure that it does not invest in firms that have a poor record in CSR (Brammer et al., 2006). There are several options for ethical investment. First: community or cause-related investment reflecting savings accounts held at socially directed organizations. Second: purchasing units in socially responsible mutual funds. Third: direct investment in corporations in order to engage in dialogue with the purpose of changing the company ethics (Hiagh and Hazelton, 2004).

The earliest SRI was employed by religious groups such as the Quakers and the Lutheran Brotherhood which exclude "sin industries" such as alcohol and tobacco (Schepers, 2003). The first SRI mutual fund, the PAX world fund, is a fund launched in 1971 that eschewed investments in military-related stocks during the Vietnam War (Fowler and Hope, 2007). A number of other funds have been initiated with similar peace-driven or other social goals. According to Social Investment Forum (2010), professionally managed assets following SRI strategies in the U.S. reached $3.07 trillion at the start of 2010. Ethical fund managers make two claims. First, that the SRI fund will influence companies to change their operations and, second, that the financial returns of SRI funds are not different from those of conventional investments in the short-term (Hiagh and Hazelton, 2004).

Prior studies (see Rudd, 1981; Grossman and Sharpe, 1986; Diltz, 1995) document that ethical and moral screening of companies affects several aspects of the fund, namely: (a) the characteristics of assets included in the mutual funds' portfolio; (b) portfolio diversification; and (c) portfolio performance. This is likely because the screening process imposes an additional set of constraints to the wealth-maximizing investor. Rudd (1981) suggests that whenever the portfolio is constrained its performance becomes weaker. He documents that socially responsible investing introduces size and other biases into the portfolio, with a consequential deterioration in long-run performance. Controversially, Statman (2000), Sauer (1997) and Goldreyer et al. (1999) all document that the application of social screens does not have a significant effect on investment performance.

Perks et al. (1992) argue that if "good" social performance is associated with poor financial performance, investors would be less likely to be inclined to invest in such socially responsible companies. Luther et al. (1992) provide weak evidence that in the UK ethical funds outperform two market indexes. They established that SRI funds perform much better when evaluated against a small company benchmark than when the Financial Times All Share index (FTSA) is used. Hamilton et al. (1993) compare the performance

3

of a sample of 32 American SRI funds to that of 170 ordinary funds for the sample period 1981–1990. Average returns for SRI funds are shown to be higher than the average returns for the ordinary funds. Their results suggest that those investors do not lose by investing in similar ethical funds. Mallin et al. (1995) use a matched pair analysis to explore returns earned by 29 UK ethical funds and 29 UK non-ethical funds, matched on the basis of age and size, from 1986 to 1993. Although ethical funds tend to outperform their matched non-ethical pairs the evidence is weak. Gregory et al. (1997) also adopt a matched pair approach, with a size adjusted measure of performance: no significant difference is determined in returns. Sauer (1997) compares the performance of the Domini Social Index (DSI 400) with that of the S&P 500 and the Center for Research in Security Prices (CRSP) value-weighted index, for the period of 1986 to 1994: the application of social responsibility screens neither increases nor decreases investment performance relative to either the S&P 500 or the CRSP value-weighted index. This study also establishes that SRI does not result in greater risk to the investor, measured as the variability of monthly returns. Goldreyer et al. (1999) compare 49 socially responsible mutual funds with a random sample of conventional funds between 1981 and 1997, and find no significant differences. Guerard (1997b) reports relatively small significant differences between the performances of socially screened and unscreened investments. An extended study by Guerard (1997a) identifies that investment screens which preclude environmental or alcohol/tobacco/gambling or nuclear stocks, yield higher average returns than unscreened investments.

To summarize, the above discussed studies show that the literature that compared conventional funds and SRI is inconclusive. In general, there is either no significant difference between conventional and SRI funds and the latter does not outperform the former.

2.2. Islamic Mutual Funds

Islamic mutual funds represent one of the fastest-growing sectors within the Islamic Finance industry, with estimated Islamic global financial assets under management of $1,033 billion (Ernst & Young, 2011). Islamic equity investment is a new phenomenon, that dates back to 1994 when a new ruling[3] was issued, allowing Muslim investors to trade in international stocks within certain parameters (Hayat and Kraeussl, 2011). Since then, the Islamic investment industry has witnessed a tremendous increase; many mainstream international players have entered the market of Islamic investing. Dow Jones, FTSE, MSCI and S&P offer hundreds of Islamic equity indices; and more than 800 managed Islamic Mutual funds were reported by the end of 2010 (Ernst & Young, 2011).

Islamic Finance is based on five main pillars, which include the prohibition of usury ($Riba$)[4]; speculation ($Maysir$); excessive uncertainty ($Gharar$); and investing in impermissible activities. The fifth pillar is encouraging sharing of risks and returns (Shanmugam and Zahari, 2009; Hayat and Kraeussl, 2011). In Islam, it is not deemed fair to apply a fixed interest rate on investment loans as it is considered discriminatory. The entrepreneur (borrower) alone has to carry the total risk, while the investor (lender) may receive the fixed income no matter the failure or success of the enterprise. In contrast, when profit is especially high, the entrepreneur will receive the major portion of the profit, while the investor's (lender's) share is minimal. This means that an uneven distribution of profit and risk has been applied (Novethic, 2009). Adherence to these principles is monitored by a Shariah Supervisory Board (SSB); a panel of

---

[3]The decree issued by International Fiqh Academy relaxed the Shariah constraints on interest-based activities for companies whose main industry sector is permissible (Halal) (Hayat and Kraeussl, 2011).

[4]$Riba$ is prohibited in "all monotheist religions" (Hossain, 2009).

Shariah scholars, which monitors the companies' or funds' compliance with Shariah guidelines.

Islamic mutual funds undergo a rigorous screening process in an attempt to select portfolios that meet both qualitative and quantitative criteria set by Shariah guidelines. Qualitative screens are used to filter out companies based on the nature of their business, securities that contain one of the Shariah-prohibited elements, or companies that conduct unethical business practices, as per Shariah. After applying qualitative filters, all remaining stocks have to be analysed on the basis of quantitative screening relevant to debt, interest-bearing securities, receivables and cash. However, there is no consensus regarding the criteria currently used to screen Shariah compliant stocks (Hayat and Kraeussl, 2011; Derigs and Marzban, 2008).[5] The Islamic investment screening procedure, astonishingly enough, does not incorporate concerns of an environmental or social nature. Unlike SRI, this screening mechanism pays no reward to positive screening criteria, e.g. protection of the environment, human rights, and investment in the community. This is a clear flouting of the intrinsic Shariah values of equity, justice and fairness. In this regard the current Islamic screening criteria are somewhat contradictory. Furthermore, it was expected that Sharia supervisory boards would promote socially responsible investment practices, and would take proactive roles in influencing companies to adopt them (Novethic, 2009; Binmahfouz, 2012).

As with SRI, extra non-financial screening standards are applied to the selection of Shariah compliant investments, during the course of which some companies and sectors are removed, regardless of the return profile or risk connected with the investments that have been excluded, simply because they do not comply with Shariah. This contradicts the theory on modern portfolios, in which rational investors selected portfolios that minimised risks while maximising returns, thus gaining the maximum expected utility. It can be argued that conventional mutual funds would probably perform better than Islamic mutual funds, which are restricted; the conventional mutual funds have the benefit of free selection of stocks and free management of their investment portfolios (Binmahfouz, 2012). Researchers are continually analysing investment characteristics of portfolios of Islamic funds that are restricted, in order to clarify their differences in performance to counterparts which are not restricted. The majority of empirical studies (Ahmad and Ibrahim, 2002; Girard and Hassan, 2008; Hashim, 2008; Albaity and Ahmad, 2008; Dharani and Natarajan, 2011; Mansor and Bhatti, 2011; BinMahfouz and Hassan, 2012) show that there is no significant difference between the performance of Islamic portfolios (restricted) and their non-restricted counterparts. In fact, some studies (e.g. Hussein and Omran, 2005) produce results of Islamic indices having outperformed their traditional counterparts.

2.3. Differences between Islamic and SRI funds

The Islamic mutual funds market is one of the fastest growing sectors within the Islamic finance industry. Even so, when compared to SRI mutual funds, Islamic mutual funds are still in the infancy stage of development. Addressing the differences between SRI and Islamic mutual funds (see Table 1), when selecting investments for their portfolio (asset allocation), SRI mutual funds can freely choose between

---

[5]Derigs and Marzban's (2008) study has shown that there is no generally accepted understanding of how to transform the prescriptive Shariah rules into a system of checkable investment guidelines. The study has also shown that applying the alternative Shariah screens, used by the most prominent Islamic Shariah-compliant funds and indexes, to a common standard asset universe (S&P 500 index), has resulted in significant differences in both size and constituents of the consequential portfolios of halal assets. In order to solve the inconsistency it is imperative, therefore for a Shariah authority of higher Islamic order, to be established, in order to resolve the unreliability of the Shariah screening procedures. This body should be set up at national level, at the least, providing screening standards that comply with Shariah guidelines for the entire country's industry. It is hoped that this will obviate conflicting interest which arises when Islamic mutual funds/indices designate their own Shariah board Binmahfouz (2012).

debt-bearing investments and profit-bearing investments, as long as the stocks chosen strictly adhere to social, moral, or environmental beliefs. This requires companies to undertake a careful screening process in order to maintain stocks that coincide with the fund's beliefs. By construction, a large part of investment from these funds is directed toward small companies with lower dividend yields (Fowler and Hope, 2007). On the opposite side, Islamic mutual funds undergo a more rigorous screen process in an attempt to select portfolios that meet both qualitative and quantitative criteria set by Shariah guidelines. The filtration process excludes companies whose business is banned under Shariah such as alcohol or companies that conduct unethical business practices, for example those engaged in biotechnology for human cloning. Islamic mutual funds also exclude investments in fixed income instruments such as corporate bonds, certificates of deposit (CDs), preferred stocks, warrants, and some derivatives (e.g., options). Even though both SRI and Islamic investment instruments employ screening strategies in their investment processes, there are distinct differences between them (Forte and Miglietta, 2007), as shown in Table 1. Islamic mutual funds are required to reflect permissible elements under the Islamic law. Invested funds must be free of interest based debt[6] or speculation. Furthermore, Shariah provides a structural guidance regarding the governance and management structure of Islamic mutual funds. Hence, it is plausible to identify Islamic mutual funds as reflecting wider pools of investments that impound socially responsible mutual funds. Examining SRI mutual funds provides the closest alternative to evaluate the performance of Islamic mutual funds. However, we expect that the above identifiable differences are likely to result in significant discrepancies in both the investment behaviour as well as the fund performance.

3. Methods

3.1. Mutual fund evaluation using frontier techniques

A growing literature focuses on the evaluation of mutual fund performance using frontier techniques. Most of this literature has been reviewed in the paper by Glawischnig and Sommersguter-Reichmann (2010a). Within this paper, the authors provide a careful assessment of the literature evaluating the performance of traditional and alternative investment funds using nonparametric methods such as DEA. They implicitly admit that, in contrast, the number of studies applying FDH to mutual fund evaluation is virtually non-existent, which might be surprising given how problematic it is to assume convexity unwarily—as DEA does, as opposed to FDH. The paper by Daraio and Simar (2006a), which is based on partial frontiers, is also mentioned in their survey, but only briefly do they indicate that Daraio and Simar's approach was "computationally demanding" (Glawischnig and Sommersguter-Reichmann, 2010a, p.297) and, therefore, one may easily infer that contributions analyzing the performance of mutual funds using partial frontiers (such as order-m and order-α) are still scant and, consequently, new contributions in this particular field would be welcome.

3.2. Using FDH for mutual fund evaluation

Both DEA and FDH share the same underpinnings and, therefore, their advantages compared with parametric methods are the same—i.e. there is no need to specify a functional form for the frontier, and no need to choose a distribution for the inefficiency. As fully nonparametric approaches, both DEA and

---

[6]Companies with a negligible level of debt financing (10% or less) may be selected, provided that the debt is not a permanent feature of the capital structure (Iqbal, 1997).

FDH require no information relating to production technology. This is an important advantage, as it is crucial information (it describes how inputs are transformed into outputs) that is frequently unavailable to researchers. Of course, if the model is wrongly specified it is likely to yield biased estimates. The main difference between FDH and DEA, as indicated by Tulkens (1993), is the convexity assumption of the attainable set of DEA, which does not exist in FDH. Therefore, if the validity of this assumption is put into question, then DEA may be statistically meaningless. Convexity is a strong assumption and it might be argued that choosing DEA can only be justified if assuming convexity is reasonable—i.e. the fact that comparisons can also be made against *virtual* funds built by combining existing funds.

One feature which makes both DEA and FDH attractive in the field of mutual fund performance evaluation is their ability to incorporate a number of different factors in a flexible way. Both approaches admit mutual fund performance indexes that can account for different risk measures and the transactions costs investments. The efficiency of a fund can be judged by the relative distance between the observed output and the efficient frontier. A fund is classified as inefficient if its output (i.e. returns) and input (i.e. risk) are below the best practice frontier. Although DEA and FDH share the same underpinnings, the removal of the convexity assumption on the attainable set under FDH suggests that this technique might be more appropriate when the convexity hypothesis is put into question. However both DEA and FDH methods share the same advantage compared with parametric methods, i.e. they are free from the "parametric straitjacket" and are relatively simple to compute. Thus, although most applications in the field have focused tightly on DEA, the greater flexibility of FDH, along with its asymptotic properties, are the reasons for our choice (Park et al., 2000).

Therefore, in the first stage of the estimation process, we evaluate the performance of mutual funds considering FDH—among other techniques which will be presented below. For this, it is necessary in the previous step to define the set of attainable combinations of inputs ($\mathbf{x}$) and outputs ($\mathbf{y}$), which determines the frontier of the possibility set.

To define the efficiency of a given fund, we use the distance between the observed values of fund variables and the frontier. The $\mathbf{\Psi}$ set of production possibilities will be the set of attainable points $(\mathbf{x}, \mathbf{y})$, defined as:

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

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

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

Its efficient boundary is then the subset of $Y(\mathbf{x})$, defined as

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

Considering expressions (1), (2) and (3), the Farrell (1957) measure of output-oriented efficiency of a given mutual fund $(\mathbf{x}, \mathbf{y})$ is defined as

$$\theta(\mathbf{x}, \mathbf{y}) = \sup\{\theta : (\mathbf{x}, \theta \mathbf{y}) \in \mathbf{\Psi}_{FDH}\} = \max\{\theta : \theta \mathbf{y} \in Y(\mathbf{x})\}, \tag{4}$$

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

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

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

and, following (Daraio and Simar, 2007, p.35), efficiencies (FDH estimators) can be estimated by solving the following integer linear program:

$$\hat{\theta}_{\mathrm{FDH}} = \max\{\theta|\mathbf{x} \geq \sum_{i=1}^{n} \gamma_i \mathbf{x}_i, \theta \mathbf{y} \leq \sum_{i=1}^{n} \gamma_i \mathbf{y}_i, \sum_{i=1}^{n} \gamma_i = 1, \gamma_i \geq 0, i = 1, \ldots, n\}. \tag{6}$$

3.3. Partial frontiers: order-m and order-α estimators

A robust alternative to DEA/FDH estimators are the partial boundaries of the production set (see Cazals et al., 2002). According to either DEA or FDH, the efficiency measure is obtained by comparison with the full frontier of all observations. According to the order-m estimators, which are based on FDH, what will actually be used as a benchmark is the *expected* maximum output achieved by any m funds chosen randomly from the population, which employs at most input level $\mathbf{x}$. When we choose a high value for m ($m \to \infty$), the order-m estimator results in the same benchmark as FDH, yielding the same results. Therefore, the most interesting cases will be those for which we define a finite value for m. In these cases the order-m does not envelop *all* the data, being more opposed to outliers in data.

The main advantage of using order-m instead of either DEA or FDH lies in their superior robustness to extremes and/or outliers in the output direction, which results in a much better estimation of the corresponding economic efficiencies. This is particularly important, as DEA and FDH are particularly sensitive to extremes and/or outliers, whether inputs or outputs. In actual fact, as indicated by Daouia et al. (2012), in most empirical examples, a naive application of DEA or FDH is impossible because real samples contain in general some anomalous data. Under these circumstances, the estimated frontier will be fully determined by these outliers, and the measurement of inefficiencies will be totally unrealistic (Daouia et al., 2012, p.286).

According to the order-m estimators the *expected* maximum output achieved by any m funds chosen randomly from the population, which employs at most input level $\mathbf{x}$ will be used as a benchmark (Pilyavsky and Staat, 2008). Therefore, for any $\mathbf{y}$, the *expected* maximum level is defined as:

$$\mathbf{y}^{\partial} = \widetilde{\theta}\mathbf{y}. \tag{7}$$

Therefore, the most interesting cases will be those for which we define a finite value for m. In these cases the order-m does not envelop *all* the data, being more robust to outliers in data. Note that the order-m efficiency scores are not bounded by 1 as it is the case under DEA or FDH. In these cases, values equal to unity correspond to *efficient* funds, whereas values higher than unity correspond to inefficient funds. According to order-m one may find values for θ lower than one, indicating that the fund operating

at the level $(\mathbf{x}, \mathbf{y})$ is more efficient than the average of m peers randomly drawn from the population of units using less inputs than $\mathbf{x}$.[7]

An estimator which shares some of the order-m underpinnings is the order-α quantile-type frontier. Whereas the m parameter of order-m serves as a trimming parameter which permits tuning of the percentage of points that will lie above the frontier, for order-α, the frontier is determined by fixing the probability $(1 - \alpha)$ of observing points above the order-α frontier. Therefore, with order-α we reverse the causality. These estimators were proposed for the multivariate case by Daouia and Simar (2007). Order-α estimators also have better properties than the usual nonparametric frontier estimators (DEA or FDH). In addition, it can be shown (see Daouia and Simar, 2007) that the order-α frontiers are more robust to extremes than the order-m frontiers (see Daraio and Simar, 2007, p.74).

The main virtue of order-α and order-m estimators lies with the fact the estimators do not envelop all the data, and thus are therefore relatively unaffected by outliers than FDH or DEA. These outliers which, in the particular output-oriented case we are dealing with will have an efficiency score below 1, will be considered as super-efficient with respect to the order-α frontier level. Therefore, analogously to order-m partial frontiers, where a mutual fund operating at $(\mathbf{x}, \mathbf{y})$ is benchmarked against the expected maximum output (recall we are in the output-oriented case) among m peers drawn randomly from the population of funds with output levels of at least $\mathbf{y}$. In the case of order-α quantile frontiers the benchmark is the output level not exceeded by $(1 - \alpha) \times 100\%$ of funds among the population of funds providing input levels of at least $\mathbf{x}$.

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

$$\theta_\alpha(\mathbf{x}, \mathbf{y}) = \sup\{\theta | F_{\mathbf{y}|\mathbf{x}}(\theta \mathbf{y}|\boldsymbol{x}) > 1 - \alpha\} \tag{9}$$

We will have that $\theta_\alpha(\mathbf{x}, \mathbf{y})$ converges to the FDH estimator $\theta(\mathbf{x}, \mathbf{y})$ when $\alpha \to 1$. As indicated in Daraio and Simar (2007), in cases where $\theta_\alpha(\boldsymbol{x}, \boldsymbol{y}) = 1$, the fund is "efficient" at the level $\alpha \times 100\%$, as it is dominated by mutual funds providing less input than $\boldsymbol{x}$ with probability $1 - \alpha$. In those cases where $\theta_\alpha(\boldsymbol{x}, \boldsymbol{y}) > 1$ then the unit $(\mathbf{x}, \mathbf{y})$ has to increase its output to the level $\theta_\alpha(\mathbf{y}, \mathbf{y})\mathbf{x}$ to achieve the output efficient frontier of level $\alpha \times 100\%$. We can also apply the plug-in principle to obtain an intuitive nonparametric estimator of $\theta_\alpha(\mathbf{x}, \mathbf{y}) = 1$ by replacing $F_{\mathbf{y}|\mathbf{x}}(\cdot|\cdot)$ with its empirical counterpart to obtain:

$$\widehat{\theta}_{\alpha,n}(\mathbf{x}, \mathbf{y}) = \sup\{\theta | \widehat{F}_{\mathbf{y}|\mathbf{x},n}(\theta \mathbf{y}|\mathbf{x}) > 1 - \alpha\} \tag{10}$$

Some recent contributions have been analyzing how the expected maximum frontier of order m and the conditional quantile-type frontier of order α are linked. Specifically, Daouia and Gijbels (2011) have

---

[7]Formally, the proposed algorithm (Cazals et al., 2002) to compute the order-m estimator takes the following steps, for n funds, $i = 1, \ldots, n$:

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

2. Obtain the efficiency measures, $\widetilde{\theta}_i$.

3. Repeat steps 1 and 2 B times and obtain B efficiency coefficients $\widetilde{\theta}_i^b (b = 1, 2, \ldots, B)$. The quality of the approximation can be tuned by increasing B, but in most applications $B = 200$ seems to be a reasonable choice.

4. Compute the empirical mean of B samples as:

$$\bar{\theta}_i^m = \frac{1}{B} \sum_{b=1}^{B} \widetilde{\theta}_i^b \tag{8}$$

9

compared both methods in terms of robustness, analyzing their breakdown values. They show that, even if the order-α quantile frontiers have global better robustness properties in terms of higher breakdown values, once they breakdown, they become less resistant to outliers than the order-m frontiers Daouia et al. (2012). Actually, Daouia and Gijbels (2011) show that the two classes of partial frontiers referred to in this section are closely related when $\alpha = \alpha(m) = (1/2)^{1/m}$, likewise the mean and median of a similar distribution do. This is an important conclusion as it may serve as a guide to answer the question of how to choose the order-α as a function of m for a possible comparison between order-α and order-m frontiers (Daouia and Gijbels, 2011, p.161) which is, precisely, one of our aims.

3.4. Analysing the determinants of mutual fund performance using regression quantiles

The central aim of classical regression is to estimate the mean of a response variable, conditional on the values of the explanatory variables. This approach may be valid in case regression assumptions are met. However, social phenomena are plagued with "nonstandard" conditions such as non-normality, or heteroskedastic data. For instance, non-financial data such as (conditional) income distribution are seldom normal. Financial or management data such as the dispersion of the annual compensation of chief executive officers tends to increase with firm size, an indicator of heteroskedasticity.[8]

In this type of scenarios quantile regression might be more appropriate, due to the relaxation of important assumptions for linear models such as those referred to above, among other reasons. In addition to this, it offers a much broader view than the conditional-mean framework, as it takes into account the effect of the explanatory variables on the (central and noncentral) location, scale, and shape of the distribution of the response variable. Specifically, typical linear models such as ordinary least squares, which are common across empirical studies, provide information on *average* impacts, i.e., the *average* impact for the *average* fund. Therefore, the researcher can only conclude whether this mean effect of a covariate on a response variable is significant or not. However, there is no reason why the impact should be the same over the entire conditional distribution.

Quantile regression makes it possible to approach these and related issues with some additional precision. Applications are growing in a variety of fields such as those referred to above, but they are still largely outnumbered by those using linear models which focus exclusively on average behaviour. Therefore, as its own name suggests, quantile regression does not confine the analysis to regression against averages (and hence it is not reliant upon its explanatory values) but rather also uses information obtained from the underlying distribution of the dependent variable. Thus its main advantage consists of being a method for estimating the conditional quantiles of a response variable distribution (in our case, the performance of mutual funds) in a linear model that provides a more complete view of possible causal relationships between variables in financial processes.

In our particular setting, the recent paper by Simar and Wilson (2011) provides a careful review of those contributions which have proposed different techniques to deal with the issue of analyzing the determinants of efficiency and productivity obtained using frontier techniques. However, the literature has evolved both in the parametric and nonparametric fields, among which Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) are, by and large, the most applied methods, Simar and

---

[8]Other examples include changes in wage distribution (Machado and Mata, 2005), educational attainment and wage distribution (Lemieux, 2006), innovation and firm growth in high-tech sectors Coad and Rao (2008), or location patterns of bank branches (Alamá and Tortosa-Ausina, 2012), among many other applications.

Wilson (2011) focus on nonparametric techniques only, which is also the family of techniques we employ.

Some years earlier, the same authors provided a heavily-cited contribution (Simar and Wilson, 2007), also reviewed in Simar and Wilson (2011), in which a bootstrap method was proposed to circumvent those problems inherent to studies mixing linear programming methods in the first stage (when efficiency scores are estimated), with parametric methods in the second stage (when the *determinants* of efficiency are explored). Whereas in the first stage most studies used DEA (much fewer employed FDH), in the second stage most studies considered either OLS or censored dependent variable methods such as Tobit. As summarized in Simar and Wilson (2011), this literature kept on expanding to include Banker and Natarajan (2008), Hoff (2007), McDonald (2009), or Ramalho et al. (2010), among others. Some alternative contributions following very different approaches and not included in Simar and Wilson's (2011) review are Tortosa-Ausina (2004), Balaguer-Coll et al. (2007), or Illueca et al. (2009). In the particular case of mutual fund performance evaluation, Daraio and Simar (2005, 2006b) also propose different nonparametric methods to overcome the problems derived from estimating regressions where the dependent variable is obtained by solving linear programming problems [9]

In this setting, the great advantage of quantile regression is that it enables the consideration the entire distribution of mutual funds' performances, either for a particular type of mutual fund or for several types of mutual funds separately. As indicated by Coad and Hölzl (2009), quantile regression is able to provide a more *complete* story of the relationship between variables—in our case, between mutual funds' performance and the relevant covariates. We can therefore specifically investigate whether, for those mutual funds whose performance is low (corresponding to the lower tail of the distribution, or to the lower quantiles), the sign and significance of the determinants is the same as for funds whose performance is high (those lying in the higher tails of the distribution, and corresponding to the highest quantiles). Under such circumstances, we would have some more detailed information for disentangling what it is that makes the performance of mutual funds to differ. Therefore, in the present study we consider that both high- and low-performance funds are of interest in their own right—we do not want them to be considered as outliers—and quantile regression allows us to analyze them in greater detail.

The main advantage of using quantile regression in the context of evaluating the determinants of mutual funds' efficiency is that, as indicated by Coad and Hölzl (2009), we know that the standard least-squares assumption of normally distributed errors does not hold for our data because the location patterns follow a fat-tailed distribution. However, while standard regression estimators are not robust to departures from normality, the quantile regression estimator is characteristically robust to outliers on the dependent variable. In addition, quantile regression also relaxes the restrictive assumption that error terms are identically distributed at all points of the conditional distribution and, avoiding this assumption will enable us to analyze the differences in the relationship between the endogenous and exogenous variables at different points of the conditional distribution of the dependent variable—mutual funds' performance.

The regression quantiles we use in this section specify the τth quantile of the conditional distribution

---

[9]Simar and Wilson revise very carefully the literature that has combined Data Envelopment Analysis (DEA) and related methods with either OLS or logistic regression in the second approach. However, in their extensive review, they did not mention the literature combining DEA with other types of regression based on more flexible methods such as nonparametric regression, conditional density estimation or, as we do here, quantile regression, which is much less abundant—in the case of quantile regression, it is entirely yet to come. Therefore, from a theoretical point of view, it remains an area to be explored. However, some examples have already been published in the literature which combine DEA and the likes (such as FDH, order-m and order-α) with these more flexible methods, including Balaguer-Coll et al. (2007), in the case of nonparametric regression, Illueca et al. (2009), in the case of conditional density estimation and, to a lesser extent Kumar and Russell (2002).

of y_i, where y_i is the variable containing the performance of mutual funds, given $\boldsymbol{x}$ as a linear function of the covariates. As described in detail by Koenker (2005), the estimation is carried out by minimizing the following equation:

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^k}{\text{Min}} \sum_{i \in \{i: y_i \geq \boldsymbol{x}'\boldsymbol{\beta}\}} \tau |y_i - \boldsymbol{x}'\boldsymbol{\beta}| + \sum_{i \in \{i: y_i < \boldsymbol{x}'\boldsymbol{\beta}\}} (1-\tau)|y_i - \boldsymbol{x}'\boldsymbol{\beta}| \tag{11}$$

where k is the number of explanatory variables, and τ represents the vector containing each quantile. The vector of coefficients $\boldsymbol{\beta}$ to be estimated will differ depending on the particular quantile.

Using regression quantiles has certain analogies with approaches such as testing the assumption of parameter's homogeneity across geographic units in measuring risk sharing. In our case, running the regressions for different quantiles (τ's) is equivalent, since the different values of τ refer to different values of the endogenous variable (the performance of mutual funds), which varies remarkable across the conditional distribution of mutual funds. Some previous contributions have also been considering the use of quantile regression in finance, although the number of contributions in this particular field is low. For instance, Meligkotsidou et al. (2009) use quantile regression to analyze how the links between hedge fund returns and risk factors change across the distribution of conditional returns. In an application to asset allocation in the popular 401(k) plans, Huberman and Jiang (2006) examine the sensitivity of choices made (the number of funds in which a participant chooses to invest all of his balance) against choices offered (the number of fund options available to employees of the plan) at different conditional quantiles. Other authors such as Chen and Huang (2012) pursue objectives which are more similar to ours, as they examine the differential impact of mutual fund governance variables (such as manager incentives and board quality) on fund performance across the entire performance distribution.

4. Data and performance measurement

4.1. Data sources

The empirical analysis is carried out for a sample of SRI and Islamic equity mutual funds in the world. Our data source is Lipper, and the first data record is from 2001. Figure 1 displays the number of mutual funds from January, 1989, until March, 2011, whereas Figure 2 displays the same information with the type of funds in different axes. Both types of funds show a similar pattern, characterized for an important increase during the "gold decade", i.e. the period 1995–2005, and a slight reduction during the following years. Taking this data into account, many funds were born but others did not survive. If only mutual funds with complete data were selected, it could yield the so-called *survivorship bias*. This bias may be important for performance analysis, because it could be the case that the non-survivor funds were just the worst performers and then, in the aggregate, results would be positively biased. To avoid this bias, the literature proposes incorporating all funds: new, survivors and non-survivors, which is precisely what we do.

However if the objective were to compare the performance between mutual funds, the inclusion of funds with limited data has several effects: (i) the low number of observations affects the robustness of the performance; (ii) there might be a bias if mutual funds' performance were correlated with the period for which data was available (for instance, the performance could be different at bullish or bearish markets); and (iii) in consequence, comparing funds with different living periods could result into some noise on the estimates. Therefore, we selected a period and sample of mutual funds in an attempt to reach a balance

between both biases. Thus, the sample period is confined to the period between December, 2000, and March, 2011, and has required that the data for any fund in the sample will cover at least 50% of the 123 months of that period. Following these criteria, 9.8% of Islamic funds has to be removed and, therefore, the initial number of 153 Islamic funds is reduced to 138. For SRI funds, 22% of the funds are removed, from 815 to 636. For the final sample, monthly returns were calculated as the variation relative to the monthly net asset value (NAV). Table 2 reports some basic statistics about the mutual funds sample.

The funds' return is the result of *passive* and *active* management. The return linked to active management is the value added by managers above the return from passive management. The return of passive management hinges critically on the fund's investment objectives. For this reason mutual funds are grouped by the *geographical area* of their investments, from Lipper data. East refers to the Pacific and Asia-Pacific area, Middle refers to Middle East, North Africa and GCC countries,[10] West refers to Europe and America, and global does not have any particular area of reference.

Table 2 shows the number of funds distributed to various areas, highlighting the absence of SRI funds, unlike the Islamic, which have as their investment target area the Middle East. Comparing SRI and Islamic funds, the averages of the percentage of data over the period sample and age are similar, something slightly more for SRI funds. For all area-objectives of funds, the average expenses of Islamic funds are higher than for SRI funds. The average of the mean of the gross returns both type of funds are quite similar for West mutual funds, around 8% annually. However those for East and Global areas are higher for SRI mutual funds. Since the expenses of SRI funds are less than those of the Islamic, the difference between both is somewhat greater if comparing net returns obtained from NAVs. For net returns, the Table 2 also shows the median of the cross-sectional distribution of the means of the returns of any fund in each area-objective. In some cases the differences are driven by mutual funds on the tails of the distribution of returns. Since the data have a risk component, Table 2 also shows the average and median of the distribution of the standard deviation for the returns of the funds. The levels of risk are similar for both types of funds. The last columns show the average of the skewness and kurtosis of the distributions of the monthly net returns of the funds in each area-objective.

Comparing the returns of all the funds, the medians of the distribution of the means of the net returns for Islamic and SRI funds are 4.95% and 6.29%, respectively. In contrast, the median of the distribution of the standard deviation is *slightly* higher for SRI funds, 20.32%, with respect to 19.92% for Islamic. Then for the 1.34% of difference, 0.45% is due to expenses (1.53% minus 1.08%) and the rest, 0.89%, is largely due to active management and the differing levels of risk, or area-objective investment, can attribute it to passive management. However, although the descriptive analysis prior drive us to approximate the behavior of the funds, it is necessary to rigorously analyze the efficiency of funds. So, the next section applies relatively new methodologies for measuring efficiency and establishing an appropriate comparison between SRI and Islamic funds.

4.2. Inputs and outputs selection

One of the main advantages of using frontier techniques to evaluate the performance of mutual funds is their ability for handling multiple inputs and outputs in the model. This has been stressed, among others, by Basso and Funari (2001), who claimed that "DEA approach allows defining mutual fund performance

---

[10]Gulf Cooperation Council countries, which include Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates.

indexes that can take into account several inputs and thus consider different risk measures (standard deviation, standard-semi deviation and beta) and redemption cost." Therefore, not only DEA but also FDH or the partial frontiers order-m and order-α approaches may include additional outputs apart from the traditional mean return measure. These would include, for instance, expected return or the expected excess return, among others. In the case of the inputs, the choice can also go beyond standard deviation and take into account other dimensions as well.

For instance, in one of the seminal papers on the topic, Murthi et al. (1997) considered computing their portfolio efficiency index with standard deviation of returns, expense ratio, loads and turnover as inputs, and mean gross return as output. Others such as, for instance, Choi and Murthi (2001) considered a similar choice, yet adopted a different DEA model. Other examples include Wilkens and Zhu (2001), who included in their study standard deviation and the percentage of periods with negative returns as inputs, whereas their choice of outputs consisted of mean return, minimum return and skewness. Chang (2004) made a proposal based on a non-standard DEA model which also required the selection of several inputs and outputs which in this case corresponded to the standard deviation, β, total assets and loads as inputs, and the traditional mean return as outputs. The number of studies using these types of approaches has flourished in relatively recent times, as indicated by Glawischnig and Sommersguter-Reichmann (2010b) and, although being able to specify several inputs and outputs is a plus, it comes with the disadvantage of having to make the selection carefully. In this line, Nguyen-Thi-Thanh (2006) considers that some investors might be more concerned with central tendencies such as mean or standard deviation whereas others might be more worried about extreme values (skewness, kurtosis).

To apply our methodological approach we must thereby define some variables as inputs and outputs. Our choice will be partly driven by the available information. Investors will usually consider as outputs those characteristics they would wish to maximize. As a main output we consider the daily mean return over the sample period (gross return, y_1). Other outputs such as skewness (y_2) have also been computed from the daily returns distribution. Regarding the inputs, the investment fund industry usually considers as inputs any characteristic that the investor is interested in minimizing. We will consider as such the risk of the fund, measured by the standard deviation of the daily returns (x_1), as well as kurtosis (x_2), also computed from the daily returns. Investors are known to dislike fat tails and, therefore, inclusion of kurtosis as an input would be reasonable. In some of the proposed models the management costs of the fund are also considered as input. In order to include these types of costs, we consider two variables. The first one is the fees paid from the fund to managers; the second is the loads, including fees and other costs incurred for operational management, e.g., for turnover. Both variables are included in expenses (x_3), which are measured as percentages (average of the sample period) of costs over the managed portfolio size. Descriptive statistics for inputs and outputs are reported in Table 3.

5. Results

5.1. FDH efficiency measures

Table 4 reports summary statistics for efficiencies obtained using FDH. Results are reported for all mutual funds evaluated jointly and also for different categories considered, splitting by type of fund (Islamic vs. socially responsible) and geographic focus (East/West/Middle/global). The joint evaluation, for all 774 mutual funds (138 Islamic and 636 SRI) is reported in the last row of Table 4. Recall that we are

considering an output-oriented approach, i.e. we assume that the objective is to maximize the level of outputs, for a given level of inputs. This implies that the best funds—the benchmarks—are those for which we obtain efficiency values closer to 1, whereas the worst funds would be those for which efficiencies are high.

Results indicate that the magnitude of the inefficiencies found for *both* Islamic and SRI funds are, *on average*, substantial. Given the characteristics of FDH, the number of efficient funds (with efficiency scores equal to 1) is very high due to the removal of the convexity assumption.

Table 4 also reports different summary statistics for the distributions of efficiency in order to provide a more comprehensive view of the results obtained. Although it could *a priori* seem remarkable that, from the first quartile backwards, all funds are efficient (for both SRI and Islamic funds), this trend is actually not surprising if one takes into account that FDH drops the convexity assumption of DEA which, in practical terms, implies that when a given unit cannot be compared with others because of their input/output combinations, it is classified as efficient (with a value of 1) *by default*. Therefore, in the particular case of FDH where we have greater flexibility compared to DEA that comes at the price of a lower discriminating power with yet more difficulties to rank funds according to their performance. This implies that, despite of the beneficial asymptotic properties of FDH referred to in previous sections, and despite being more flexible than DEA, FDH has difficulties for both *discriminating* and, more importantly (especially in the context of mutual fund performance evaluation) for *ranking* units. As indicated earlier, it suits very well that those contexts are ones in which the analyst attempts to ascertain which the most obvious cases of inefficiency are.

Although we have specified a *common* frontier or *metafrontier* (Battese et al., 2004), as indicated above we also report results considering both a classification of funds according to type (Islamic vs. SRI) and geographical focus (East/West/Middle/global). Results indicate that, *on average*, ethical funds perform better than Islamic funds—although it should be properly tested to measure if the differences are significant or not. The different summary statistics considered in Table 4 point out that the discrepancies are not constant across all the distributions of efficiency. Specifically, for the first quartile, which include the best performers (recall we are considering an output orientation), there are no differences—the two types of funds are efficient. For the third quartile, the differences are also very small (1.5632 in the case of Islamic funds and 1.5320 in the case of SRI funds). This descriptive analysis, based entirely on summary statistics, suggests that a quantile approach, which will be carried out below, might add statistical robustness to these initial guesses.

The geographical analysis also reveals remarkable differences, especially for some categories. In the case of funds focusing in the Middle East, North Africa and GCC countries, the *average* performance is much worse (the average efficiency is much higher) than that corresponding to the rest of geographical areas considered. Again, it should be tested if these differences are significant or not. Particularly, the *average* performance of those funds investing in the Middle East, North Africa and GCC countries (Middle) is 191.87%. For the rest of categories, the magnitudes are more moderate and, on average, the differences are not as marked. Again, this requires a more in-depth analysis in order to ascertain whether the differences we find are statistically significant or not. This will be the main attempt of the analysis in Section 5.3.

15

5.2. Expected order-m and order-α efficiency estimates

The order-m estimators partly overcome the difficulties of FDH and DEA for ranking *efficient* funds—i.e., those with a value of 1. We have computed the order-m estimates for different values of m ($m = 25$, $m = 75$ and $m = 150$). These frontiers are nested and, therefore, for $m' > m$, the order-m frontier is below the order-m' frontier. Although the choice of the m parameter might seem somewhat arbitrary, as shown by Sampaio de Sousa and Schwengber (2005) the impact of the decision might not be so relevant when plotting the order-m efficiencies for different values of m—which usually show that they are highly correlated. Indeed the choice should not be so difficult if one follows Cazals et al. (2002), who suggest that "a few values of m could be used to guide the manager of the production unit to evaluate its own performance". In addition, as indicated by Simar (2003), it is also important to notice the difference between m and n. Whereas m is a "trimming" parameter fixed at any desired level defining the level of the benchmark, n is the sample size and, accordingly, there are no *a priori* links between m and n. This idea of trimming is not new in statistics, but its use in boundary estimation actually is.

Summary statistics for the results obtained using order-m are displayed in Table 5. We report results for the three values of m considered ($m = 25$, $m = 75$ and $m = 150$) in three panels displayed vertically. Regardless of the value of m, the trimming parameter which allows tuning of the percentage of points that will lie above the order-m frontier, the *mean* is always lower than for the FDH case—they are 108.24%, 120.33% and 126.19% for $m = 25$, $m = 75$ and $m = 150$, respectively. This could suggest a superior ability of order-m to rank observations, as it turns out to be the case. Overall, and regardless of the fund classification and geographical focus considered, these results indicate that, on average, the performance of mutual funds could improve much more than what FDH predicts, since many of the *efficient* funds under FDH are not efficient under order-m.

Results for the order-α partial frontier are reported in Table 6 for $\alpha = .90$, $\alpha = .95$ and $\alpha = .99$. It reports analogous information to that reported in Table 5 for the case of the order-m frontiers. Although in this case the impact of the α parameter seems *a priori* stronger, this only occurs because the range of variation is not exactly equivalent to that chosen for m. As indicated in Table 6 the average efficiency for $\alpha = .90$ is even lower than 100 (97.05%), due to the presence of super-efficient funds whose efficiencies are driving the average to values below unity. For the $\alpha = .95$ and $\alpha = .99$ cases, analogously to the order-m case, the average increases, with values of 106.25% and 123.51%, respectively, as expected. One of the main differences one may perceive between results for order-m and order-α is the great impact of α, which makes the order-α to converge with FDH efficiencies much faster.

However, perhaps it is more important to establish that results are robust regardless of the estimator considered, i.e. order-m, order-α, or even the FDH estimator. SRI funds are *on average*, more efficient than Islamic funds and, regarding the geographical focus, those funds investing in the Middle East are also, *on average*, less efficient. More importantly, this result holds regardless of the *tuning* parameter used, both for the order-m and order-α case. Therefore, those funds which perform better regardless of the α parameter considered should be the best candidates to be selected by the investor and, although results vary depending on either the estimator considered or the m and α parameters selected, the rankings are generally preserved.

5.3. Using regression quantiles to understand mutual funds' performance differentials

In the previous section it has been argued that, according to our results, SRI funds always perform better than Islamic funds. Regarding the geographical focus of the funds, we found that those investing in the Middle East were the worse performers, regardless of the estimator considered to measure efficiency (order-m, order-α or FDH), or the tuning parameter (m, α) chosen. However, these conclusions were based on descriptive statistics only. In this section we provide a more detailed analysis of the differences among different types of funds, or the differences among funds focusing on different geographical areas, extending also the analysis to other likely determinants.

Results on the determinants of mutual fund performance, considering all different methods to measure performance, are provided in tables 7 and 8 for order-m and order-α, respectively. We confine the analysis to these two estimators only for space reasons. In both cases results are provided for the different tuning parameters considered (m in the case of order-m and α in the case of order-α), and these results are arrayed vertically. Bootstrapped standard errors are provided in parentheses, and the bootstrap method considered has been the standard xy-pair bootstrap (see Koenker, 2005).

The first covariate we will consider is the type of fund ($TYPE$), which enters as a dichotomous variable in the quantile regression analysis, taking the value of 0 for SRI funds and 1 for Islamic funds. According to the results reported in both tables 7 and 8, the lower performance found for Islamic funds compared to mutual funds across the different methods is, in general, not significant. This result is obtained for most quantiles, i.e. for most values of τ. Please note that, since lower efficiency scores indicate best performance (because of the output orientation adopted), in both tables 7 and 8 results for the best funds are reported in the first columns (lower values of τ), and for those funds with worst performance these are reported in the last ones (higher values of τ).

Both tables also show the benefits of performing the analysis using quantile regression. In the case of order-m (Table 7), regardless of the m parameter considered ($m = 25$, $m = 75$ or $m = 150$), the performance differentials between Islamic and SRI mutual funds are found to be significant for some quantiles only—i.e. except for $\tau = .10$ when $m = 25$, and $\tau = .50$ (the median) and $\tau = .75$ when $m = 75$ and $m = 150$, although in the last two cases only at the 10% significance level. This implies that, since we are maximizing for obtaining our efficiency scores, only for the funds with the worst performance, the superior performance of SRI funds is significant. Across the rest of the quantiles, the differences are not significant.

We obtain further results using quantile regression. For instance, the magnitude of the quantile regression coefficients increases with the quantiles, implying that the performance differentials between both types of funds change depending on the conditional distribution of efficiency. This result is robust for the different values of the tuning parameter (m) considered. For $\tau < .50$ (median regression), the coefficients obtained are in the vicinity of 0. In contrast, for $\tau \geq .50$, the magnitude of the coefficient increases sharply for the highest quantiles, however in these cases ($\tau = .90$ and $\tau = .95$) it is not significant.

Although the values change slightly for some specific quantiles, the results obtained for order-m are generally robust when compared to those others yielded by the order-α estimator. As indicated in Table 8, the performance differences between both types of funds are only significant for some quantiles ($\tau = .50$ and $\tau = .75$ when $\alpha = .99$) and, analogously to what we found for order-m, the values for the estimated coefficients increases with τ—in absolute terms. Therefore, results are generally robust regardless of the

method used to measure efficiency and the value of the tuning parameter.[11]

It is particularly important to note the relative performance of SRI and Islamic funds involved; as neither show the sign of the relationship nor its significance they remain constant across quantiles. The magnitude of the coefficient also varies. Results as these are concealed by methodologies focusing on *average* impacts only such as OLS. Of special note is the finding that, although tables 4, 5 and 6 showed that the relative performance of SRI funds was higher, on average, than that of Islamic funds, the quantile regression analysis (tables 7 and 8) reveals this only happens for some particular quantiles, whereas in other cases, Islamic funds perform better.

Specifically, the highest differentials are found for $\tau = .95$, representing the lowest performance, for which the coefficient is positive and significant. Since $TYPE$ is a dummy variable whose value is 1 for Islamic funds, and we are adopting an output orientation (i.e. the highest quantiles represent worse performance), this would indicate that for this particular case (quantile) Islamic funds' performance are significantly worse compared to SRI funds. This result is robust across both methodologies (order-m, order-α) and tuning parameters (m, α). However, for the rest of the quantiles, there are more instances in which the performance of SRI funds are superior (although in others the opposite holds), and the differences found are, in general, not significant.

For the rest of the variables the quantile regression analysis also provides useful insights. The geographical focus variables, namely, $EAST$, $WEST$ and $MIDDLE$ are dichotomous variables taking the value of one for investments performed in those geographical areas, and zero otherwise. According to the descriptive analysis performed in the preceding sections, one of the most remarkable results was that, regardless of the estimator considered (either FDH, order-m or order-α), the performance of those funds focusing on the Middle East, North Africa and GCC countries ($MIDDLE$) was relatively worse than that obtained for the rest. According to the results displayed in both Table 7 and Table 8, this result is significant throughout—for all quantiles (τ) and for all methods considered.

However, some of the findings would have been very difficult to forecast not only for a descriptive analysis such as the one carried out in the preceding sections, but also from the point of view of regression methods based on the average—such as OLS. Specifically, similarly to the results obtained for the $TYPE$ variable, as indicated in both Table 7 and 8, the magnitude of the estimated coefficients increases strongly with τ, suggesting that, among the worst performers, those focusing on the Middle East, North Africa and GCC countries are much worse than the rest. In contrast, for the best performers (corresponding to low values of τ, since we are maximizing), those focusing on the Middle East are closer to their peers. Similarly to what we obtained for the $TYPE$ variable, this result is invisible to some techniques which provide estimates of the *average* impact such as OLS.

The impact of the other geographical focus variables on performance is more modest. In the case of the $EAST$ (Asia and Asia-Pacific area) it is generally found to be negative (implying better performance), but only significant for the central quantiles. However, these results are mostly robust across methods and tuning parameters. In the case of $WEST$ (Europe and America), although the sign of the relationship

---

[11]Regarding the choice of the tuning parameters, although it is convenient to choose them based on the possible extent of outliers in the data, according to the available literature it is not entirely clear how the selection of the tuning parameters should be done. On this respect, Daraio and Simar (2007, p.78) provide some rough guidance in some particular settings and, in a more recent paper, Daouia and Gijbels (2011) derive an algorithm to disentangle how the two classes of partial frontiers are related (under some particular circumstances) thereby helping in choosing the α parameter as a function of m for a possible comparison between order-α and order-m frontiers. Given the problems associated to the selection of the tuning parameters, we are more interested in whether our results are robust across them.

varies slightly with τ, it is mostly negative throughout (implying a positive impact on performance, due to the output orientation adopted), and with higher values in absolute terms, for the highest quantiles. However, for these cases ($\tau \geq .50$) it is not significant. Given that we are maximizing (output orientation), the negative and significant sign found for the lowest quantiles would indicate that, among the best performers, those focusing on the West are better than the rest. This result is generally found to be robust across methods, (order-m, order-α) and tuning parameters (m, α), although it changes slightly for some particular quantiles ($\tau = .50$).

Finally, for the AGE variable, representing the age of the fund, we obtain an impact that varies both in terms of significance and sign according to the quantile considered. Specifically, as indicated in both Table 7 and 8, they are negative and significant, but only for the highest quantiles. For the lowest ones, the impact is either positive, non-significant, or negligible. This implies that, among the worst performing funds (with higher values of τ), the impact of the age is negative implying that, since we are maximizing, older funds perform better. However, for the best performers, age is irrelevant.

6. Conclusions

Over the last few decades, the literature on SRI performance has grown remarkably, paralleling the world-wide growth of the socially responsible mutual fund industry. However, the literature that compares the relative performance of SRI and conventional funds is not conclusive. Islamic funds have also been growing at exponential rates, but only a handful of studies analysed their performance relative to conventional funds. Although SRI and Islamic funds share some of their underpinnings and ethical orientation, we are not aware of previous studies that compared the relative performance of these two types of funds.

In this study, we have provided a direct comparison of SRI and Islamic funds in a two-stage analysis. In the first stage, the performance of the funds is evaluated using partial frontier methods (order-m and order-α), which provide a robust analysis of performance. In the second stage we use quantile regression methods to test explicitly for the comparative performance of SRI and Islamic funds. These methods offer several advantages over OLS or logistic regression methods, due to their robustness to departures from normality but, more importantly, for providing a specific analysis for each particular quantile of the conditional distributions of performance.

According to our results, sourced from a sample of funds between 2001 and 2011, the average efficiency of socially responsible SRI funds is slightly higher than that of Islamic funds. Although this result is robust across the partial frontier methods considered to measure performance efficiency, further testing performed in the second stage indicates that the differences are, in general, not statistically significant. When they are significant, it only happens for some quantiles of the distribution of efficiencies. We found that for the most inefficient funds the superior performance of SRI funds is really significant. On the other hand, for the best mutual funds this evidence vanished and even Islamic funds perform better than SRI. These results show the benefits of performing the analysis using quantile regression.

When geographical focus is considered, quantile regression reveals that for the higher efficiency quantiles Middle East (MENA and GCC) funds' performance is similar to that of the rest of the funds. Only for the worst mutual funds is the performance of the Middle East. In brief, results reveal disparities in the performance of funds within the Middle East region. Results also vary greatly according to each particular quantile, indicating that some of the funds focusing on this region perform especially well. With respect

to the age of the funds, the quantile regression shows that only within the worst funds is evidence found that older funds perform better.

From a methodological point of view, the study attempts to make a contribution by using two stage analysis of mutual fund performance. Applying different frontier methods (in this case; partial frontiers) contributes to making the evaluation of performance more robust. In addition, using quantile regression methods, in the second stage, is useful for analyzing the determinants of the efficiency scores obtained in the first stage for every quantile. The use of two stage analysis enable us to observe that, in general, for the most efficient funds we find no statistically significant differences among funds based on their investment types (either SRI or Islamic), or due to their geographical focus or age. This implies that, for the best funds, these characteristics would be irrelevant and it would be their managers' value added that mattered more. However, the second stage analysis shows that for the worst funds these characteristics are actually relevant. For instance, among the funds with the worst efficiency, SRI funds focusing on the West and with higher experience—if we choose age as a proxy for experience. are the best performers and the Islamic funds focusing on the Middle East region and with less experience are the worst performers. Therefore, for the worst funds, performance is a result not only of their managers' skills but also of other factors related to the environment, such as their geographical focus, their investment type constraints and their age. In sum, we find an asymmetry regarding the determinants of efficiency for the best and for the worst funds.

Our results have some implications for investors' decisions, as their first concern is to find fund managers who can actually provide added value within a predetermined style or investment objectives. In our context, although there are differences between funds, as belonging to a different geographical area, or investment constraints or experience, managers are able to obtain superior results for investors. From an investor's point of view, the arguments above imply that their first concern is to find managers who can actually add value. Comparing Islamic to SRI funds, we found that the different screening criteria do not affect efficiency if managers can actually add value. However, if managers are not able to significantly add value to the fund, we find that other factors have an effect on a varying scale, depending on the fund's efficiency.

Our results provide significant contribution to the previous literature that has compared the performance of conventional funds with that of SRI funds or with that of Islamic funds. Globally evaluated, these studies have not found conclusive evidence in one or another direction. Evaluating previous literature jointly with our results, it could well be the case that the differences between conventional funds and those with screening criteria only existed for the worst funds, but not for the best. This would imply that the screening criteria might affect efficiency only if the manager is not able to add value in his/her management of the fund. We think these results represent an interesting contribution that opens up opportunities for further research in this particular area.

We also believe that our methodological and results contribution is relevant to other topics in the mutual fund literature. Specifically, there is a vast literature on mutual funds analyzing the determinants of their performance. In most of these applications, the methods applied have not considered that results might vary for different parts of the conditional distribution of performance—e.g., for the upper and lower tails. Our results point out the existence of an asymmetry in the determinants depending on the success of the fund. The implementation of our proposed methodology could shed light in some cases, particularly in those instances in which the literature has been inconclusive about the determinants of mutual fund

performance.

Our results are also relevant to the literature that has examined mutual funds' performance persistence which, in general, is not conclusive as to whether performance persistence exists. From our results' perspective, the asymmetry in the determinants of efficiency might have some implications for the factors explaining mutual funds' performance persistence. In this respect, we envisage that when assessing the persistence of performance for the best funds, we only need to consider only the fund's value added. However, for the worst funds, since a larger share of their efficiency is explained by several determinants, the persistence attributable to their poor performance would be a consequence of the persistence in the factors explaining their efficiency. For a given level of financial market efficiency, it is easier to assume that the environmental factors will be more persistent than the manager's value added. In this case, it will be more probable to find persistence among the worst funds rather than the best ones. We recommend further research to be developed according to this evidence.

References

Ahmad, Z. and Ibrahim, H. (2002). A study of performance of the KLSE Shariah Index. *Malaysian Management Journal*, 6:25–34.

Alamá, L. and Tortosa-Ausina, E. (2012). Financial exclusion and bank branch geographic location patterns in Spain: 1988-2008. *Growth and Change*, forthcoming.

Albaity, M. and Ahmad, R. (2008). Performance of Syariah and composite indices: Evidence from Bursa Malaysia. *Asian Academy of Management Journal of Accounting and Finance*, 4(1):23–43.

Balaguer-Coll, M. T., Prior, D., and Tortosa-Ausina, E. (2007). On the determinants of local government performance: A two-stage nonparametric approach. *European Economic Review*, 51(2):425–451.

Banker, R. D. and Natarajan, R. (2008). Evaluating contextual variables affecting productivity using Data Envelopment Analysis. *Operations Research*, 56(1):48–58.

Basso, A. and Funari, S. (2001). A Data Envelopment Analysis approach to measure the mutual fund performance. *European Journal of Operational Research*, 135(3):477–492.

Battese, G. E., Rao, D., and O'Donnell, C. J. (2004). A metafrontier production function for estimation of technical efficiencies and technology gaps for firms operating under different technologies. *Journal of Productivity Analysis*, 21(1):91–103.

Bauer, R., J. Derwall, J., and Otten, R. (2007). The ethical mutual fund performance debate: New evidence from Canada. *Journal of Business Ethics*, 70:111–124.

Binmahfouz, S. (2012). *Investment Characteristics of Islamic Investment Portfolios: Evidence from Saudi Mutual Funds and Global indices*. PhD thesis, Durham University.

BinMahfouz, S. and Hassan, M. K. (2012). A comparative study between the investment characteristics of Islamic and conventional equity mutual funds in Saudi Arabia. *The Journal of Investing*, 21(4):128–143.

Brammer, S., Brooks, C., and Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management*, 35:97–116.

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

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

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

Chen, C. R. and Huang, Y. (2012). Mutual fund governance and performance: A quantile regression analysis of Morningstar's Stewardship Grade. *Corporate Governance: An International Review*, 9(4):311–333.

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

Coad, A. and Hölzl, W. (2009). On the autocorrelation of growth rates. *Journal of Industry, Competition and Trade*, 9(2):139–166.

Coad, A. and Rao, R. (2008). Innovation and firm growth in high-tech sectors: A quantile regression approach. *Research Policy*, 37(4):633–648.

Daouia, A., Florens, J. P., and Simar, L. (2012). Regularization of nonparametric frontier estimators. *Journal of Econometrics*, 168:285–299.

Daouia, A. and Gijbels, I. (2011). Robustness and inference in nonparametric partial frontier modeling. *Journal of Econometrics*, 161(2):147–165.

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

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

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

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

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

Derigs, U. and Marzban, S. (2008). Review and analysis of current Shariah-compliant equity screening practices. *International Journal of Islamic and Middle Eastern Finance and Management*, 1(4):285–303.

Dharani, M. and Natarajan, P. (2011). Equanimity of risk and return relationship between shariah index and general index in India. *Journal of Economics and Behavioral Studies*, 2(5):213–222.

Diltz, J. D. (1995). The private cost of socially responsible investing. *Applied Financial Economics*, 5(2):69–77.

Ernst & Young (2011). Islamic Funds & Investments Report 2011: Achieving Growth in Challenging Times.

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

Forte, G. and Miglietta, F. (2007). Islamic mutual funds as faith-based funds in a socially responsible context. Technical report, Social Science Research Network. Available at SSRN: http://ssrn.com/abstract=1012813 or http://dx.doi.org/10.2139/ssrn.1012813.

Fowler, S. J. and Hope, C. (2007). A critical review of sustainable business indices and their impact. *Journal of Business Ethics*, 76(3):243–252.

Girard, E. C. and Hassan, M. K. (2008). Is there a cost to faith-based investing: Evidence from FTSE Islamic indices. *The Journal of Investing*, 17(4):112–121.

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

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

Goldreyer, E. F., Ahmed, P., and Diltz, J. D. (1999). The performance of socially responsible mutual funds: Incorporating sociopolitical information in portfolio selection. *Managerial Finance*, 25(1):23–36.

Gregory, A., Matatko, J., and Luther, R. (1997). Ethical unit trust financial performance: Small company effects and fund size effects. *Journal of Business Finance and Accounting*, 24(5):705–725.

Grossman, B. R. and Sharpe, W. F. (1986). Financial implications of South African divestment. *Financial Analysts Journal*, 42(4):15–29.

Guerard, J. B. (1997a). Additional evidence on the cost of being socially responsible in investing. *Journal of Investing*, 6(4):31–35.

Guerard, J. B. (1997b). Is there a cost to being socially responsible? *Journal of Investing*, 6(2):11–18.

Hamilton, S., Jo, H., and Statman, M. (1993). Doing well while doing good? The investment performance of socially responsible mutual funds. *Financial Analysts Journal*, pages 62–66.

Hashim, N. (2008). The ftse global islamic and the risk dilemma. Working Paper Series 2008-08, American International University, Office of Research and Publications (ORP), Bangladesh.

Hayat, R. and Kraeussl, R. (2011). Risk and return characteristics of Islamic equity funds. *Emerging Markets Review*, 12(2):189–203.

Hiagh, M. and Hazelton, J. (2004). Financial markets: A tool for social responsibility. *Journal of Business Ethics*, 52(1):59–71.

Hoff, A. (2007). Second stage DEA: Comparison of approaches for modelling the DEA score. *European Journal of Operational Research*, 181(1):425–435.

Hossain, M. Z. (2009). Why is interest prohibited in Islam? A statistical justification. *Humanomics*, 25(4):241–253.

Huberman, G. and Jiang, W. (2006). Offering versus choice in 401(k) plans: equity exposure and number of plans. *The Journal of Finance*, 61(2):763–801.

Hussein, K. and Omran, M. (2005). Ethical investment revisited. *The Journal of Investing*, 14(3):105–126.

Illueca, M., Pastor, J. M., and Tortosa-Ausina, E. (2009). The effects of geographic expansion on the productivity of Spanish savings banks. *Journal of Productivity Analysis*, 32(2):119–143.

Iqbal, Z. (1997). Islamic financial systems. *Finance and Development*, 34:42–45.

Koenker, R. (2005). *Quantile Regression*. Cambridge University Press.

Koenker, R. (2012). *quantreg: Quantile Regression*. R package version 4.79.

Kumar, S. and Russell, R. R. (2002). Technological change, technological catch-up, and capital deepening: Relative contributions to growth and convergence. *American Economic Review*, 92(3):527–548.

Lemieux, T. (2006). Postsecondary education and increasing wage inequality. *The American Economic Review*, 96(2):195–199.

Luther, R. G., Matatko, J., and Corner, D. C. (1992). The Investment performance of UK ethical Unit Trusts. *Accounting Auditing & Accountability Journa*, 5:57–70.

Machado, J. A. F. and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of applied Econometrics*, 20(4):445–465.

Mallin, C., Saadouni, B., and Briston, R. (1995). The financial performance of ethical investment trusts. *Journal of Business Finance and Accounting*, 22(4):483–496.

Mansor, F. and Bhatti, I. (2011). Risk and return analysis on performance of the Islamic mutual funds: evidence from Malaysia. *Global Economy and Finance Journal*, 4(1):19–31.

McDonald, J. (2009). Using least squares and tobit in second stage DEA efficiency analyses. *European Journal of Operational Research*, 197(2):792–798.

Meligkotsidou, L., Vrontos, I. D., and Vrontos, S. D. (2009). Quantile regression analysis of hedge fund strategies. *Journal of Empirical Finance*, 16:264–279.

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

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

Novethic (2009). Islamic finance and SRI: any crossover? Unpublished manuscript.

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

Perks, R. W., Rawlinson, D. H., and Ingram, L. (1992). British accounting review. *British Accounting Review*, 24(1):43–65.

Pilyavsky, A. and Staat, M. (2008). Efficiency and productivity change in Ukrainian health care. *Journal of Productivity Analysis*, 29(2):143–154.

Ramalho, E. A., Ramalho, J. J. S., and Henriques, P. D. (2010). Fractional regression models for second stage dea efficiency analyses. *Journal of Productivity Analysis*, 34(3):239–255.

Rudd, A. (1981). Social responsibility and portfolio performance. *California Management Review*, 23(4):55–61.

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

Sauer, D. A. (1997). The impact of social-responsibility screens on investment performance: Evidence from the Domini 400 Social Index and Domini Equity Mutual Fund. *Review of Financial Economics*, 6(2):137–149.

Schepers, D. H. (2003). A critique of social investing's diversity measures. *Business & Society Review*, 108(4):487–508.

Shanmugam, B. and Zahari, Z. (2009). A primer on Islamic finance. Research foundation publications, CFA Institute, Charlottesville.

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

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of productive processes. *Journal of Econometrics*, 136(1):31–64.

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

Simar, L. and Wilson, P. W. (2011). Two-stage DEA: caveat emptor. *Journal of Productivity Analysis*, 36(2):205–218.

Social Investment Forum (2010). *Report on Socially Responsible Investing Trends in the United States*. Available at: http://ussif.org/resources/research/.

Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal*, 56:30–39.

Tortosa-Ausina, E. (2004). An alternative conditioning scheme to explain efficiency differentials in banking. *Economics Letters*, 82:147–155.

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

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

Table 1: Comparison of the screening criteria of SRI and Islamic funds[a]

	CRITERIA	ISLAMIC FUNDS	SOCIALLY RESPONSIBLE INVESTMENT FUNDS (SRI)
Similarities	Broad objectives	Both types of funds have ethical, social and financial objectives.	
	Negative screening	Both types of funds use negative screening (filtering) criteria in the selection of stocks to be included in their portfolios.	
	Shareholder advocacy	Shareholders in both types of funds are encouraged to formally express any negative opinion regarding certain practices.	
Differences	Source of guidance	*Shariah*, however, owing to the lack of a global Shariah supervisory body, differing interpretations between funds are currently used.	Historically, SRI originated with religious groups avoiding investing in Sin stocks. However, the concept of SRI evolved over time, and became further developed and broadened by environmental strategies, anti-war projects, and human rights activism. There are no universally recognised definitions for SRI and ESG (Environmental, Social and Governance) principles.
	Supervisory Committee	*Shariah Supervisory board* (SSB), whose opinion is <u>binding</u>.	*Ethical Committee*, acting as an <u>advisory</u> body.
	Restrictions on investment activities and instruments	Yes. This type of fund excludes investments in fixed-income instruments, such as corporate bonds, certificates of deposit (CDs), preferred stocks, warrants, and some derivatives (e.g., options). <u>Equity</u> mutual funds represent the largest portion of Islamic funds.	No. SRI mutual funds can freely choose between debt-bearing investments and equity-bearing investments, as long as the stocks chosen adhere to SRI and ESG principles.
	Financial Ratio Screens	Yes. Financial filters, determined by the SSB, are applied during the stock selection process. The core principles to which the filters are related are: leverage, presence of interest-bearing assets and liabilities, high level of debt, and credit.	There are no financial parameters that determine the inclusion of an asset in the SRI index. These are freely decided by the fund manager.
	Purification process	Yes. Purification is the process of eliminating or cleaning the portfolio of income or gain resulting from interest or other impermissible revenue sources. Impermissible portfolio income is donated to charities and non-profit organisations.	No.
	Positive Screening	No.	No.
	Sector exclusion	Yes, sectors considered not compliant with *Shariah* are excluded. Examples include: alcoholic beverages, tobacco, weapons production or distribution, gambling, pornography, conventional financial services, leisure, etc.	Yes. Sectors not compliant with SRI criteria are excluded. Examples include: alcoholic beverages, tobacco, weapons production or distribution, gambling, and pornography.
	Best-in-class	No. There is a general distinction between admissible and prohibited assets. The strategy is in-out.	Yes. Some funds include firms operating in sectors generally forbidden, if they exhibit a commitment to SRI principles.
	Screens based on environmental filters	No.	Yes.
	Screens related to human rights	No.	Yes.
	Screens associated with transparent corporate practices	No.	Yes, but not in all cases.

[a] Adapted from Forte and Miglietta (2007).

Table 2: Descriptive statistics for Islamic and SRI funds distributed by geographical area (2001–2011)

Panel A: Islamic equity mutual funds sample										
Geographical focus[a]	Number	% Date	Age (years)	Annual expenses	Annual gross return	Annual net return		Std.dev. annual net return		Monthly net returns	
						Average	Median	Average	Median	Skewness	Kurtosis
East	51	77%	10.16	1.53%	10.87%	9.35%	9.78%	20.78%	20.11%	−0.48	2.78
Middle	39	69%	8.07	1.69%	0.63%	−1.06%	3.06%	31.27%	17.27%	−1.00	4.97
West	19	82%	10.57	1.41%	8.00%	6.58%	−1.31%	18.01%	27.85%	−0.59	1.92
Global	29	86%	11.40	1.40%	3.97%	2.57%	5.54%	18.99%	18.10%	−1.80	11.06
All funds	138	77%	9.88	1.53%	6.13%	4.60%	4.95%	22.99%	19.92%	−0.92	5.02

Panel B: SRI equity mutual funds sample										
Geographical focus[a]	Number	% Date	Age (years)	Annual expenses	Annual gross return	Annual net return		Std.dev. annual net return		Monthly net returns	
						Average	Median	Average	Median	Skewness	Kurtosis
East	73	75%	10.36	1.26%	13.84%	12.58%	13.18%	24.11%	24.71%	−0.73	2.52
West	350	86%	12.62	0.89%	8.01%	7.12%	5.78%	21.81%	18.56%	−0.87	3.84
Global	213	85%	10.90	1.32%	7.60%	6.28%	5.71%	19.98%	20.57%	−0.89	4.98
All funds	636	84%	11.78	1.08%	8.54%	7.46%	6.29%	21.46%	20.32%	−0.86	4.07

[a] East refers to the Pacific and Asia-Pacific area, Middle refers to Middle East, North Africa and GCC countries (Gulf Cooperation Council countries, which include Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates), West refers to Europe and America, and global does not have any particular area of reference.

Table 3: Descriptive statistics for inputs and outputs, Islamic and SRI funds (2001–2011)[a]

Class: Islamic		Mean	1st quartile	Median	3rd quartile	Std.dev.
Outputs	Gross return (y_1)	0.0051	0.0023	0.0054	0.0093	0.0075
	Skewness (y_2)	−0.9166	−1.0763	−0.5727	−0.2446	1.4745
Inputs	Std.dev. (x_1)	0.0664	0.0499	0.0575	0.0758	0.0280
	Expenses (x_2)	0.0153	0.0150	0.0150	0.0175	0.0048
	Kurtosis (x_3)	5.0176	0.6650	1.5104	3.5814	13.0013
Number of funds: 138						
Class: SRI		Mean	1st quartile	Median	3rd quartile	Std.dev.
Outputs	Gross return (y_1)	0.0071	0.0035	0.0061	0.0100	0.0056
	Skewness (y_2)	−0.8614	−1.0260	−0.8229	−0.5883	1.0122
Inputs	Std.dev. (x_1)	0.0620	0.0519	0.0587	0.0687	0.0151
	Expenses (x_2)	0.0108	0.0050	0.0113	0.0150	0.0071
	Kurtosis (x_3)	4.0704	1.6535	2.5617	3.6268	7.9252
Number of funds: 636						

Table 4: FDH efficiencies, mutual funds (2001–2011)

		Mean	1st quartile	Median	3rd quartile	Std.dev.
Fund classification	Islamic	1.4720	1.0000	1.2020	1.5632	0.7789
	SRI	1.3237	1.0000	1.2842	1.5320	0.3561
Geographic focus	East	1.1878	1.0000	1.0639	1.2665	0.2932
	West	1.2863	1.0000	1.1919	1.5087	0.3729
	Middle	1.9187	1.1095	1.7235	2.1549	1.1863
	Global	1.4391	1.1479	1.4444	1.5691	0.3622
All funds		1.3502	1.0000	1.2667	1.5340	0.4636

Notes: the numbers represent descriptive statistics for the efficiency scores yielded by Free Disposal Hull (FDH). Since we are adopting an output orientation, efficiencies closer to one indicate better performance, and if a given fund achieves an efficiency score of 1 it implies it is fully efficient (best performance).

Table 5: Order-m efficiencies, mutual funds (2001–2011)

		Mean	1st quartile	Median	3rd quartile	Std.dev.
			$m = 25$			
Fund classification	Islamic	1.2123	0.9831	1.0436	1.2644	0.5000
	SRI	1.0542	0.9133	1.0577	1.1961	0.2737
Geographical focus	East	1.0667	0.9701	0.9999	1.1569	0.2180
	West	1.0146	0.8633	0.9990	1.1805	0.3014
	Middle	1.5062	1.0254	1.3239	1.5613	0.7616
	Global	1.1255	0.9920	1.1190	1.2171	0.2424
All funds		1.0824	0.9349	1.0550	1.2063	0.3309
			$m = 75$			
Fund classification	Islamic	1.3141	1.0000	1.1150	1.3577	0.5673
	SRI	1.1793	0.9976	1.1550	1.3147	0.2659
Geographical focus	East	1.1383	1.0000	1.0451	1.2499	0.2231
	West	1.1487	0.9795	1.0573	1.2957	0.2854
	Middle	1.6572	1.0713	1.5145	1.7465	0.8636
	Global	1.2468	1.0825	1.2334	1.3511	0.2614
All funds		1.2033	0.9993	1.1497	1.3271	0.3432
			$m = 150$			
Fund classification	Islamic	1.3694	1.0000	1.1649	1.4457	0.6146
	SRI	1.2386	1.0000	1.2120	1.3912	0.2856
Geographical focus	East	1.1647	1.0000	1.0579	1.2658	0.2439
	West	1.2076	0.9993	1.1077	1.3784	0.3035
	Middle	1.7440	1.0963	1.6479	1.8942	0.9270
	Global	1.3167	1.1352	1.3182	1.4167	0.2849
All funds		1.2619	1.0000	1.2065	1.3950	0.3694

Notes: the numbers represent descriptive statistics for the efficiency scores yielded by order-m, and for three tuning parameters ($m = 25$, $m = 75$, $m = 150$). Since we are adopting an output orientation, efficiencies closer to one indicate better performance, and if a given fund achieves an efficiency score of 1 it implies it is fully efficient (best performance). Efficiency scores lower than 1 indicate that the unit under analysis is an outlier.

Table 6: Order-α efficiencies (maximizing), mutual funds (2001–2011)

		Mean	1st quartile	Median	3rd quartile	Std.dev.
			$\alpha = .90$			
Fund classification	Islamic	1.0603	0.8901	1.0000	1.0964	0.4016
	SRI	0.9476	0.8502	0.9861	1.0535	0.2312
Geographical focus	East	0.9448	0.8000	1.0000	1.0261	0.1996
	West	0.9315	0.8000	0.9655	1.0471	0.2628
	Middle	1.3071	1.0000	1.1664	1.3464	0.5802
	Global	0.9761	0.8667	1.0000	1.0559	0.2116
All funds		0.9705	0.8549	1.0000	1.0585	0.2778
			$\alpha = .95$			
Fund classification	Islamic	1.1477	1.0000	1.0000	1.1672	0.4428
	SRI	1.0409	0.9344	1.0102	1.1489	0.2345
Geographical focus	East	1.0335	0.9504	1.0000	1.1178	0.1984
	West	1.0334	0.9329	1.0056	1.1333	0.2611
	Middle	1.3906	1.0000	1.1667	1.4603	0.6762
	Global	1.0593	0.9627	1.0348	1.1512	0.2276
All funds		1.0625	0.9531	1.0056	1.1509	0.2919
			$\alpha = .99$			
Fund classification	Islamic	1.3042	1.0000	1.1114	1.3575	0.5463
	SRI	1.2175	1.0000	1.1835	1.3333	0.2497
Geographical focus	East	1.1597	1.0000	1.0595	1.2645	0.2209
	West	1.2073	1.0000	1.1667	1.3333	0.2785
	Middle	1.6298	1.0933	1.4541	1.7174	0.8457
	Global	1.2437	1.0869	1.2098	1.3402	0.2425
All funds		1.2351	1.0000	1.1667	1.3404	0.3332

Notes: the numbers represent descriptive statistics for the efficiency scores yielded by order-α, and for three tuning parameters ($\alpha = 0.90$, $\alpha = 0.95$, $\alpha = 0.99$). Since we are adopting an output orientation, efficiencies closer to one indicate better performance, and if a given fund achieves an efficiency score of 1 it implies it is fully efficient (best performance). Efficiency scores lower than 1 indicate that the unit under analysis is an outlier.

Table 7: Determinants of mutual funds' performance, regression quantiles, order-m

			$m = 25$				
Covariates	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	$\tau = 0.95$
(Intercept)	0.767 (0.044)	0.835 (0.028)	0.986 (0.019)	1.117 (0.020)	1.252 (0.030)	1.372 (0.039)	1.546 (0.100)
$TYPE$	0.056 (0.046)	0.077 (0.034)	0.029 (0.024)	−0.011 (0.021)	−0.040 (0.034)	0.028 (0.126)	0.408 (0.234)
$EAST$	0.009 (0.070)	−0.008 (0.033)	−0.048 (0.023)	−0.107 (0.022)	−0.056 (0.033)	0.015 (0.056)	−0.147 (0.105)
$WEST$	−0.300 (0.057)	−0.190 (0.050)	−0.127 (0.017)	−0.115 (0.018)	−0.036 (0.026)	−0.008 (0.031)	−0.156 (0.101)
$MIDDLE$	0.073 (0.058)	0.048 (0.041)	0.001 (0.053)	0.228 (0.087)	0.386 (0.162)	1.143 (0.441)	1.163 (0.820)
AGE	0.004 (0.003)	0.003 (0.002)	0.001 (0.001)	0.000 (0.001)	−0.002 (0.003)	−0.003 (0.002)	−0.002 (0.003)

			$m = 75$				
Covariates	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	$\tau = 0.95$
(Intercept)	0.920 (0.022)	0.969 (0.015)	1.081 (0.024)	1.251 (0.030)	1.394 (0.040)	1.518 (0.042)	1.768 (0.118)
$TYPE$	0.004 (0.031)	0.024 (0.025)	0.006 (0.009)	−0.060 (0.023)	−0.044 (0.033)	0.063 (0.158)	0.425 (0.252)
$EAST$	0.003 (0.040)	−0.007 (0.023)	−0.087 (0.025)	−0.170 (0.025)	−0.094 (0.037)	−0.010 (0.064)	−0.221 (0.122)
$WEST$	−0.098 (0.029)	−0.068 (0.018)	−0.104 (0.025)	−0.184 (0.033)	−0.040 (0.026)	−0.028 (0.034)	−0.201 (0.117)
$MIDDLE$	0.053 (0.033)	0.005 (0.031)	−0.032 (0.085)	0.359 (0.139)	0.417 (0.124)	1.270 (0.628)	1.328 (1.127)
AGE	0.002 (0.002)	0.000 (0.001)	0.000 (0.001)	−0.001 (0.002)	−0.004 (0.003)	−0.003 (0.003)	−0.004 (0.005)

			$m = 150$				
Covariates	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	$\tau = 0.95$
(Intercept)	0.978 (0.012)	0.999 (0.005)	1.128 (0.028)	1.333 (0.029)	1.483 (0.042)	1.640 (0.050)	1.861 (0.104)
$TYPE$	0.014 (0.017)	0.006 (0.007)	0.000 (0.003)	−0.090 (0.029)	−0.049 (0.043)	0.061 (0.151)	0.524 (0.256)
$EAST$	0.003 (0.015)	−0.005 (0.007)	−0.128 (0.029)	−0.225 (0.028)	−0.148 (0.046)	−0.064 (0.067)	−0.274 (0.110)
$WEST$	−0.035 (0.015)	−0.030 (0.008)	−0.128 (0.029)	−0.201 (0.051)	−0.026 (0.038)	−0.033 (0.040)	−0.219 (0.115)
$MIDDLE$	0.006 (0.019)	−0.005 (0.020)	−0.049 (0.083)	0.416 (0.172)	0.503 (0.187)	1.310 (0.617)	1.294 (1.127)
AGE	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	−0.001 (0.002)	−0.005 (0.003)	−0.006 (0.003)	−0.003 (0.006)

Notes: the numbers represent coefficients for the estimated linear quantile regression for each of the selected quantiles (τ). Bootstrapped standard errors are provided in parentheses. See Koenker (2012) for further details. Since we are adopting an output orientation, lower values of τ are associated with best performance, and vice versa.

Table 8: Determinants of mutual funds' performance, regression quantiles, order-α

	$\tau = 0.05$	$\tau = 0.10$	$\tau = 0.25$	$\tau = 0.50$	$\tau = 0.75$	$\tau = 0.90$	$\tau = 0.95$
				$\alpha = 0.90$			
Covariates							
(Intercept)	0.599 (0.046)	0.728 (0.041)	0.860 (0.032)	1.004 (0.018)	1.080 (0.019)	1.173 (0.036)	1.335 (0.081)
$TYPE$	0.052 (0.052)	−0.004 (0.049)	−0.020 (0.042)	−0.001 (0.021)	−0.044 (0.027)	0.049 (0.091)	0.391 (0.181)
$EAST$	−0.067 (0.053)	−0.042 (0.049)	−0.074 (0.030)	0.001 (0.029)	−0.003 (0.029)	−0.002 (0.035)	−0.150 (0.081)
$WEST$	−0.194 (0.056)	−0.154 (0.049)	−0.070 (0.036)	−0.031 (0.019)	−0.009 (0.023)	0.046 (0.034)	−0.084 (0.085)
$MIDDLE$	0.137 (0.105)	0.180 (0.068)	0.154 (0.047)	0.167 (0.059)	0.338 (0.159)	1.019 (0.462)	1.018 (0.646)
AGE	0.006 (0.003)	0.003 (0.003)	0.001 (0.003)	−0.001 (0.002)	−0.002 (0.002)	−0.003 (0.002)	−0.004 (0.006)
				$\alpha = 0.95$			
Covariates							
(Intercept)	0.797 (0.063)	0.813 (0.046)	0.958 (0.030)	1.034 (0.016)	1.181 (0.023)	1.289 (0.051)	1.419 (0.122)
$TYPE$	−0.017 (0.083)	0.057 (0.048)	0.061 (0.028)	0.000 (0.007)	−0.062 (0.041)	0.028 (0.084)	0.364 (0.210)
$EAST$	−0.003 (0.063)	−0.013 (0.043)	−0.019 (0.028)	−0.034 (0.014)	−0.032 (0.036)	−0.015 (0.054)	−0.109 (0.126)
$WEST$	−0.076 (0.076)	−0.041 (0.030)	−0.028 (0.033)	−0.028 (0.016)	−0.010 (0.024)	0.008 (0.040)	−0.112 (0.104)
$MIDDLE$	0.213 (0.138)	0.131 (0.054)	−0.018 (0.041)	0.133 (0.048)	0.370 (0.179)	1.058 (0.492)	1.096 (0.910)
AGE	−0.003 (0.005)	0.000 (0.004)	0.000 (0.002)	0.000 (0.001)	−0.002 (0.002)	−0.003 (0.003)	0.001 (0.008)
				$\alpha = 0.99$			
Covariates							
(Intercept)	1.000 (0.011)	1.000 (0.001)	1.086 (0.019)	1.238 (0.035)	1.371 (0.029)	1.508 (0.056)	1.773 (0.121)
$TYPE$	0.000 (0.023)	0.000 (0.005)	0.000 (0.000)	−0.101 (0.035)	−0.092 (0.039)	0.018 (0.143)	0.316 (0.197)
$EAST$	0.000 (0.009)	0.000 (0.001)	−0.086 (0.019)	−0.113 (0.033)	−0.075 (0.035)	0.012 (0.075)	−0.129 (0.112)
$WEST$	−0.056 (0.024)	0.000 (0.009)	−0.086 (0.019)	−0.040 (0.032)	0.002 (0.025)	−0.008 (0.042)	−0.040 (0.130)
$MIDDLE$	0.000 (0.028)	0.000 (0.025)	0.001 (0.089)	0.330 (0.117)	0.456 (0.212)	1.276 (0.513)	1.266 (1.011)
AGE	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	−0.001 (0.002)	−0.002 (0.003)	−0.003 (0.003)	−0.011 (0.006)

Notes: the numbers represent coefficients for the estimated linear quantile regression for each of the selected quantiles (τ). Bootstrapped standard errors are provided in parentheses. See Koenker (2012) for further details. Since we are adopting an output orientation, lower values of τ are associated with best performance, and vice versa.

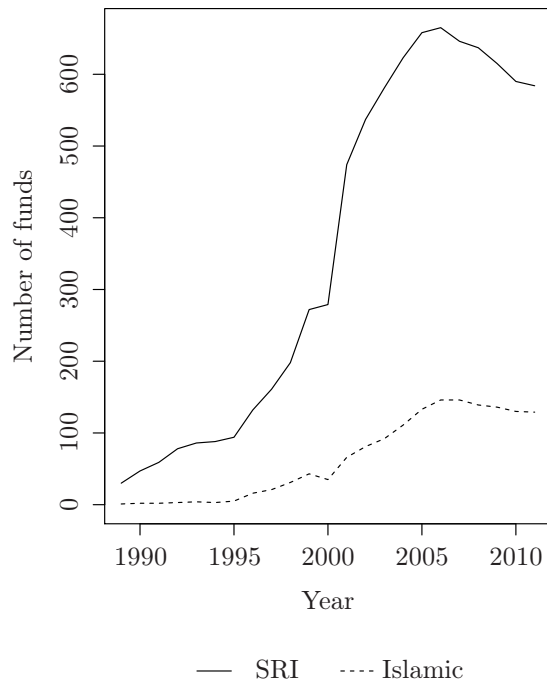**Figure 1:** Evolution of the number of SRI funds and Islamic funds, 1989–2011

Figure 2: Evolution of the number of SRI funds vs. Islamic funds, 1989–2011