Ethical strategy focus and mutual fund management: performance and persistence

November 10, 2018

Abstract

The aim of this study is to analyze whether managers, practitioners and individual investors could obtain higher risk-adjusted returns by allocating their investments to funds that integrate specific levels of socially responsible (SR) criteria in their portfolios. This is achieved by comparing the performance of mutual funds according to their SR characteristics: environmental, governance, social, and sustainability attributes. For a large sample of 3,920 equity SR mutual funds around the world, performance is measured using a multifactor model that incorporates relevant benchmarks according to the fund investment objectives, and using Carhart's (1997) methodology to measure mutual fund performance persistence. In general, fund performance is not significant, the average being negative and close to zero. Funds achieving relatively high levels of SR attributes in their portfolios seem to experience overall worse performances. This evidence is, however, mainly driven by the behavior of worst-performing funds. Moreover, investing in the previous best SR funds could lead investors to greater overall returns in most areas and levels of SR attributes considered. This evidence highlights the role of managers in enhancing the returns of a portfolio with a well-defined SR investment policy. Therefore, there is no incompatibility between pursuing higher ethical (and sustainable) values as well as greater financial performances from investments—provided managers have the skills necessary to choose the right SR funds.

Keywords: environmental, mutual fund, performance, persistence, SRI, sustainability

JEL Classification: G2, N25, Z12

1. Introduction

Sustainable investing is transforming the panorama of the global financial industry. Interest in the perspective of long-term returns is growing, and the recently observed trends in favor of integrating sustainability concerns into investment worldwide are indicative of its momentum resulting in an explosive growth. Thus, according to the Global Sustainable Investment Alliance (GSIA)'s review,[1] from 2014 to 2016 the trend for sustainability embedded in investment represents a significant share of the market not only in Europe and Australia, where it accounts for approximately 50 percent of their professionally managed assets, but also in the United States and Canada, where its share of the market ranges from 22 to 38 percent. Investors and investment managers are aware of the opportunities provided by the Socially Responsible Investment (SRI); specifically, in 2017, more than one quarter of the world's professionally managed assets—roughly US$22.9 trillion—have some sort of sustainable investing mandate, with about US$8.7 trillion of that in the United States, US$12 trillion in Europe and the rest shared among other regions.[2]

In this context, global growth in sustainable investing is aligned with an emergence of specialized literature on the expansion of organizations that meet environmental, social and governance (ESG) criteria into their strategies (see Humphrey et al., 2012; O'Rourke, 2003, among others). Specifically, the motivations for centering attention on the Socially Responsible (SR) mutual funds have been investigated by a number of authors, such as Renneboog et al. (2008b), Capelle-Blancard and Monjon (2012), and Silva and Cortez (2016), to cite just a few. Likewise, Busch et al. (2016) build on the idea that a reorientation towards a long-term paradigm for sustainable investments is important; their main argument is that embedding sustainable values into investments requires a progressive change in mindset toward long-term risk and opportunities arising from the incorporation of such perspective are countless. Moreover, Laurel-Fois (2016) provides evidence of a positive relationship between responsible investment and financial performance as per a risk mitigation effect derived from high screening intensity. The author argues that intensive screening let fund managers to gain advantages from selection strategies.

Despite this, there is a lack of studies analyzing the performance experienced by SR funds according to their SR strategy. In contrast to previous literature, our main goal is not

---

[1] Global Sustainable Investment Alliance, 2017: "2016 Global Sustainable Investment Review".
[2] Morgan Stanley's 2018 edition of the Sustainable Signals series: "Asset Owners Embrace Sustainability".

to compare SR to conventional funds[3] but rather to analyze the performance of SR funds with similar characteristics. In other words, we aim to analyze whether managers, practitioners and individual investors could obtain higher risk-adjusted returns by allocating their investments to funds that integrate specific levels of SR criteria in their portfolios. This is achieved by comparing the performance of socially responsible (SR) mutual funds according to their SR characteristics: environmental, governance, social, and sustainability attributes.

To address this issue, we first compare the performance of SR funds according to their SR strategy, reflected in the level of SR attributes accomplished in their portfolio. These characteristics refer to Environmental, Governance, Social, and Sustainability attributes. SR funds not reporting information on these characteristics are also included in the analyses as "undefined" funds. With the aim of avoiding any local bias related to the investment geographical area, we build groups of funds and repeat the analysis for each of them.

The empirical work focuses on equity SR mutual funds around the world for the period from January, 2000, to March, 2018. The sample is made by 3,920 funds. Among them, only 180 funds have data for the full sample period, meaning that the rest of the funds either disappeared or were set up during this period. Our analysis is therefore free from survivorship bias. We apply a multifactor model to these data in order to estimate abnormal performance.

In sum, this study contributes to the previous literature in several ways. Firstly, we analyze the risk-adjusted returns experienced by SR funds. Rather than comparing their financial results with those achieved by their conventional peers, we assess the SR fund performance in relation to the level of SR criteria integrated in their investment strategies by distinguishing among SR funds with higher and lower SR attributes in their portfolios. Secondly, we also differentiate SR funds according to their investment area in order to avoid any potential local bias. Bearing this aim in mind, we evaluate whether mutual funds achieving higher scores in their SR attributes underperform funds that focus less on their SR strategies. The analysis will go beyond the average (in order to prevent the emergence of effects mainly driven by the performance of the worst-performing funds), for which we follow a nonparametric approach based on kernel density estimation. As far as we know, this methodology has not been applied to analyze the differences in performance with regard to the SR characteristics of mutual funds, and constitutes the third

---

[3]E.g., Bauer et al. (2005), Renneboog et al. (2008b) and Nofsinger and Varma (2014).

contribution of our study. Finally, our fourth contribution lies in assessing the SR funds' performance persistence through the recursive portfolio approach. This methodology has been frequently been applied to analyze the persistence of conventional funds but not of SR funds (see, for instance Bollen and Busse, 2005, 2001; Busse and Irvine, 2006; Cremers and Petajisto, 2009; Cremers et al., 2013; Gottesman and Morey, 2007). It reveals whether the performance differences between the best and the worst SR funds are persistent over time, and for which SR categories and areas analyzed this persistence holds. Providing responses to the objectives of the study might, *a priori*, be useful to professionals and investors aiming to enhance their risk-adjusted returns in order to identify skilled managers who can provide them with greater overall performances while integrating a relevant degree of ethical values in their investment portfolios.

The remainder of the paper is structured as follows: Section 2 presents a brief review of relevant literature and states the main hypotheses to be tested. Section 3 provides details on the methodologies used to measure fund performance and persistence. Section 4 describes the data used in the study, while Section 5 reports the results. Finally, Section 6 presents some concluding remarks.

2. Literature review and hypotheses development

O'Rourke (2003) states that there are several reasons that explain the growth of SR funds in the financial markets around the world. Firstly, SR funds are a sophisticated option that adds value to the investment by following a non-purely financial orientation as other parameters are also considered (such as reputation, good corporate governance practice and environmental responsibility, among others). In this regard, Koellner et al. (2005) propose the basic principles and methods on which a sustainability rating for mutual funds could be based; specifically, they state that a variety of impacts—economic, social and ecological—should be considered in order to embed sustainability into investment processes. In addition, some authors analyze the intention to invest in a socially responsible manner (Palacios-González and Chamorro-Mera, 2018), and show that environmental as a sustainable criteria matter in the decisions of mutual funds because of their effect on the company's future financial performance (Said et al., 2013; Cai and Li, 2018). Moreover, Helminen (2000) introduces the concept of "eco-efficiency" linked to sustainable development as the integration of ecological, economic and ethical dimensions at the firm level;

indeed, ethics have also been integrated into mainstream business as a competitive strategy. In this sense, some studies highlight the impact of integrating ethics into the business strategy (Friede et al., 2015).

However, Jansson and Biel (2014) conduct a survey on SR fund industry and show that future SR investment is not influenced by social and environmental concerns. Rather, their findings suggest that financial beliefs about risk, as well as beliefs about increased market shares, are a relevant driving force underlying SR investment. From a financial perspective, SR mutual funds are investment vehicles which investors can easily access to, and their returns are comparable to those of conventional funds. Indeed, numerous studies have analyzed the financial performance of these portfolios by comparing them to their conventional peers, finding mixed results. On the one hand, some studies argue that SR funds outperform conventional funds (Galema et al., 2008; Kempf and Osthoff, 2007). Nonetheless, it appears that socially responsible investors are less sensitive to poor performance and the overall implicit benefit to the SRI practice dilutes any negative impact, as postulated by Bollen (2007), Benson and Humphrey (2008) and Renneboog et al. (2011), among others. On the other hand, studies such as Bauer et al. (2005) and Renneboog et al. (2008a,b) conclude that, in general, there are no significant differences between the performance of SR mutual funds and their conventional counterparts.

Up to this point, it is worth to note that SR mutual funds usually invest under different core values related to their ethical strategy. For instance, a SR mutual fund with a focus on environmental criteria does invest in companies that potentially contribute to the environment. Hence, attributes such as sustainability or environment should be implicitly related to the main investment objectives of a SR portfolio. Given that SR funds aim to maximize their financial results while achieving their SR goals, it is therefore of interest to observe the behavior of SR funds with the same characteristics or attributes. Most of the previous literature, however, does not focus on SR attributes.

Some studies have compared the performance of different types of SR mutual funds, mainly in relation to green funds. In this line Mallett and Michelson (2010) find no differences in their comparison of returns of green and other SR funds. This same result is found by Climent and Soriano (2011), although this evidence varies over time. Furthermore, observing the level of SR attributes may help investors to distinguish between funds with the same label, but very different SR scores. This could be a solution to Silva and Cortez's (2016) proposal that investors should pay attention to the social performance of

green funds.

Our interest therefore lies in analyzing the relationship between SR attributes and performance. In this study, we aim to fill this gap in the literature by analyzing the performance of mutual funds focusing on similar SR attributes. Firstly, it should be noted that funds in the same SR categories (e.g., environmental) could achieve different scores in their SR portfolios attributes. In other words, two green funds, for example, can differ in the level of environmental characteristics related to the assets held in their portfolios. That is, one green fund could be considered "greener" than another. Accordingly, we should expect the first fund to yield higher SR scores related to environmental attributes than the second. Different levels of SR scores could therefore drive the performance distribution. For instance, in their comparison of US green funds with conventional funds Chang et al. (2012) find that green performance is not uniformly distributed across fund types. In another comparison, Nofsinger and Varma (2014) find that SR funds outperform (underperform) conventional funds during crisis (non-crisis) periods, but their most striking finding is that this asymmetric pattern is driven by SR attributes rather than differences in management or the characteristics in the fund portfolios.

We are interested in analyzing what can be an intricate relationship between SR scores and performance, in which it is not easy to comprehend what should be expected from it. As Sharpe (1992) points out, the performance of a fund depends mainly on the evolution of the asset class it invests in. Therefore, different SR scores could lead to differences in performance. In this sense, higher SR scores in the portfolio could imply higher policy constraints that restrict their investment to very specific types of assets. As a result, the performance of the SR funds would be linked to the behavior of the specific constrained SR securities. In this line, Jin and Han (2018) found that green funds tend towards industry concentration rather than diversification. Additionally, previous evidence for conventional funds (Huang et al., 2011) shows that increasing investment concentration leads to higher specific risks and worse performance. Consequently, SR funds with high scores in their attributes, and therefore with more concentrated portfolios investing in specific securities, should face more difficulties in providing investors with greater financial performance. The analysis of the differences in performance according to SR levels is therefore the first hypothesis of our study:

Hypothesis 1 (H1): *There are no differences in the performance of SR funds according to their SR attributes.*

5

In order to better understand the behavior of SR funds over time, we are also interested in analyzing the persistence of their performance. This analysis is useful to distinguish whether mutual fund results are due to luck or to managers' ability. Thus, if there is persistence, the best-performing funds during a given period should experience greater subsequent performances than the worst-performing funds. The literature on performance persistence in conventional mutual funds is not conclusive. Some studies do not find general evidence of persistence (see Carhart, 1997; Cuthbertson et al., 2008; Massa and Patgiri, 2009) whereas other have uncovered some evidence of performance persistence in the mutual fund industry (see Brown and Goetzmann, 1995; Elton et al., 1996; Cremers and Petajisto, 2009).

The literature on the persistence of SRI funds is still scant. Leite and Cortez (2013), using Carhart's (1997) methodology, do not find evidence of persistence for a sample of French SRI funds; Lean et al. (2014) draw a similar conclusion—for a sample of SRI funds from the Asia-Pacific region. However, in Lean et al.'s (2015) study, although they found weak evidence of performance persistence for a sample of European and North American SRI funds (also using Carhart's methodology), the evidence is stronger when applying contingency tables. In this vein, Matallín-Sáez et al. (2016) show how, compared to using Carhart's method, contingency tables are biased towards finding evidence of persistence too easily. Unlike some of the previous literature, we do not analyze the persistence for a regional sample but for a global sample of SRI funds by applying a more robust methodology—also based on Carhart (1997). In our study, we contribute by incorporating information on the funds' SR attributes into the persistence analysis. In fact, and as we will see, we obtain that persistence results are different depending on the level of mutual funds' SR scores. Accordingly, the analysis of the performance persistence constitutes the second hypothesis of our study:

Hypothesis 2 (H2): *The differences in the abnormal performance between the best and the worst mutual funds, with similar SR attributes scores, are not persistent over time.*

3. Methodology

3.1. Performance measurement

This section is devoted to a succinct description of the measurement of mutual funds' performance and their persistence, for which we consider a linear model which adjusts each fund's returns for a set of given risk factors. This is a very popular approach in the literature, based on one of its seminal contributions (Jensen, 1968), although a successive number of contributions in the field have proposed some variations of it, in order to include more factors (Fama and French, 2015; Carhart, 1997). Multifactor models attempt to avoid the omitted benchmark bias in the performance evaluation. This bias, as pointed out by Pástor and Stambaugh (2002) and Matallín-Sáez (2006), is present when the performance model is not considered a relevant benchmark that proxies some asset classes in which the mutual fund invests. It must also be noted that the fund's return is the result of passive and active management. The return linked to active management is the value added by managers over the return from passive management. The return of passive management hinges critically on the funds' investment objectives. Thus, considering the characteristics of the mutual funds in the sample, and following Sharpe (1992) and Elton et al. (1993), we propose a multifactor model with specific benchmarks. Given that we aim at evaluating funds with a specific investment strategy as well as a broad geographical scope for investment, we have adopted a linear model such as the following one:

$$r_{p,t} = \alpha_p + \beta_{p,w} r_{w,t} + \beta_{p,s} r_{s,t} + \beta_{p,m} r_{m,t} + \varepsilon_{p,t} \tag{1}$$

In the above expression, $r_{p,t}$ is the excess return over the risk-free asset of the assessed fund, the constant in the model, α_p, measures the fund's abnormal performance, and the risk factors are the excess returns corresponding to: (i) a global benchmark, which represents investment in different markets around the world ($r_{w,t}$); (ii) a specific benchmark, representing investment constrained by SRI fundamentals ($r_{s,t}$); and (iii) a specific benchmark for investment in the emerging markets, taking into account the characteristics of some of the funds being evaluated ($r_{m,t}$).

3.2. Performance persistence measurement

In order to measure performance persistence we will consider the so-called recursive portfolio approach (Carhart, 1997), which is probably the most popular approach in the literature to measure mutual fund performance persistence. Some successful variations of this approach have been proposed by Busse et al. (2010) and, most notably, Fama and French (2010).

Carhart (1997) proposes to evaluate persistence by analyzing the abnormal performance of portfolios that invest according to mutual funds' past performance. Persistence is then calculated for two semiannual (half-yearly) symmetrical windows. The first of these windows estimates past performance, whereas the second one *rebalances* the recursive portfolio. In addition, when estimating performance of a non-overlapping rolling window we will be allowing the model parameters to vary over time. This is an interesting feature, due to the substantial amount of literature on time-varying systematic risk.

Similarly to Abdelsalam et al. (2017), we apply the recursive portfolio approach by means of the following <u>algorithm</u>:

1. In the first step the performance of the SR funds is estimated by means of Equation (1) for the first sample period.

2. SR funds are ranked in increasing order according to the performance achieved in the period in order to form quintiles within each group of funds—according to investment area and SR attributes.

3. At the start of the following period we form five equally weighted portfolios according to quintile past performance, $Q_1, \ldots, Q_5$, where the first portfolio (Q_1) invests in the worst performing funds in the previous period and, conversely, the last portfolio (Q_5) invests in the previous period's best funds. The same investment strategy is followed for the other deciles.

4. This process is repeated at the beginning of each period (i.e., we would restart in step 1). Therefore, each portfolio would represent a *dynamic* investment strategy that *rebalances* selected funds according to their previous performance.

5. We therefore compute the daily return of the five portfolios and then estimate the abnormal performance of the portfolio, also using model (1).

Our hypothesis is therefore that, should persistence in mutual fund performance exist, a portfolio with an investment strategy based on a poor (good) past performance will show a negative (positive) abnormal performance.

4. Data

The data used in this study are from equity mutual funds with SR conditioned investment policies. Specifically, we analyze 3,920 SR equity mutual funds in the world according to investment geographical area. Attending to the representativeness on the global sustainable investment arena, we cover separately the following zones: Europe, US and Canada, and "Other" (including emerging economies) to prevent from any local bias. Morningstar database provides information on daily returns for these funds. The sample period analyzed spans from January 1, 2000 to March 29, 2018. We report some characteristics of the sample funds in Tables 1 through 4.

Table 1 shows how the number of SR funds varies over the sample period because some funds disappear and new ones are incorporated. The net balance of all SR funds since inception shows that the number of funds generally increases. The annual increase in the number of funds was positive over the 2000–2018 period, with an annual average of 11.01%. It is also remarkable the strong impact of the global financial crisis on the clean energy production and the underlying SR funds, shown by the dramatic decrease in the total number of funds from 2009 to 2013 coinciding with downturns and upturns in the economic cycle (1,115 and 1,097, respectively). Although the recovery times have varied around the world, it is evidenced the remarkable overall increase experienced from 2014 to 2018 and reaching an all-time high with a positive inter-annual change of the number of funds with an average value 22.15%. In general, considering the full period 2000-2018, on average, the highest growth rate is mainly concentrated in the category that includes emerging countries, grouped in "Other" with a value of 20.25%, being practically double compared to "Europe" and "US and Canada" (10.06% and 8.56%, respectively).

Since both surviving and non-surviving mutual funds are considered in the study, there is no survivorship bias in the results for performance and persistence. Rohleder et al. (2011) reviews the effects of this bias in mutual fund performance measurement. However, avoiding survivorship bias may also lead to other problems that are not usually addressed in the literature. Specifically, the inclusion of funds with limited data may reduce the ro-

bustness of the analysis. In this regard, Rohleder et al. (2011) pointed out how individual fund performance measurement requires a return history of a certain length to generate reliable regression estimates. In addition, comparing funds with different periods of existence could add some bias if the mutual funds' performance is correlated with the period for which data are available—for instance, the performance could differ depending on the economic cycle or for bull and bear markets (Kacperczyk et al., 2009; Kosowski, 2011; Sun et al., 2013). In order to avoid this type of problem, our empirical strategy will take into account the following: we present performance and persistence results in two ways, i.e. for all mutual funds and for survivor funds only.

Table 2 reports information on the fund sub-types considered in the study according to their survivorship characteristics. We create five categories of mutual funds. First, we split the funds into two sets, survivors (S) and non-survivors (NS). The difference between the two sets is that the first (second) includes all the mutual funds with (without) a net asset value as of March 29th 2018. Considering the number of semesters for which data for mutual funds are available, we distinguish three subgroups for S mutual funds: $S = 36.5$, when the fund shows data for the whole sample period, i.e., 36.5 semesters; $S \geq 4$ for survivor mutual funds with at least four semesters of data, and $S < 4$ for survivors less than four semesters old. Also, considering semesters with data, non-survivor mutual funds are collated into two groups: $NS \geq 4$ comprises mutual funds with at least four semesters of data, and $NS < 4$, the rest.

As indicated in Table 2, only 4.59% (180/3,920) of the SR funds have complete data over the sample period. The largest group is that denoted by $S \geq 4$, specifically 42.78% (1,677/3,920) for the funds. Table 2 also shows the survival characteristics of the mutual fund sample according to their Morningstar SR attributes scores as provided by the Morningstar mutual fund database. Specifically, we consider the following portfolio SR attributes: Sustainability, Environmental, Social, Governance and Undefined. The first one, Portfolio Sustainability Score combines a fund's Portfolio ESG Score and its Portfolio Controversy Score jointly; the second, Environmental Score is based on the asset-weighted average of the Sustainalytics Environmental Scores for the covered holdings in its portfolio; the third one, Social Score based on the asset-weighted average of the Sustainalytics Social Scores for the covered holdings in its portfolio; the fourth, Governance Score is based on the asset-weighted average of the Sustainalytics Governance Scores for the covered holdings in its portfolio; and finally, fifth, Undefined, covers funds that do not focus on specific

SR attributes. Then, for each category the funds with the 30% highest/lowest portfolio SR attributes scores are classified as High/Low. Thus, we group top and bottom 30% of funds with "High" and "Low" scores, respectively, and the middle 40% (between 30% and 70%) are categorized as "Mid". Interestingly, there are noticeable differences in the areas under analysis and Europe appears to concentrate the highest scores, while the lowest lie in "Other". For the US and Canada case the vast majority of funds are classified as "Mid", whereas "Low" outnumber "High".

Table 3 reports some summary statistics corresponding to the mutual funds' sample. Regarding geographical area of investment, most funds focus on Europe, the US and Canada; specifically, 77.09% (3,022/3,920). A mean-variance analysis reveals that in Europe funds with "High" scores perform better since, on average, they show a higher net return and lower risk. And also notable is that this finding is revealed for the four SR attributes. However, for the US and Canada this evidence is not as clear, although Low scores concentrate higher risk, the return is mixed depending on the attributes observed. And contrary to Europe, the "Other" group reports a better balance for the "Low" scores.

By connecting the previous evidence from Table 2 where the number of predominant funds is high (low) for Europe ("Other"), Table 3 brings insights in regard the risk-return trade-off and reinforces the argument in favor of the maturity of Europe to be a more developed and experienced area in the SR management industry. Nevertheless, the "Other" group does not appear to follow adequate SR investment strategies since results are worse than in the case of "Low" scored funds. These descriptive statistics give an idea of both the segmentation and disparate evolution of the SR fund industry in different locations.

As mentioned in the methodology section, to evaluate mutual fund performance we apply the linear model (1) where funds' excess returns are adjusted to the excess returns corresponding to the types of assets in which the funds invest. Note that because the analyzed funds invest in very different geographical areas, the first benchmark is a global index representing global investments, specifically the FTSE World. We selected the DJ Sustain World to represent investments under SR conditions. A number of funds invest in less mature and developed markets, so we also included the FTSE index for emerging markets. For these indices we calculate daily returns from information provided by Morningstar. We compute the excess return using the one-month Treasury bill rate as the risk-free asset.[4]

---

[4]Obtained from Ken French's website (mba.tuck.dartmouth.edu/pages/faculty/ken.french/, accessed March 2018).

Table 4 reports some summary statistics for the benchmarks used in expression (1). For the analyzed sample period, the most globalized indexes (for which financial markets in more advanced economies weigh more) show a more conservative mix of average return and risk than those for emerging markets (FTSE Emerging) for which there is higher risk and average return.

5. Results

5.1. SRI mutual fund performance

In Table 5 we report results on funds' performance, not only for all funds (last row of the table), and also split for Europe, US and Canada, and "other". Within each of these three categories we also group results considering fund sub-categories and according to portfolio SR attributes: Sustainability, Environmental, Social, Governance and Undefined. The table is also split vertically in two panels, the left-hand panel indicating the percentage of funds in each category and sub-category for which results were either positive or negative (and whether these results were significant or not), and the right-hand panel reporting the annualized performance, considering both unweighted and weighted (by fund size and fund life) averages.

Unlike Tables 2 and 3 the grouping criterion in "High", "Mid" or "Low" attributes scores is not made on the whole of the totality of the funds, but within each geographical area (Europe, Usa and Canada, and "Other"). This procedure is due the lack of uniformity of the scores within the geographical zones. For instance, in Europe there were more "High" funds and in "Other" more "Low" funds. As one of the objectives of the study is to compare the performance of the funds based on their higher or lower score, there could be some bias in this comparison (i.e., if we maintained the previous grouping method, since comparing "High" vs. "Low" could mean comparing in some way Europe with "Other"). Thus, if there is a specific component in these geographical areas that is not captured by the linear model that adjusts to systematic risk sources, the High-Low comparison could be biased by that component.

Results indicate that, on average, the number of funds with significant performance different from zero is small. For all the funds, in the bottom row of Table 5, the percentage of funds with negative performance was slightly higher than that corresponding to funds with positive performance (53.34% vs. 46.66%). However, the percentage of funds for which

results were significant was very low in both cases—5.46% and 3.27%, respectively. In addition, the funds do not outperform the market, there is an overall unweighted negative performance (-0.16%). This result is not very different from that achieved by most of the literature that has evaluated investment funds: a negative performance near zero is evident.

Results differ depending on the geographical investment area. On average, they are particularly bad for US and Canada (-0.85%), whereas for Europe they were only slightly negative (-0.46%). In contrast, for "Other", the average performance was positive (1.58%). In view of these results, we can infer that more mature markets (USA and Canada, and Europe) involve some complexity for managers to beat the market, while in emerging markets, perhaps less efficient, there may be more opportunities for managers to provide greater added value to their management.

In contrast, looking at the weighted average performance by fund size, we observe that the total performance improves notably in Europe (0.69%) and in the USA and Canada (0.66%), so the negative average seems to be driven by the smaller funds. For funds in the "Other" area the weighted average by size is reduced, taking the highest value instead (1.14%). Therefore, part of the differences found between the different zones is partly due to the behavior of smaller funds; in Europe and, US and Canada there are more negative results while in "Other" area better results are found. It should be borne in mind that smaller funds have a greater capacity for active management than funds with large assets that implicitly approximate the average behavior of the market, given that they are an important part of it (Sharpe, 1992). In line with the above mentioned, it seems that more mature markets, such as Europe and, the US and Canada, do not provide value-added opportunities to these smaller funds, which can be inferred as a measure of the efficiency of that market. However, for funds from "Other" zones there are such opportunities.

5.2. Performance and SR attributes' scores

With the aim of testing the first hypothesis of this study, we now focus on the performance differences of funds with different levels of SR attributes. The specific comparisons between funds grouped according to attributes scores are reported in Table 6. In general, the sign of the difference between the performance of funds with higher and lower scores is negative, which indicates that on average the performance of funds with "High" scores is worse than those with "Low" scores. This evidence is clearly significant for the "Other" zone, and in most cases for the US and Canada. Whereas in Europe the sign of the difference

in performance is generally negative, in no case it is significant. After the diagnosis of the three geographic areas, we find a common pattern and this is an interesting finding that contributes to the SR literature: there are differences in performance among Sustainability, Environmental, Social and Governance attributes scores, in that the higher (lower) the SR attributes score, the worst (best) its performance. We can therefore reject the first hypothesis proposed in Section 2. In other words, we can infer that funds with weaker socially-conscious schemes outperform those with a strong SR orientation. This result is consistent with Muñoz et al. (2014) for a sample of US and European socially responsible mutual funds in which underperformance is identified.

In this line, Silva and Cortez (2016) expanded the analysis of US and European towards green funds and their results highlight the idea that green funds present a tendency to underperform the benchmark. One explanation for these findings draws on the tenets of modern portfolio theory, which holds that the negative financial consequences from investing in socially responsible screened companies are due to increased information costs and the inherent difficulty of diversifying ethical fund portfolios (Martí-Ballester, 2015); both motives may lead to poorer risk-adjusted performance mainly for the US and Canada, but also in "Other" areas—as observed. However, concerning the general relationship between financial and socially responsible performance from the Morningstar SR attributes scores approach, the lack of research identified means this remains an open question.

Figures 1–4 provide a clearer view of the tendencies. Performance varies depending on the investment area as well as the attributes' score, as commented on in the preceding paragraphs. However, the figure also reveals some hitherto undetected features. Of particular interest is the higher dispersion found for funds investing in "other" areas (Figures 1c, 2c, 3c and 4c), evidenced in the way probability mass is more spread along the OX axis, regardless of the attributes' scores—sustainability (Figure 1c), environmental (Figure 2c), social (Figure 3c) or governance (Figure 4c). In contrast, for funds investing in Europe (Figures 1a, 2a, 3a and 4a) and US and Canada (Figures 1b, 2b, 3b and 4b) the performance of the funds in each category is more alike, particularly for the latter, as shown by the much tighter densities. However, in the case of Europe and US and Canada, despite the higher homogeneity, densities reveal that some funds perform particularly well, as shown by some bumps in the upper end of the distributions—although some pockets of poor performance also exist, as revealed by some (fewer) bumps in the lower tail of the distribution. This extreme behavior is also present in the "Other" geographical investment area, and

14

particularly for the "Low" category.

In this sense, some specific "Low" funds perform particularly badly in Europe, as shown by the lower tail stretching beyond -0.3 for Europe funds (Figures 1a, 2a, 3a, 4a), whereas for US and Canada funds (Figures 1b, 2b, 3b and 4b) and "Other" (Figures 1c, 2c, 3c and 4c) this phenomenon is less pronounced. However, at the other extreme, the best funds are more frequently found in the US and Canada category, as shown by the relatively long tails corresponding to the solid and dashed lines in Figures 1b, 2b, 3b and 4b. This overperformance is also found for the "Other" category, and particularly for "Low" funds. This result highlights the asymmetry in the performance differences according to "High"/"Low" categories—i.e., the main driver of the differences found is the behavior of the worst funds.

Table 7 reports analogous results to Table 6 for survivor funds only—those surviving for at least four semesters. The survivor-only counterparts to the densities shown in Figures 1–4 are reported in Figures 5–8. Therefore, the tables and the figures give the analysis additional robustness. The total number of surviving funds is 1,677, representing 43% of total funds in our sample. The comparison of results between all funds (Table 6, Figures 1–4) and nonsurviving funds only (Table 7, Figures 5–8) indicates that, in general, the tendencies are maintained, since the sign of the difference in annualized performance is virtually the same in most instances—all of them except Portfolio Environmental in US and Canada, for which the sign is reversed. However, results for this particular comparison are not significant. Therefore, it seems that the results achieved in Table 6 and Figure 1–4 are robust in the sense that they have not been driven by the performance of non-surviving funds.

5.3. Performance persistence and SR attributes' scores

We will now analyze whether funds' performance *persists* over time in order to test the second hypothesis of the study. As explained in the methodology section, we applied the recursive portfolio approach (Carhart, 1997). Then, we form portfolios which follow investment strategies based on past performance. Should performance exist, we expect that the quintile-portfolio that invests in the worst (best) funds in the past obtains a worst (best) performance. Results are shown in Figure 9 and Table 8 for all funds, and Figure 10 and Table 9 for surviving funds only.

Figure 9 shows, for funds grouped according to SR attributes scores and investment

15

area, the performance of the quintile-portfolios. Should persistence exist then we will expect lines with positive slopes. In general, this positive slope exists in most cases and the difference $Q5 - Q1$ is positive, which means that investing in the best funds ($Q5$) in the past provides better results than investing in the worst funds ($Q1$). These results support performance persistence evidence. More specifically, Table 8 reports the performance of the extreme quintile-portfolios ($Q1$ and $Q5$), and the magnitude and significance of the difference between both of them. For most of the cases being analyzed the difference is positive, implying that investing in the best past funds yields better performance the worst past funds. However, for the Europe zone not in all cases it occurs with the same intensity and significance. Specifically, for funds with "High" scores in no case there is significance, in fact the difference $Q5 - Q1$ takes smaller values—even negative. On the contrary, the persistence takes high and significant values for the funds with "Low" scores. Thus, investing in funds with lower scores that performed better in the past provides a differential with respect to the worst between 5.64% per annum in the case of the Social attribute and 5.01% in the Governance attribute. This differential is mainly due to the worst performance of the funds in the first quintile, which ranges from -4.24% for the Sustainability attribute and -3.65% for the Environmental attribute. That is, funds with Low scores have a clear persistence in their bad results. This evidence also holds, in general, for surviving funds (Table 9 and Figure 10).

In the case of the US and Canada, previous finding is still maintained as funds with "High" scores are the ones with the lowest values of difference $Q5 - Q1$, and in no case they are significant. Unlike the previous analysis with Europe, the groups of funds with greater persistence and significance are the "Mid" over the "Low". Finally for the "Other" zone, those funds with High scores are the ones that have the highest persistence and significance. Thus, the difference in the performance of $Q5 - Q1$ takes values that oscillate between 7.24% for the Environmental attribute and 4.16% for the Governance. These results are especially driven by the high performance of funds in $Q5$. For the "Mid" and "Low" funds although in some cases the difference is positive and weakly significant, in others it takes negative but not significant values (Social and Governance Scores).

In conclusion, we find greater persistence in Europe and the US and Canada than in "Other", leading us to reject the second hypothesis of the study (H2, see Section 2). In other words, the differences in the abnormal performance between the best and the worst funds with similar SR attribute scores persist over time. In particular, it is observed that

"Low" funds are the ones with the greatest persistence, while in "Other" those scored as High are the predominant.

6. Conclusions and implications for theory and practice

6.1. Conclusions

The performance of SR funds has been and continues to be an issue of interest worldwide. The popularity and reputation that these ethically oriented assets have been experiencing for decades brings to academics and investors a wide variety of research and investment opportunities, respectively. For both groups the issue as to the likely underperformance of SRI when compared to their conventional peers has always been at stake, and contributions in this particular field are diverse, being this question far from consensus.

The fact that SR investors are less sensitive to poor past performance suggests that social, environmental or ethical issues are prioritized in their investment decisions. This approach looks beyond the economic dimension and the mere pursuit of competitive returns to encompass the investment strategy as non-financial attributes are taken into consideration. This point, coupled with the evidence about its positive impact on societal concerns with no extra charges involved when integrating ethics or moral values into the investment choices justifies its extraordinary development and relevance within the international investment arena.

Our study analyzes the financial performance of a sample of mutual funds around the world that integrate different levels of SR attributes in their investment portfolios. These attributes are related to several concerns, such as environmental issues, governance, social matters, and sustainability. Our results lead us to reject the first study hypothesis, indicating that most of the mutual funds achieving higher scores in these SR attributes, especially those investing in less-developed areas, experience lower overall performances than other similar funds that focus less on their SR strategies. This is in line with the higher constraints these funds experience in their investments, in relation to other SR funds characterized with lower SR grades.

The financial value added by fund managers, however, is essential in relation to the performance experienced in the fund portfolio, as shown in previous studies analyzing conventional funds. Then, assuming that funds that achieve greater extents of SRI in their portfolios will obtain lower performances is a distorted conclusion. In this line, we show

17

that their overall underperformance is mainly driven by the worst-performing funds in each SR category.

A carefully analysis of SR fund performance over time leads us to reject the hypothesis related to the lack of persistence of SR fund performance by showing that some skilled managers are able to persistently provide investors with greater risk-adjusted returns. For instance, investing in the previous worst-performing funds usually leads to negative alphas, while investing in the previous best funds, whose managers are likely to show greater managerial skills, results in positive risk-adjusted returns. Moreover, the performance differences between the previous best and the previous worst funds are significantly positive for most of the regions and SR categories analyzed.

In sum, the evidence reported in this study shows that managers, practitioners and individual investors could obtain higher risk-adjusted returns by allocating their investments to the previous best-performing funds that integrate specific levels of SR criteria in their portfolios.

It follows, therefore, that a potential SR investor should be more careful when picking the SR funds for investing in. Given that some SR funds experience higher constraints in their investment decisions, the role of managers is then crucial to overcome these restrictions and to add value through their management. Hence, investors and managers willing to incorporate higher levels of SR criteria in their portfolios without experiencing worse risk-adjusted returns should invest in the previous best-performing funds in order to achieve both goals. In other words, pursuing higher ethical values in our portfolios should not imply a worsening in our financial objectives if we choose the right SR funds to invest in.

6.2. Implications for theory and practice

This study highlights the relevance of distinguishing between SR funds achieving higher and lower SR attributes in their portfolios, such as environmental or sustainability characteristics, among others. Overall, we find that funds achieving higher scores on their SR attributes seem to underperform their non-SR intensively oriented counterparts. This evidence is, however, mainly driven by the overall behavior of the worst-performing funds, and it is not associated with the sustainable levels accomplished in the funds' portfolios. Accordingly, several funds with high SR scores do experience greater risk-adjusted returns over time, implying that assessing correctly the financial performance is essential

to improve the SR funds' results. These findings have several theoretical and practical implications.

For instance, the characteristics and specific risks related to the portfolio can determine the overall returns of a mutual fund. This is especially relevant for funds aiming to invest in sustainable assets to a greater extent, achieving higher socially-conscious objectives in their investments. Therefore, managers and investors should consider the specific features and SR scores related to each portfolio when assessing the relative performance of their funds; otherwise, their results and conclusions will be biased, potentially affecting the consistency and adequacy of their subsequent decisions.

Moreover, this study also highlights the importance of SR funds in the financial sector. As shown throughout the paper, several funds oriented to high levels of sustainability and SR treats are shown to persistently provide investors with greater performances. Accordingly, the access to information related to socially-responsible and financial features of these investment vehicles plays an important role in improving the investment decisions of all individuals who participate in the market. Therefore, the efforts made by policymakers must be focused towards an enhancement of the transparency, quality and availability of the data reported in the SR mutual fund industry. This implication would lead managers and investors to optimize their financial and sustainable investment decisions.

Finally, it should be noted that our study only assesses the financial performance of funds with specific levels of environmental, sustainable and other SR attributes in their portfolios. Nonetheless, this perspective also offers many avenues for future research lines, especially those related to managerial perceptions. Some examples would be to assess managers' abilities to detect "green" opportunities, or determine the sensitivity of inflows and outflows in SR portfolios to new environmental information. Moreover, given the evidence from this study about the persistence in the abnormal performance of socially-conscious portfolios, developing sustainable performance measures that combine both ethical and financial criteria in the mutual fund industry is a potentially key topic for academics, investors, and other stakeholders who want a better understanding of the behavior of sustainable portfolios.

References

Abdelsalam, O., Duygun, M., Matallín, J. C., and Tortosa-Ausina, E. (2017). Is ethical money sensitive to past returns? The case of portfolio constraints and persistence of Islamic funds. *Journal of Financial Services Research*, 51(3):363–384.

Bauer, R., Koedijk, K., and Otten, R. (2005). International evidence on ethical mutual fund performance and investment style. *Journal of Banking & Finance*, 29(7):1751–1767.

Benson, K. L. and Humphrey, J. E. (2008). Socially responsible investment funds: Investor reaction to current and past returns. *Journal of Banking & Finance*, 32(9):1850–1859.

Bollen, N. and Busse, J. (2001). On the timing ability of mutual fund managers. *Journal of Finance*, 56(3):1075–1094.

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

Bollen, N. P. B. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, 42(03):683–708.

Brown, S. J. and Goetzmann, W. N. (1995). Performance persistence. *Journal of Finance*, 50:679–698.

Busch, T., Bauer, R., and Orlitzky, M. (2016). Sustainable development and financial markets: Old paths and new avenues. *Business & Society*, 55(3):303–329.

Busse, J. and Irvine, P. (2006). Bayesian alphas and mutual fund persistence. *Journal of Finance*, 61(5):2251–2288.

Busse, J. A., Goyal, A., and Wahal, S. (2010). Performance and persistence in institutional investment management. *Journal of Finance*, 65(2):765–790.

Cai, W. and Li, G. (2018). The drivers of eco-innovation and its impact on performance: Evidence from China. *Journal of Cleaner Production*, 176:110–118.

Capelle-Blancard, G. and Monjon, S. (2012). Trends in the literature on socially responsible investment: Looking for the keys under the lamppost. *Business Ethics: a European Review*, 21(3):239–250.

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

Chang, C. E., Nelson, W. A., and Witte, H. D. (2012). Do green mutual funds perform well? *Management Research Review*, 35(8):693–708.

Climent, F. and Soriano, P. (2011). Green and good? The investment performance of US environmental mutual funds. *Journal of Business Ethics*, 103(2):275–287.

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

Cremers, M., Petajisto, A., and Zitzewitz, E. (2013). Should benchmark indices have alpha? Revisiting performance evaluation. *Critical Finance Review*, 2(1):1–48.

Cuthbertson, K., Nitzsche, D., and O'Sullivan, N. (2008). UK mutual fund performance: Skill or luck? *Journal of Empirical Finance*, 15(4):613–634.

Elton, E. J., Gruber, M. J., and Blake, C. R. (1996). The persistence of risk-adjusted mutual fund performance. *Journal of Business*, 69(2):133–157.

Elton, E. J., Gruber, M. J., Das, S., and Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies*, 6(1):1–22.

Fama, E. F. and French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance*, 65(5):1915–1947.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Friede, G., Busch, T., and Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4):210–233.

Galema, R., Plantinga, A., and Scholtens, B. (2008). The stocks at stake: Return and risk in socially responsible investment. *Journal of Banking & Finance*, 32(12):2646–2654.

Gottesman, A. A. and Morey, M. R. (2007). Predicting emerging market mutual fund performance. *Journal of Investing*, 16(3):111–122.

Helminen, R.-R. (2000). Developing tangible measures for eco-efficiency: the case of the Finnish and Swedish pulp and paper industry. *Business Strategy and the Environment*, 9(3):196.

Huang, J., Sialm, C., and Zhang, H. (2011). Risk shifting and mutual fund performance. *Review of Financial Studies*, 24(8):2575–2616.

Humphrey, J. E., Lee, D. D., and Shen, Y. (2012). Does it cost to be sustainable? *Journal of Corporate Finance*, 18(3):626–639.

Jansson, M. and Biel, A. (2014). Investment institutions' beliefs about and attitudes toward Socially Responsible Investment (SRI): A comparison between SRI and non-SRI management. *Sustainable Development*, 22(1):33–41.

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

Jin, J. and Han, L. (2018). Assessment of Chinese green funds: Performance and industry allocation. *Journal of Cleaner Production*, 171:1084–1093.

Kacperczyk, M., Van Nieuwerburgh, S., and Veldkamp, L. (2009). Rational attention allocation over the business cycle. Working Paper 15450, National Bureau of Economic Research, Cambridge, MA.

Kempf, A. and Osthoff, P. (2007). The effect of socially responsible investing on portfolio performance. *European Financial Management*, 13(5):908–922.

Koellner, T., Weber, O., Fenchel, M., and Scholz, R. (2005). Principles for sustainability rating of investment funds. *Business Strategy and the Environment*, 14(1):54–70.

Kosowski, R. (2011). Do mutual funds perform when it matters most to investors? US mutual fund performance and risk in recessions and expansions. *Quarterly Journal of Finance*, 1(3):607–664.

Laurel-Fois, D. (2016). Beyond appearances the risk-reducing effects of responsible investment practices. *Business & Society*, forthcoming.

Lean, H. H., Ang, W. R., and Smyth, R. (2014). Performance persistence of socially responsible investment funds in the Asia Pacific region. In Gregoriou, G. and David, L., editors,

Handbook of Asian finance REITs, trading, and fund performance, volume 2, pages 377–392. Academic Press Inc.

Lean, H. H., Ang, W. R., and Smyth, R. (2015). Performance and performance persistence of socially responsible investment funds in Europe and North America. *The North American Journal of Economics and Finance*, 34:254–266.

Leite, P. and Cortez, M. C. (2013). Performance and performance persistence of European socially responsible funds: French evidence. Technical report, School of Management, Polytechnic Institute of Cavado and Ave., Portugal.

Mallett, J. and Michelson, S. (2010). Green investing: is it different from socially responsible investing? *International Journal of Business*, 15(4):395–410.

Martí-Ballester, C. P. (2015). Can socially responsible investment for cleaner production improve the financial performance of Spanish pension plans? *Journal of Cleaner Production*, 106:466–477.

Massa, M. and Patgiri, R. (2009). Incentives and mutual fund performance: Higher performance or just higher risk taking? *Review of Financial Studies*, 22(5):1777–1815.

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

Matallín-Sáez, J. C., Soler-Domínguez, A., and Tortosa-Ausina, E. (2016). On the robustness of persistence in mutual fund performance. *The North American Journal of Economics and Finance*, 36:192–231.

Muñoz, F., Vargas, M., and Marco, I. (2014). Environmental mutual funds: Financial performance and managerial abilities. *Journal of Business Ethics*, 124(4):551–569.

Nofsinger, J. and Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance*, 48:180–193.

O'Rourke, A. (2003). The message and methods of ethical investment. *Journal of Cleaner Production*, 11(6):683–693.

Palacios-González, M. M. and Chamorro-Mera, A. (2018). Analysis of the predictive variables of the intention to invest in a socially responsible manner. *Journal of Cleaner Production*, 196:469–477.

Pástor, L. and Stambaugh, R. F. (2002). Investing in equity mutual funds. *Journal of Financial Economics*, 63(3):351–380.

Renneboog, L., Ter Horst, J., and Zhang, C. (2008a). Socially responsible investments: Institutional aspects, performance, and investor behavior. *Journal of Banking & Finance*, 32(9):1723–1742.

Renneboog, L., Ter Horst, J., and Zhang, C. (2008b). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3):302–322.

Renneboog, L., Ter Horst, J., and Zhang, C. (2011). Is ethical money financially smart? Nonfinancial attributes and money flows of socially responsible investment funds. *Journal of Financial Intermediation*, 20(4):562–588.

Rohleder, M., Scholz, H., and Wilkens, M. (2011). Survivorship bias and mutual fund performance: Relevance, significance, and methodical differences. *Review of Finance*, 15(2):441–474.

Said, R. M., Sulaiman, M., and Ahmad, N. N. N. (2013). Do fund managers perceive environmental information useful? An empirical study from Malaysia. *Journal of Cleaner Production*, 52:281–288.

Sharpe, W. F. (1992). Asset allocation: management style and performance measurement. *Journal of Portfolio Management*, 18(2):7–19.

Sheather, S. J. and Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Ser.B,53(3):683–690.

Silva, F. and Cortez, M. C. (2016). The performance of US and European green funds in different market conditions. *Journal of Cleaner Production*, 135:558–566.

Sun, Z., Wang, A., and Zheng, L. (2013). Only when the tide goes out: Downside returns and hedge fund performance. Working paper, SSRN. Available at SSRN: http://ssrn.com/abstract=2249033 or http://dx.doi.org/10.2139/ssrn.2249033.

Table 1: Evolution of the number of SR funds, 2000–2018[a]

| Year | Number of funds | | | | % Change | | | |
| | Morningstar equity category | | | | Morningstar equity category | | | |
	Europe	US and Canada	Other	Total	Europe	US and Canada	Other	Total
2000	216	229	40	485	—	—	—	—
2001	255	266	56	577	18.06	16.16	40.00	18.97
2002	285	275	65	625	11.76	3.38	16.07	8.32
2003	308	274	69	651	8.07	–0.36	6.15	4.16
2004	325	286	80	691	5.52	4.38	15.94	6.14
2005	355	314	90	759	9.23	9.79	12.50	9.84
2006	401	345	100	846	12.96	9.87	11.11	11.46
2007	435	438	142	1,015	8.48	26.96	42.00	19.98
2008	467	505	162	1,134	7.36	15.30	14.08	11.72
2009	469	494	152	1,115	0.43	–2.18	–6.17	–1.68
2010	487	497	156	1,140	3.84	0.61	2.63	2.24
2011	505	487	150	1,142	3.70	–2.01	–3.85	0.18
2012	508	473	138	1,119	0.59	–2.87	–8.00	–2.01
2013	604	535	314	1,097	18.90	13.11	127.54	–1.97
2014	654	653	365	1,672	8.28	22.06	16.24	52.42
2015	730	796	454	1,980	11.62	21.90	24.38	18.42
2016	805	963	540	2,308	10.27	20.98	18.94	16.57
2017	928	1200	737	2,865	15.28	24.61	36.48	24.13
2018	925	1193	725	2,843	–0.32	–0.58	–1.63	–0.77
Mean % change 2000–2018					8.56	10.06	20.25	11.01
Mean % change 2000–2008					10.18	10.68	19.73	11.32
Mean % change 2009–2013					5.49	1.33	22.43	–0.65
Mean % change 2014–2018					9.03	17.79	18.88	22.15

[a] Note, number of funds measured on December 31, except for 2018 that is measured on March 29.

Table 2: Survival characteristics of the mutual funds in the sample

Morningstar equity category	Morningstar SR attributes scores		$S = 36.5$[a]	$S \geq 4$[b]	$S < 4$[c]	$NS \geq 4$[d]	$NS < 4$[e]	Total
Europe	Portfolio Sustainability	High	52	316	171	70	17	626
		Mid	36	150	11	64	1	262
		Low	8	36	2	8	2	56
	Portfolio Environmental	High	48	294	160	64	16	582
		Mid	42	120	16	59	1	238
		Low	6	88	8	19	3	124
	Portfolio Social	High	42	320	169	63	17	611
		Mid	39	118	11	47	1	216
		Low	15	64	4	32	2	117
	Portfolio Governance	High	52	317	169	65	17	620
		Mid	33	145	12	61	1	252
		Low	11	40	3	16	2	72
	Undefined		5	91	47	182	35	360
	Total category		101	593	231	324	55	1304
US and Canada	Portfolio Sustainability	High	8	78	61	32	5	184
		Mid	28	345	268	84	38	763
		Low	35	182	104	87	9	417
	Portfolio Environmental	High	10	104	67	52	11	244
		Mid	34	331	257	71	27	720
		Low	27	170	109	80	14	400
	Portfolio Social	High	10	78	70	39	4	201
		Mid	33	376	270	93	43	815
		Low	28	151	93	71	5	348
	Portfolio Governance	High	11	58	64	35	0	168
		Mid	30	390	263	89	43	815
		Low	30	157	106	79	9	381
	Undefined		1	51	32	233	37	354
	Total category		72	656	465	436	89	1,718
Other	Portfolio Sustainability	High	1	29	13	24	2	69
		Mid	3	89	46	11	7	156
		Low	1	196	175	23	19	414
	Portfolio Environmental	High	0	22	18	8	2	50
		Mid	4	98	81	27	13	223
		Low	1	194	135	23	13	366
	Portfolio Social	High	0	39	6	22	1	68
		Mid	4	69	51	12	13	149
		Low	1	206	177	24	14	422
	Portfolio Governance	High	2	52	9	24	4	91
		Mid	2	49	41	12	5	109
		Low	1	213	184	22	19	439
	Undefined		2	114	56	66	21	259
	Total category		7	428	290	124	49	898
	All funds		180	1,677	986	884	193	3,920

[a] $S = 36.5$: total survivor (36.5 semesters).
[b] $S \geq 4$: mature survivor (it has value for $s36.5$), with at least four semesters.
[c] $S < 4$: new survivor (it has value for $s36.5$), with less than four semesters.
[d] $NS \geq 4$: no survivor (no value for $s36.5$), with at least four semesters.
[e] $NS < 4$: no survivor (no value for $s36.5$), with less than four semesters.

Table 3: Mutual fund summary statistics, 2000–2018

Morningstar equity category	Morningstar SR attributes scores		Number of funds	Average annual-ized net return (%)	Average annual-ized s.d. (%)	Average size (US$)
Europe	Portfolio Sustainability	High	626	7.87	16.16	63,493,482
		Mid	262	6.40	17.78	125,907,792
		Low	56	4.05	20.42	83,442,207
	Portfolio Environmental	High	582	7.89	16.03	84,810,779
		Mid	238	6.76	17.90	86,621,608
		Low	124	5.06	18.83	51,478,513
	Portfolio Social	High	611	7.86	16.03	77,698,530
		Mid	216	6.56	18.15	101,281,895
		Low	117	5.17	18.86	56,924,132
	Portfolio Governance	High	620	7.97	16.21	66,753,485
		Mid	252	6.56	17.27	110,411,021
		Low	72	3.23	21.08	103,147,120
	Undefined		360	3.90	18.58	77,112,319
	Total category		1,304	6.31	17.34	80,063,077
US and Canada	Portfolio Sustainability	High	184	8.49	13.55	34,120,169
		Mid	763	9.60	13.90	71,724,629
		Low	417	8.67	16.85	90,319,640
	Portfolio Environmental	High	244	8.73	14.04	53,402,803
		Mid	720	10.02	13.98	83,011,176
		Low	400	7.89	16.60	62,614,355
	Portfolio Social	High	201	9.09	13.56	36,635,807
		Mid	815	9.34	14.21	78,871,351
		Low	348	8.80	16.72	76,265,091
	Portfolio Governance	High	168	8.03	14.64	34,312,381
		Mid	815	9.13	13.91	78,650,540
		Low	381	9.74	16.62	73,956,266
	Undefined		354	1.90	21.83	30,345,123
	Total category		1,718	7.67	16.21	64,080,875
Other	Portfolio Sustainability	High	69	9.93	22.02	29,799,194
		Mid	156	10.27	16.89	31,846,150
		Low	414	12.90	15.03	32,205,593
	Portfolio Environmental	High	50	10.15	18.62	27,641,400
		Mid	223	11.84	18.71	25,385,487
		Low	366	12.24	14.40	36,594,742
	Portfolio Social	High	68	7.79	22.00	22,218,351
		Mid	149	10.56	16.50	38,225,058
		Low	422	13.09	15.21	31,095,988
	Portfolio Governance	High	91	9.02	21.89	23,820,437
		Mid	109	10.87	15.46	40,431,008
		Low	439	12.80	15.26	31,256,875
	Undefined		259	8.55	17.79	32,691,782
	Total category		898	10.96	16.68	32,068,387
	All funds		3,920	7.97	16.70	62,086,961

Table 4: Summary statistics for the benchmarks

Factors	Average annualized net return	Average annualized s.d.
FTSE World	6.04%	16.15%
DJ Sustain World NR USD	4.81%	17.80%
FTSE Emerging TR USD	9.46%	19.22%

Table 5: Performance

Morningstar equity category	Morningstar SR attributes scores		Number of funds	Percentage of total number of funds in group				Annualized performance		
				< 0	p-value ≤ 0.05	> 0	p-value ≤ 0.05	Mean (un-weighted average)	Weighted average by fund size	Weighted average by fund life
Europe	Portfolio Sustainability	High	280	46.43%	1.07%	53.57%	3.21%	0.66%	1.40%	1.84%
		Mid	380	48.42%	1.84%	51.58%	4.21%	0.04%	1.39%	1.29%
		Low	284	44.37%	7.04%	55.63%	8.45%	0.17%	1.21%	0.28%
	Portfolio Environmental	High	279	50.18%	2.15%	49.82%	2.51%	0.45%	0.77%	1.31%
		Mid	379	50.40%	1.58%	49.60%	4.22%	-0.20%	1.37%	1.17%
		Low	286	38.11%	6.29%	61.89%	9.09%	0.69%	1.59%	0.70%
	Portfolio Social	High	282	49.65%	0.00%	50.35%	3.55%	0.10%	1.12%	1.63%
		Mid	378	49.47%	2.65%	50.53%	4.23%	0.19%	1.63%	1.25%
		Low	284	39.79%	7.04%	60.21%	8.10%	0.51%	0.96%	0.48%
	Portfolio Governance	High	280	47.14%	0.00%	52.86%	1.07%	-0.16%	1.92%	1.79%
		Mid	377	48.54%	3.18%	51.46%	6.63%	0.60%	1.15%	1.43%
		Low	287	43.55%	6.27%	56.45%	7.32%	0.22%	1.01%	0.18%
	Undefined		360	67.50%	8.33%	32.50%	1.67%	-2.35%	-1.26%	-1.44%
	Total category		1304	52.38%	4.60%	47.62%	4.22%	-0.46%	0.69%	0.44%
US and Canada	Portfolio Sustainability	High	394	66.50%	10.91%	33.50%	0.00%	-1.06%	-0.94%	-1.37%
		Mid	557	60.50%	6.46%	39.50%	2.87%	-0.68%	1.27%	-0.68%
		Low	413	49.15%	4.12%	50.85%	4.36%	-0.03%	1.30%	-0.21%
	Portfolio Environmental	High	404	63.86%	8.66%	36.14%	0.00%	-0.70%	-0.38%	-0.85%
		Mid	545	56.15%	6.97%	43.85%	3.30%	-0.16%	0.85%	-0.58%
		Low	415	57.35%	5.54%	42.65%	3.86%	-1.05%	1.70%	-0.58%
	Portfolio Social	High	406	59.61%	10.34%	40.39%	0.00%	-0.23%	-0.30%	-1.04%
		Mid	545	64.40%	7.16%	35.60%	3.12%	-1.37%	1.00%	-0.91%
		Low	413	50.61%	3.63%	49.39%	4.12%	0.08%	1.59%	-0.14%
	Portfolio Governance	High	404	67.57%	11.14%	32.43%	3.22%	-1.19%	-0.65%	-1.26%
		Mid	548	57.30%	5.84%	42.70%	0.91%	-1.05%	1.64%	-0.45%
		Low	412	52.18%	4.61%	47.82%	3.88%	0.61%	0.65%	-0.39%
	Undefined		354	66.95%	11.58%	33.05%	2.82%	-1.85%	-0.90%	-1.47%
	Total category		1718	60.48%	7.97%	39.52%	2.56%	-0.85%	0.66%	-0.82%
Other	Portfolio Sustainability	High	190	40.00%	3.16%	60.00%	3.68%	0.80%	2.54%	0.86%
		Mid	255	52.94%	3.53%	47.06%	1.96%	0.79%	-0.02%	-0.05%
		Low	194	34.54%	0.00%	65.46%	5.67%	2.96%	0.60%	3.46%
	Portfolio Environmental	High	188	38.30%	3.19%	61.70%	2.13%	1.57%	1.46%	0.77%
		Mid	249	46.99%	1.20%	53.01%	1.61%	0.51%	1.02%	1.09%
		Low	202	44.06%	2.97%	55.94%	7.43%	2.49%	0.14%	1.43%
	Portfolio Social	High	179	39.66%	3.35%	60.34%	3.35%	0.74%	2.17%	0.78%
		Mid	265	47.92%	3.40%	52.08%	1.13%	0.89%	-0.07%	0.39%
		Low	195	41.03%	0.00%	58.97%	7.18%	2.87%	1.60%	2.66%
	Portfolio Governance	High	191	39.79%	3.66%	60.21%	3.66%	1.36%	2.61%	0.97%
		Mid	255	52.55%	3.14%	47.45%	1.96%	0.31%	-0.74%	-0.09%
		Low	193	35.23%	0.00%	64.77%	5.70%	3.06%	2.08%	2.90%
	Undefined		259	35.14%	0.77%	64.86%	2.32%	1.89%	1.77%	1.67%
	Total category		898	41.09%	1.89%	58.91%	3.23%	1.58%	1.14%	1.25%
	All funds		3920	53.34%	5.46%	46.66%	3.27%	-0.16%	0.73%	0.05%

Table 6: Comparative performance according to SR attributes' scores

Morningstar equity category	Morningstar SR attributes scores		Difference in annualized performance	*p*-value[a]
Europe	Portfolio Sustainability	High – Low	0.49%	0.155
	Portfolio Environmental	High – Low	–0.24%	0.307
	Portfolio Social	High – Low	–0.41%	0.187
	Portfolio Governance	High – Low	–0.38%	0.186
US and Canada	Portfolio Sustainability	High – Low	–1.03%	0.006
	Portfolio Environmental	High – Low	0.36%	0.190
	Portfolio Social	High – Low	–0.30%	0.210
	Portfolio Governance	High – Low	–1.80%	0.000
Other	Portfolio Sustainability	High – Low	–2.17%	0.000
	Portfolio Environmental	High – Low	–0.92%	0.089
	Portfolio Social	High – Low	–2.13%	0.000
	Portfolio Governance	High – Low	–1.70%	0.010

[a] Tests for significance were run by boostrapping one-sided *p*-values.

Table 7: Comparative performance according to SR attributes' scores (only surviving funds)

| Morningstar equity category	Morningstar SR attributes scores		Difference in annualized performance	p-value[a]
Europe	Portfolio Sustainability	High – Low	0.48%	0.175
	Portfolio Environmental	High – Low	–0.88%	0.027
	Portfolio Social	High – Low	–0.38%	0.197
	Portfolio Governance	High – Low	–0.02%	0.488
US and Canada	Portfolio Sustainability	High – Low	–0.94%	0.005
	Portfolio Environmental	High – Low	-0.14%	0.368
	Portfolio Social	High – Low	–0.77%	0.016
	Portfolio Governance	High – Low	–0.97%	0.001
Other	Portfolio Sustainability	High – Low	–4.00%	0.000
	Portfolio Environmental	High – Low	–1.28%	0.075
	Portfolio Social	High – Low	–3.12%	0.000
	Portfolio Governance	High – Low	–4.17%	0.000

[a] Tests for significance were run by boostrapping one-sided p-values.

Table 8: Performance persistence

Morningstar equity category	Morningstar SR attributes' scores		Performance Q1 (%)	Performance Q5 (%)	Difference Q5 − Q1 (%)	p-value[a]
Europe	Portfolio Sustainability	High	1.94%	1.32%	−0.62%	0.760
		Mid	−1.38%	2.89%	4.26%	0.009
		Low	−4.24%	1.30%	5.54%	0.013
	Portfolio Environmental	High	−0.27%	0.80%	1.07%	0.538
		Mid	−0.92%	1.88%	2.80%	0.104
		Low	−3.65%	1.98%	5.63%	0.014
	Portfolio Social	High	1.76%	2.00%	0.24%	0.895
		Mid	−0.54%	1.78%	2.33%	0.212
		Low	−3.95%	1.69%	5.64%	0.010
	Portfolio Governance	High	2.32%	1.17%	−1.16%	0.505
		Mid	−1.58%	2.84%	4.41%	0.017
		Low	−4.04%	0.98%	5.01%	0.020
	Undefined		−3.81%	0.36%	4.17%	0.014
US and Canada	Portfolio Sustainability	High	0.45%	0.83%	0.38%	0.861
		Mid	−2.39%	0.90%	3.29%	0.022
		Low	−1.22%	0.52%	1.75%	0.239
	Portfolio Environmental	High	−1.16%	0.51%	1.67%	0.375
		Mid	−2.29%	0.33%	2.61%	0.049
		Low	−1.74%	0.99%	2.73%	0.106
	Portfolio Social	High	−1.17%	0.72%	1.90%	0.299
		Mid	−1.83%	0.82%	2.65%	0.086
		Low	−1.49%	0.63%	2.12%	0.155
	Portfolio Governance	High	0.31%	1.79%	1.48%	0.511
		Mid	−1.80%	0.81%	2.61%	0.045
		Low	−1.72%	0.17%	1.88%	0.177
	Undefined		−4.52%	−2.19%	2.33%	0.297
Other	Portfolio Sustainability	High	−1.09%	4.02%	5.12%	0.028
		Mid	0.11%	−6.15%	−6.26%	0.220
		Low	0.85%	6.41%	5.56%	0.071
	Portfolio Environmental	High	−2.53%	4.71%	7.24%	0.020
		Mid	2.64%	2.88%	0.24%	0.932
		Low	−0.92%	2.87%	3.80%	0.094
	Portfolio Social	High	−0.86%	3.93%	4.79%	0.039
		Mid	−4.14%	−3.35%	0.79%	0.819
		Low	4.55%	−2.29%	−6.83%	0.795
	Portfolio Governance	High	0.40%	4.56%	4.16%	0.087
		Mid	−1.13%	−7.05%	−5.92%	0.782
		Low	0.06%	−4.38%	−4.44%	0.572
	Undefined		−0.72%	−0.62%	0.10%	0.980

[a] Tests for significance were run by boostrapping one-sided p-values.

Table 9: Performance persistence, survivors

Morningstar equity category	Morningstar SR attributes' scores		Performance Q1 (%)	Performance Q5 (%)	Difference Q5 − Q1 (%)	p-value[a]
Europe	Portfolio Sustainability	High	1.73%	0.38%	−1.35%	0.523
		Mid	−0.29%	3.60%	3.89%	0.010
		Low	−4.30%	2.07%	6.37%	0.005
	Portfolio Environmental	High	−1.55%	0.50%	2.05%	0.288
		Mid	0.26%	2.49%	2.23%	0.135
		Low	−3.34%	2.76%	6.10%	0.010
	Portfolio Social	High	1.85%	1.40%	−0.45%	0.799
		Mid	0.39%	1.83%	1.43%	0.446
		Low	−3.80%	2.52%	6.32%	0.005
	Portfolio Governance	High	2.41%	1.61%	−0.80%	0.669
		Mid	−1.33%	3.19%	4.53%	0.009
		Low	−3.87%	2.01%	5.88%	0.007
	Undefined		2.14%	3.70%	1.56%	0.443
US and Canada	Portfolio Sustainability	High	−0.39%	−0.41%	−0.02%	0.992
		Mid	−1.53%	1.63%	3.16%	0.040
		Low	−1.19%	1.05%	2.25%	0.094
	Portfolio Environmental	High	−2.24%	−0.18%	2.07%	0.212
		Mid	−1.85%	0.50%	2.34%	0.093
		Low	−1.30%	2.73%	4.03%	0.021
	Portfolio Social	High	−1.95%	0.45%	2.40%	0.157
		Mid	−1.56%	0.36%	1.92%	0.214
		Low	−0.97%	1.95%	2.92%	0.052
	Portfolio Governance	High	−0.49%	1.18%	1.67%	0.379
		Mid	−1.70%	1.98%	3.68%	0.024
		Low	−0.73%	0.95%	1.68%	0.177
	Undefined		1.29%	−2.88%	−4.17%	0.225
Other	Portfolio Sustainability	High	0.33%	4.46%	4.13%	0.062
		Mid	−0.47%	−6.50%	−6.03%	0.255
		Low	−0.75%	−2.48%	−1.73%	0.420
	Portfolio Environmental	High	−1.33%	4.87%	6.20%	0.057
		Mid	2.77%	2.87%	0.09%	0.974
		Low	−3.63%	3.54%	7.17%	0.147
	Portfolio Social	High	0.60%	4.32%	3.73%	0.088
		Mid	−2.39%	−3.83%	−1.45%	0.731
		Low	3.65%	−2.67%	−6.32%	0.855
	Portfolio Governance	High	2.12%	5.07%	2.94%	0.184
		Mid	−1.91%	−7.22%	−5.31%	0.874
		Low	−1.69%	−3.95%	−2.26%	0.923
	Undefined		0.52%	−5.13%	−5.64%	0.334

[a] Tests for significance were run by boostrapping one-sided p-values.

Figure 1: Performance densities, Portfolio Sustainability Score

(a) Europe   **(b)** US and Canada   **(c)** Other
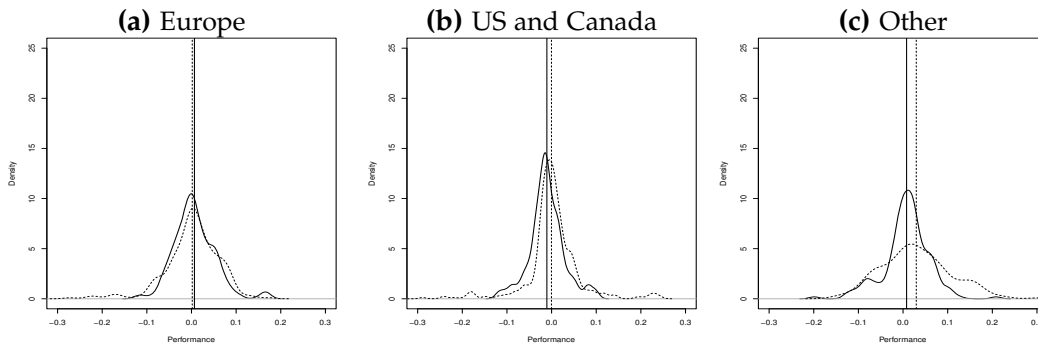
Figure 2: Performance densities, Portfolio Environmental Score

(a) Europe   **(b)** US and Canada   **(c)** Other
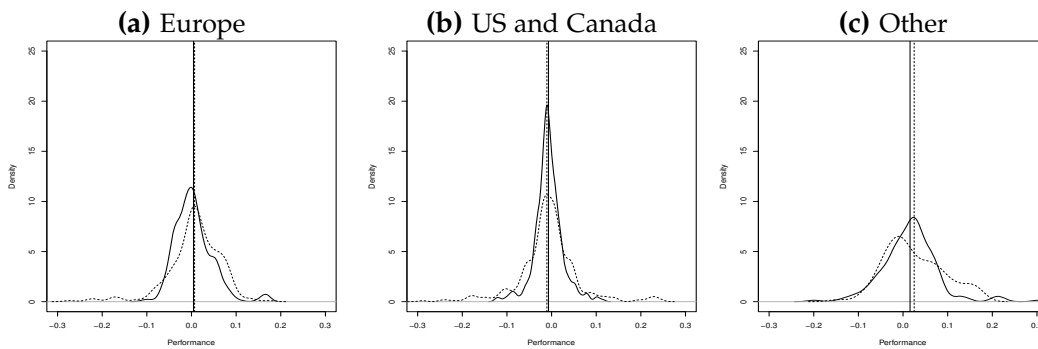
Figure 3: Performance densities, Portfolio Social Score

(a) Europe   **(b)** US and Canada   **(c)** Other
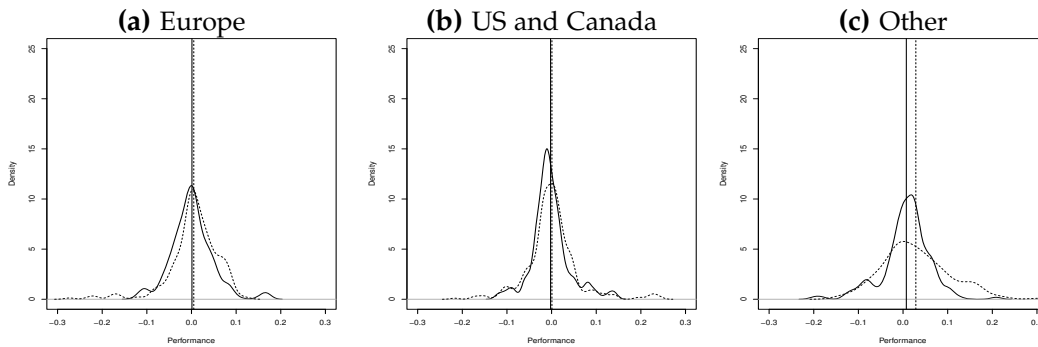
Figure 4: Performance densities, Portfolio Governance Score
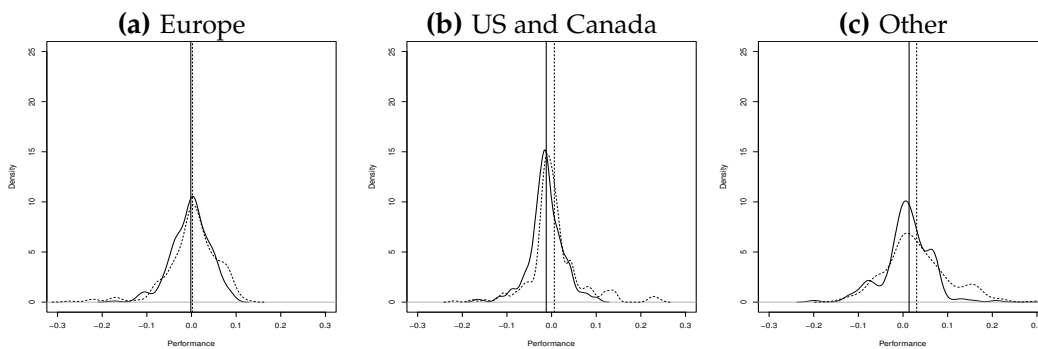
(a) Europe   **(b)** US and Canada   **(c)** Other

High —— Low ------

Notes: All figures contain densities estimated using kernel density estimation for the selected funds. We chose a Gaussian kernel, and the bandwidths were implemented using the plug-in methods of Sheather and Jones (1991). The vertical lines represent the average for each category.

Figure 5: Performance densities, Portfolio Sustainability Score, survivor funds only

(a) Europe   **(b)** US and Canada   **(c)** Other
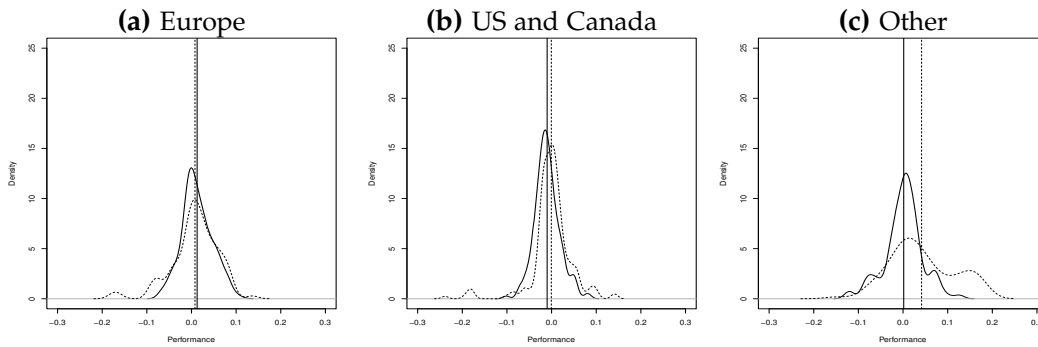
Figure 6: Performance densities, Portfolio Environmental Score, survivor funds only

(a) Europe   **(b)** US and Canada   **(c)** Other
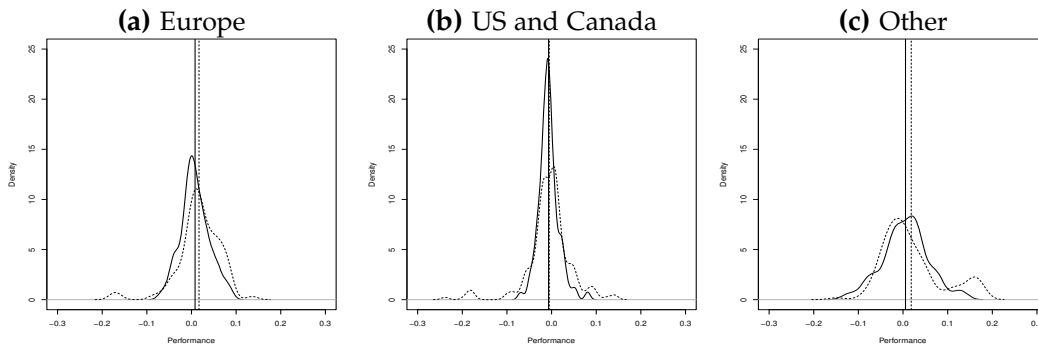
Figure 7: Performance densities, Portfolio Social Score, survivor funds only

(a) Europe   **(b)** US and Canada   **(c)** Other
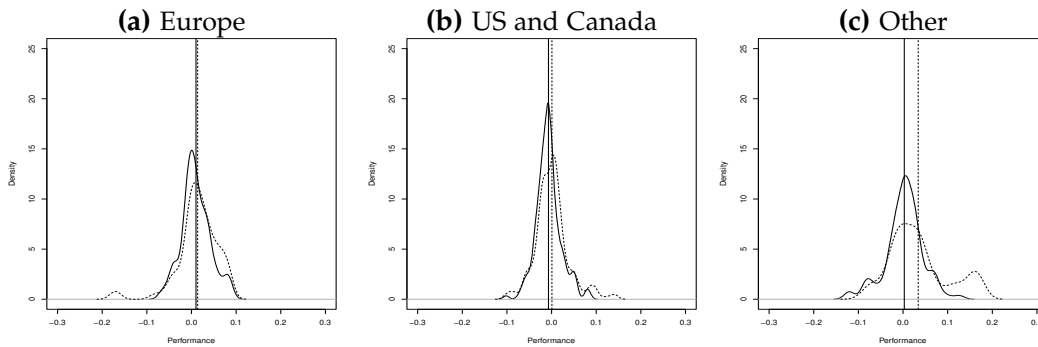
Figure 8: Performance densities, Portfolio Governance Score, survivor funds only
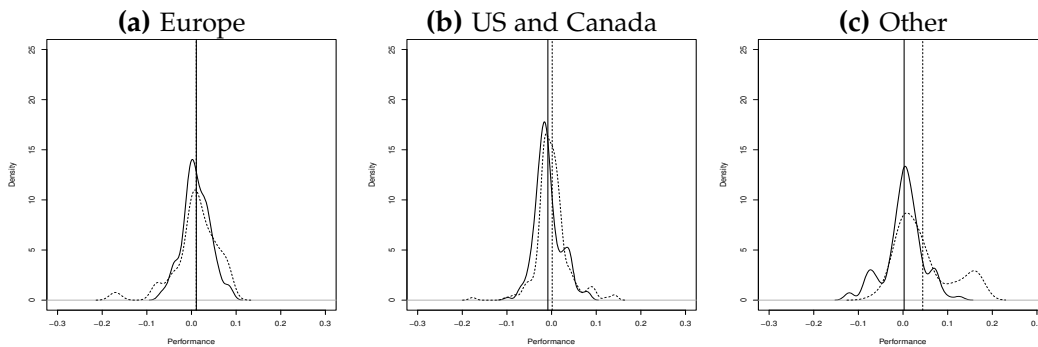
(a) Europe   **(b)** US and Canada   **(c)** Other

High —— Low ------

Notes: All figures contain densities estimated using kernel density estimation for the selected funds. We chose a Gaussian kernel, and the bandwidths were implemented using the plug-in methods of Sheather and Jones (1991). The vertical lines represent the average for each category.
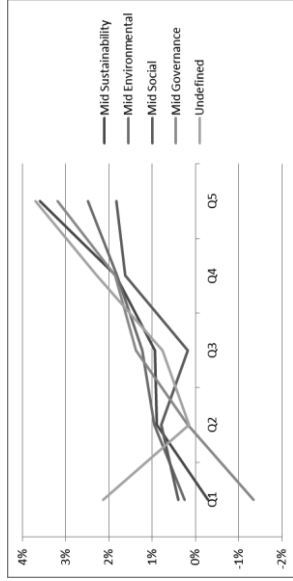
Figure 9: Performance persistence, all funds

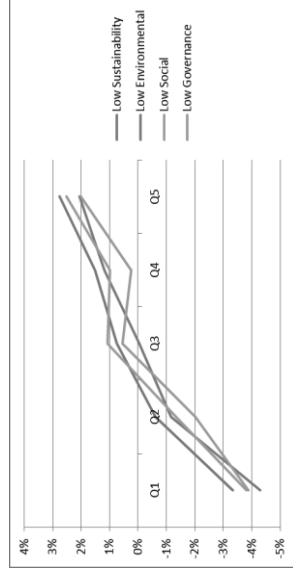**Figure 10**: Performance persistence, surviving funds
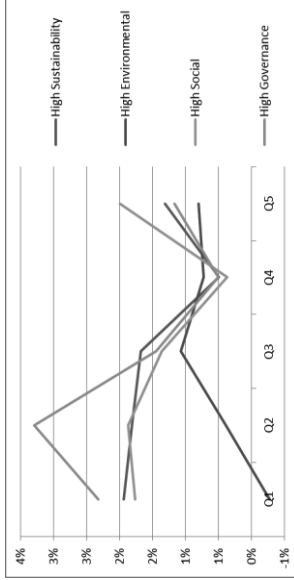
Funds with high SR scores
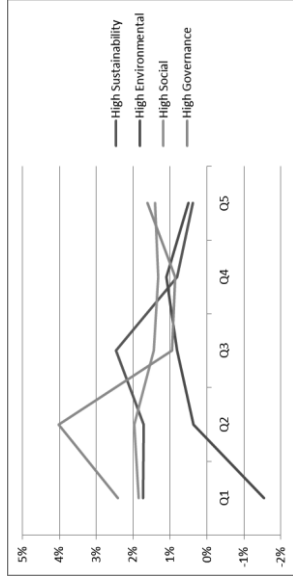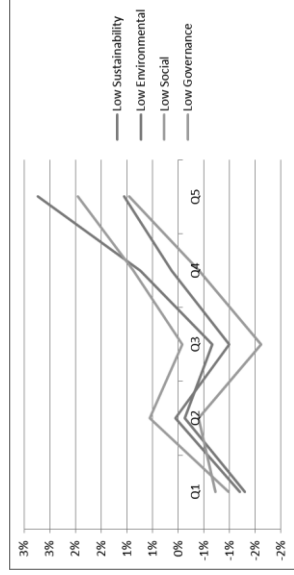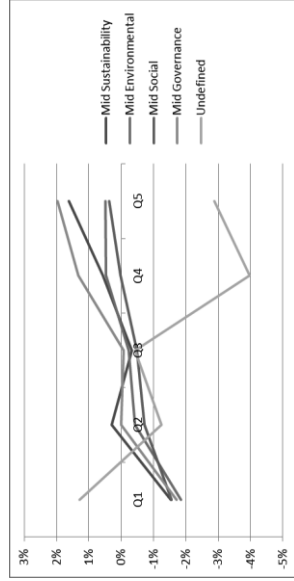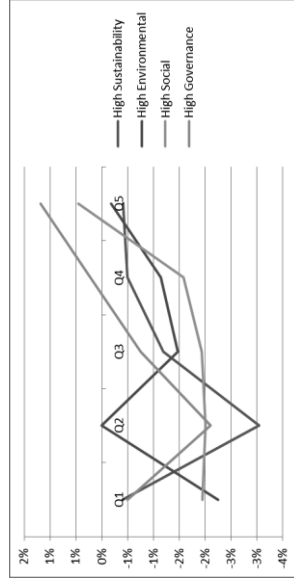
Funds with mid SR scores (and undefined)

Funds with low SR scores