

Does Sustainability Score Impact Mutual Fund Performance?

Pablo Durán-Santomil, University of Santiago de Compostela

(pablo.duran@usc.es)

Luís. A Otero-González, University of Santiago de Compostela

(luis.otero@usc.es)

Renato Heitor Correa-Domingues, Universidade Lusíada do Porto

(renatodomingues@por.ulusiada.pt)

Juan Carlos Reboredo-Nogueira, University of Santiago de Compostela

(juancarlos.reboredo@usc.es)

Abstract: Given that sustainable investing constitutes a major force across global financial markets, in 2016 Morningstar began reporting Morningstar Sustainability and ESG scores. We use these scores to study the effects of Socially Responsible Investments (SRI) on European equity fund performance. Sustainability score and the different pillars of ESG scores (environmental, social, and governance) impact negatively on performance. We also test the effect on mutual fund flows and risk. The sustainability score is significant on the flows, so higher-rated funds receive a larger volume of funds. In terms of risk, the level of sustainability is negatively related to the VaR (value at risk) of the fund, supporting that higher scored mutual funds offer better protection against extreme losses.

Keywords: mutual fund performance, mutual fund flows, Sustainability Score, ESG score, Socially Responsible Investment, SRI.

JEL Classification: G11; G23.

1. Introduction

Socially Responsible Investment (SRI), also known as sustainable, responsible, and impact investing, is “an investment discipline that considers environmental, social and corporate governance (ESG) criteria to generate long-term competitive financial returns and positive societal impact” (US SIF, n.d.¹). According to the 2016 Global Sustainable Investment Review (GSIA, 2017), in 2016 there were \$22.89 trillion assets being professionally managed under SRI strategies in the world. Bilbao-Tero, Álvarez-Otero, Bilbao-Tero and Cañal-Fernández (2017) conclude that the SRI label in the mutual fund industry is valued favorably by the market, which is an important factor that drives this growth. Barreda-Tarrazona, Matalín-Sáez and Balaguer-Franch (2011) conclude that social preference (instead of financial performance) is the primary factor for investors choosing SRI mutual funds.

The growing interest in SRI in recent years has led to several organizations assessing mutual funds on how well the underlying companies perform on ESG issues. In 2016, Morningstar launched a Morningstar Sustainability Rating. The idea of the Morningstar Sustainability Rating is classifying mutual funds about ESG factors relative to their Morningstar category peers. The advantage of this product is that it makes it possible to find sustainable funds even if they aren't labelling themselves specifically as funds that support an SRI approach. The use of these scores shows an important difference to previous studies, which compare SRI funds with an index, or the most advanced studies apply a so-called matching approach, i.e. they compare the performance of SRI and non-SRI investment funds with similar characteristics (fund size, fund age, expenses, et cetera.) to properly considered management and transaction costs for both SRI funds and conventional funds (see, among others, Mallin, Saadouni and Briston, 1995; Gregory,

Matatko and Luther, 1997; Statman, 2000; Kreander, Gray, Power and Sinclair, 2002; Kreander, Gray, Power and Sinclair, 2005).

One important research question in the mutual fund industry about SRI investing is to know how SRI mutual funds perform. There are several studies that have demonstrated that companies with social responsibility policies and practices are good investments. For example, a recent paper of Friede, Busch and Bassen (2015) conducted a meta-analysis of about 2,200 unique primary empirical studies. They found that the majority of studies show a positive correlation between ESG factors and financial performance. But despite the investigations carried out to date, there is still a debate about whether these types of investments can create value for investors or not and why they put their money here. Although according to Lewis and Mackenzie (2000) and Webley, Lewis and Mackenzie (2001), some investors in SRI funds are willing to accept lower returns for their moral stance, the performance of SRI funds and conventional funds is still an open question. As Junkus and Berry (2015) sustain, after a review of the most recent work in major finance journals on SRI, “the performance of SR mutual funds and indexes are not generally significantly different to conventional funds or indexes, but again these results are also highly dependent on model specification, time period, benchmark, and other characteristics of the study”.

Authors such as Luther, Matatko and Corner (1992) and Mallin, Saadouni and Briston (1995) support the idea that SRI funds outperform market indexes. But the more conventional theory is that SRI mutual funds have the same return as any other funds, and authors such as Hamilton, Jo and Statman (1993), Kreander, Gray, Power and Sinclair (2002, 2005), Gregory and Whitakker (2007) and Bauer, Derwall and Otten (2007), Humphrey, Warren and Boon (2016) and Syed (2017) are in line with this theory. Another theory defends that choosing SRI funds are basically a “trade off” between investing in

SRI and returns, so SRI investments underperform the benchmark (for example, White, 1995). One important recent paper is Nofsinger and Varma (2014), which provides a new perspective because they found that the difference between socially responsible (SR) and conventional mutual funds depends on the state of the market. SR mutual funds outperform conventional mutual funds during periods of market crisis, but in non-crisis periods, SR funds underperform conventional funds.

Previous research has studied the effect of sustainability on performance exclusively using a dichotomous variable to differentiate between socially responsible funds and conventional funds. However, the results could be biased because under "socially responsible", they could have funds with very different levels of sustainability. Statman and Glushkov (2016) conclude that there is a lack of clearly defined criteria to distinguish mutual funds as "socially responsible" results in inconsistently applied classifications that make it difficult to measure the performance of SRIs. Traditional methodology in empirical research is benchmarking with indices or, most recently, matched pair analysis, which was initially applied by Maillin, Saadouni and Briston (1995) and is based on comparing returns of SRI funds and conventional funds with similar characteristics in terms of volume of assets, inception dates, et cetera. For this reason, the inclusion of sustainability scores in our work allows us to evaluate whether the degree of sustainability of the portfolio in which the funds are invested has a positive effect on performance. As far as we know, only Dolvin, Fulkerson and Krukover (2017) and El Ghoul and Karoui (2017) analyzed this effect. Dolvin, Fulkerson and Krukover (2017) conclude that funds with higher Morningstar Sustainability scores have similar alphas from those with lower Sustainability scores. Authors also observe that there is little difference in the performance or Sustainability scores between self-proclaimed SRI funds versus those that fall in the top 50 and top 20 percent of Morningstar's Sustainability

scores. Finally, they observe that mutual funds with higher Morningstar Sustainability metrics do not appear to be more attractive to investors compared to low scoring funds. In contrast, self-proclaimed SRI funds have performed significantly better regarding fund flows. El Ghoul and Karoui (2017) use CSR (Corporate Social Responsibility) scores to study the effect on fund performance and flows, concluding that higher values display poorer performance and weaker performance-flow relation. From an investor point of view, the advantage of using Sustainability scores is that they can select their SRI, taking into consideration the funds with better scores, whether or not they are declared as an SRI fundⁱⁱ.

This paper adds to the growing literature on SRI by specifically examining the effect of the degree of sustainability, measured through Morningstar Sustainability scores included in Morningstar Direct in 2016. In particular, we assess the effect of sustainability scores and the different dimensions in which the score is subdivided (environmental, social, and governance) in the performance, in addition to the downside risk and the flow of funds. On the other hand, the conventional dichotomous variable has been added to the models to evaluate to what extent the results may differ. Our empirical evidence also contributes to the literature on mutual funds that discusses whether applying a particular investment screening in portfolio selection affects the mutual fund performance (see, for example, Bauer, Derwall and Otten, 2007 or Muñoz, Vargas and Marco, 2014). SRI portfolios are subject to both positive and negative social screens (Rivoli, 2003). cThe Portfolio theory argues that narrowing the universe of assets restricts diversification opportunities and thus the risk-adjusted performance (Rudd, 1981); whereas Hill, Ainscough, Shank and Manullang (2007) and Chegut, Schenk, and Scholtens, 2011) consider that restricting investment screening allows the identification of companies with higher growth potential and better management, therefore leading to a better financial

performance and risk profile. Sustainable mutual funds apply a specific portfolio screening by concentrating investments in socially conscious businesses. Although there is profuse empirical literature on the impact of social responsibility of the performance, little is known about the sustainability-based screening.

Our empirical results show that a large number of funds are not declared sustainable but their portfolio is comparable to sustainable mutual funds. Furthermore, the Sustainability score is significant in explaining the level of performance, downside risk, and flows. We also achieved equivalent results for the three dimensions of sustainability (environmental, social and corporate). The signs are different on performance and downside risk when the conventional dummy to declare social mutual funds is used.

The remainder of this paper is laid out as follows. In Section 2, we review the related literature on SRI performance, in Section 3 we describe our data and the performance evaluation metrics, in Section 4 we describe our empirical methods and results, in Section 5 we conduct robustness tests and, finally, we draw conclusions from our research.

2. Literature Review

Over the last few years, SRI investment research has been growing. The CFA Institute, which is a global association for investment professionals, states that “a key idea in the discussion of ESG issues is that systematically considering ESG issues will likely lead to more complete analyses and better-informed investment decisions” and “that every investment analyst should be able to identify and properly evaluate investment risks, and ESG issues are a part of this evaluation” (CFA Institute, 2015). For this association, there are basically two investors interested in considering ESG issues: *value-*

motivated and *values-motivated investors*. We focus on the first kind of investors concerned with the financial performance of their SRI funds.

Hamilton, Jo and Statman (1993) developed three hypotheses regarding the performance of SRI mutual funds. The first hypothesis is that SRI fund performance equals that of conventional funds, which is consistent with a market that does not regard the social responsibility feature. The second hypothesis is that SRI fund performance is lower than that of conventional funds, which is consistent with a market that values the social responsibility feature. Finally, the third hypothesis is that SRI fund performance is higher than that of conventional funds. There are several arguments which could explain why SRI mutual funds can outperform, in financial terms, the conventional funds (which do not consider ESG factors). First, SRI mutual funds have a higher proportion of their portfolio in the segment of small companies; these companies are better adapted to market changes (Luther, Matatko and Corner, 1992; Gregory, Matatko and Luther, 1997) and may also be more profitable in the long run. Second, social companies are more efficient, better managed and develop better in the market (Hamilton, Jo and Statman, 1993). From a theoretical point of view (for example, Margolis, Elfenbein and Walsh, 2009 or Flammer, 2015), social companies can reduce costs (penalties, etcetera.) or increase revenues (innovative products, greater employee effort, better public perception, increasing the likelihood that consumers will purchase the company's products or its share price, attract socially conscious customers, etcetera.). In contrast, one important argument of the detractors of SRI funds is that the universe of possible investments of these funds (individual companies) is small, so they assume a higher investment risk because of the lack of diversity (Chegut, Schenk and Scholtens, 2011). Humphrey and Tan (2014) replicate 10,000 pairs of SRI and conventional portfolios to test the impact of SRI screening on performance, finding no significant difference in the risk-adjusted

return of screened and unscreened portfolios. They conclude that a typical SRI fund will neither gain nor lose from screening its portfolio. But Trinks and Scholtens (2017) find that negative screening implies an opportunity cost, because excluding controversial stocks for an investment portfolio may reduce financial performance. Authors such as Kurtz (1997) or Goldreyer and Diltz (1999) argue that SRI mutual funds managers need more information than conventional funds about the companies in which they invest; they base their decisions on deeper, more complete, and higher quality information, resulting in a significant reduction in the risk of their investment decisions. Empirical evidence of some authors, such as Luther, Matatko and Corner (1992) and Maillin, Saadouni and Briston (1995), support the idea that SRI funds outperform conventional investments. But there is also evidence to support the idea that SRIs are neutral to financial performance (Hamilton, Jo and Statman, 1993; Kreander, Gray, Power and Sinclair, 2005; Gregory and Whittaker, 2007; Bauer, Derwall and Otten, 2007; among others), or that SRI funds underperform conventional investments (for example, White, 1995).

The first study about SRI investment was done by Luther, Matatko and Corner (1992), where these authors found that SRI investment funds did not under or outperform the index benchmark. They used 15 British Ethical funds, finding weak evidence that 15 UK SRI funds outperformed two stock market indices. Hamilton, Jo and Statman (1993) conducted a similar study where the difference of means of excess returns was not significant and only one of 17 mutual funds had a positive Jensen's alpha. Luther and Matatko (1994) improved their prior work by including a small market index and they concluded that the excess returns of SRI funds are strongly influenced by the low capitalization of the small cap stocks. The study also shows that SRI funds have a neutral effect on performance. White (1995) researches US and German mutual funds using a simple regression against an environmental market index, showing that the SRI investments underperform the benchmark in terms of three performance measures

(Jensen's alpha, the Treynor ratio, and the Sharpe ratio). In this research, the author used a sample of six US funds and five German SRI Investment funds.

All previous studies used an index as benchmark, so they have the problem of what is the appropriate index. Mallin, Saadouni and Briston (1995) avoided this problem by using a matched pair analysis to compare SRI mutual funds and conventional funds in the UK. The authors matched 29 SRI mutual funds to conventional ones using the size and the age of the funds as criteria. Their results showed no differences in the performance of both samples using the Sharpe and Treynor ratios as performance measures, but they found that ethical funds did better than the non-ethical funds when the Jensen performance measure was used. Gregory, Matatko and Luther (1997) studied 18 SRI funds where the investment area and the fund type were considered. They did not find differences in performance against conventional funds. Statman (2000) studied the performance of 31 US SRI mutual funds and the Domini 400 Social-Index (DSI) from 1990 to 1998. The results show that only some SRI funds could underperform the benchmark (S&P 500 or DSI). But, in general, SRI funds obtained a similar performance to S&P 500, DSI, and conventional funds. Kreander, Gray, Power and Sinclair (2002) used a matching procedure and the age, size, country and investment universe of the fund as variables. The study included mutual funds from Sweden, the Netherlands, Norway, Germany, the UK and Switzerland, and Jensen's alpha and the Sharpe and Treynor ratios as performance metrics. Their results showed that SRI funds' performance was very similar to those of conventional funds. Kreander, Gray, Power and Sinclair (2005) studied the performance of 30 European SRI funds from four countries, finding that there is no difference between SRI funds and conventional funds.

Bello (2005) studied 42 SRI U. mutual funds; he found no evidence of a performance difference between SRI and conventional funds. Both underperformed the

Domini 400 Social Index and S&P 500 during the study period (1994 – 2001). Bauer, Koedijk and Otten (2005) investigated the performance of 32 British, 16 German and 55 US SRI funds, they used Jensen and Carhart's alpha and found that German and US SRI mutual funds underperformed in both their relevant indexes and the conventional funds, whereas UK funds slightly outperformed, however the differences were not significant. Scholtens (2005) investigated the performance of Dutch SRI funds and found that these funds outperformed conventional funds but with no statistically significant difference. Also Barnett and Salomon (2006) studied 61 SRI funds tracked by the US Social Investment Forum (USSIF). They found that the relationship between financial and social performance is neither strictly negative, nor strictly positive. Instead, they found a curvilinear relationship, suggesting that the two viewpoints may be complementary. Risk-adjusted performance varies with the types of social screens used. Community relations screening (excludes firms that do not invest in and/or develop economically depressed communities) increased financial performance, but environmental and labor relations screening (excludes firms with a record of poor environmental performance and firms with a record of poor labor relations practices, respectively) decreased financial performance.

Bauer, Otten and Rad (2006) investigated the performance of Australian ethical funds, and Bauer, Derwall and Otten (2007) investigated evidence from Canada, finding no statistical difference in performance between conventional and SRI funds. Gregory and Whittaker (2007), in the UK market, found that neither SRI nor non-SRI funds exhibited significant under performance. Renneboog, Ter Horst and Zhang (2008) found that SRI funds in the US, the UK, and in many continental European and Asia-Pacific countries underperformed their domestic benchmarks. However, with the exception of France, Japan, and Sweden, the risk-adjusted performance of SRI funds is not statistically

different from the performance of conventional funds. Gil-Bazo, Ruiz-Verdú and Santos (2010) found that during the period 1997–2005, US SRI funds had better performance (gross and net Carhart's alphas) than conventional mutual funds with similar characteristics. Authors find that the differences are driven exclusively by SRI funds run by management companies specializing in SRI, while funds run by companies not specializing in SRI underperform conventional funds

Climent and Soriano (2011) compare the investment performance of US environmental or green mutual funds, finding that in the 1987–2009 period, green funds had lower performance levels than conventional funds but, in the period 2001–2009 green funds achieved risk-adjusted returns not significantly different from the rest of SRI and conventional mutual funds. Cortez, Silva and Areal (2011) performed a study focused on 88 SRI funds from the European market from 1996 to 2007. They concluded that the performance of SRI funds is similar to the performance of both conventional and socially responsible indexes. Cortez, Silva and Areal (2012) studied seven European and US markets from 1996 to 2008. They found in several European markets (Belgium, France, Germany, Italy, the Netherlands, and the UK) that SRI funds showed similar performances compared to both conventional and benchmarks. In contrast, the US and Austrian funds showed evidence of underperformance.

Pérez-Gladish, Méndez Rodríguez, M'Zali and Lang (2013) compared 46 randomly selected US-based large-cap equity mutual funds (25 are members of the SIF and 21 are conventional funds) finding there were no significant performance differences between conventional and SRI mutual funds employing Data Envelopment Analysis. Nofsinger and Varma (2014) found that SRI mutual funds outperformed conventional funds in the global financial crisis, so they can be an optimal choice for investors who want to protect themselves from downside risk. They also found that SRI funds

underperform at other times. Leite and Cortez (2014) performed a multi-country study focused on 54 international SRI funds located in eight European markets (Austria, Belgium, France, Germany, Italy, the Netherlands, the UK, and Spain); they applied the five-factor model and found a similar performance between socially responsible funds and conventional funds. Muñoz, Vargas and Marco (2014) studied 89 European green funds and 18 US funds from 1994 to 2013. They applied the Carhart four-factor model and stated that, for the US market, green funds did not perform any worse than the market, but with a global equity portfolio green funds showed evidence of underperformance. Becchetti, Ciciretti, Dalo and Herzel (2015) found no clear-cut dominance over the entire period analyzed (1992-2012), but also found that SRI funds generally did better than conventional funds in the period following the global financial crisis of 2007. Leite and Cortez (2015), focusing on the French market, found that SRI funds underperformed slightly more than their matched samples according to different models, but differences in alphas are not statistically significant in most cases. They only found significance in one of the estimated models at the 10% significance level. Humphrey, Warren and Boon (2016) found that SRI managers have longer tenure and are more likely to be female, but they did not find any significant difference in the performance of SRI and conventional funds. Ibikunle and Steffen (2017) conducted a comparative financial performance analysis on European green, conventional, and black mutual funds; they concluded that there was no difference in the performance of the green and the conventional funds and that green funds are beginning to significantly outperform black funds.

Dolvin, Fulkerson and Krukover (2017) is the only reference, to our knowledge, that employs Morningstar Sustainability scores in their analysis. The authors conclude that funds with higher Morningstar Sustainability scores have similar alphas from those with lower Sustainability scores. The authors also observe that there is little difference in

the performance or Sustainability scores between self-proclaimed SRI funds versus those that fall in the top 50 and top 20 percent of Morningstar's Sustainability scores. Finally, El Ghouli and Karoui (2017) employed a CSR score, which is an asset-weighted composite CSR fund score. They showed the effects of CSR on fund performance; compared to low-CSR funds, high-CSR funds displayed a poorer performance.

3. Empirical study

3.1. Sample

Our sample contains 1,593 European equity funds rated by Morningstar Sustainability in November 2016. The funds are the "open funds" type with an ESG score in the investment area of Europe. Furthermore, to avoid problems of multicollinearity, we have selected only an equivalent class for each fund. We obtained for each equity mutual fund several measures of performance and other variables such as size, volatility, socially conscious, expenses, and age. We also used the Morningstar style-box to control the effect of the different categories which are included in the sample. The number of funds varied when we considered the costs where the sample reduces from 1,593 to 571 motivated for the lack of data available in Morningstar Direct.

3.2. Variables construction

Our sustainable variables have been obtained from Morningstar Direct (original source Sustainalyticsⁱⁱⁱ). We will employ five variables: three are the pillars scores [Environment score variable (Envscore), Social score variable (Socscore) and Government score variable (Govscore)], the fourth is the ESG score of a portfolio (ESGscore), and finally, the Portfolio Sustainability Score (Sustscore), which is the ESG

score minus the Portfolio Controversy Score. Bos (2017) examines the process by which Morningstar Sustainability and ESG scores are calculated.

The Morningstar Portfolio Sustainability Score is a measure developed in 2016 for scoring mutual funds and ETFs about ESG risks and opportunities. The subsequent Morningstar Sustainability Rating is a comparison relative to their Morningstar Category peers and is derived from the Morningstar Portfolio Sustainability Score. The Morningstar Portfolio Sustainability Score (Sustscore) is defined as follows (Morningstar, 2016a and 2016b):

$$\text{SustScore} = \text{Portfolio ESG Score} - \text{Portfolio Controversy Deduction}$$

In order to receive a portfolio sustainability score, a portfolio must have a portfolio ESG score and a portfolio controversy score, which, according to Morningstar (2016b), at least 50% of a portfolio's assets under management must have these scores. The Morningstar Portfolio ESG Score (ESGscore)^{iv} is calculated as:

$$\text{ESGscore} = \sum_{i=1}^n w_i \text{ESGNorm}_i$$

Where:

ESGNorm_i = the normalised ESG score of company i ,

n = the number of securities in the portfolio,

w_i = the asset weight on security i , so the sum $\sum_{i=1}^n w_i = 100\%$.

Normalised company ESGs are aggregated to a portfolio ESG score using an asset-weighted average of all covered securities. Sustainalytics tracks and categorizes ESG-related incidents in companies, which are called “controversies”. A single company may be involved in multiple ESG-related incidents at any given time, so Sustainalytics makes a controversy score of company i . Morningstar employs company controversy scores of Sustainalytics, creating a Morningstar Portfolio Controversy Score (MContr_p), as follows:

$$MContr_p = \sum_{i=1}^n w_i SCont_i$$

Where:

w_i = the asset weight on security i

$SCont_i$ = the Sustainalytics controversy score of company i .

We have divided our sample funds into two groups based on whether the ESG scores are below or above the median. Then, we estimated the means and their differences between both groups. Table 1 reports the results of the univariate analysis. As can be observed, the differences are very significant between the two groups for the different scores, with a difference of approximately five points in favour of the funds included in the high score group.

Table 1- Sustainability and ESG scores for different groups (A)

Variable	Low ESG score	High ESG score	t-statistic
Sustscore	52.63	58.58	-39.29***
ESGscore	57.44	63.99	-33.78***
Envscore	56.66	62.36	-30.70***
Socscore	57.01	62.95	-34.81***
Govscore	55.28	60.41	-31.97***

This table reports the values of sustainability variables considered in the analysis obtained from Morningstar direct database. The funds are classified into low or high groups depending on whether their score is above or below the median. The t-statistic for difference of means is reported in the third column. Sustscore is the level of sustainability of the mutual fund measured by Morningstar. ESGscore is the ESG score of a fund. EnvScore, Socscore and GovScore are the mutual fund scores for the three dimensions (environment, social and corporate governance). *Significant at 10%; ** significant at 5% and *** significant at 1%.

Table 2 compares the funds declared sustainable (Sociallyconscious) and those that result from dividing the sample according to a low ESGscore or a high ESGscore criteria. As you can see, there are big differences but, in general, mutual funds that are declared sustainable are from an ESGscore point of view too. However, there are many funds that are not declared sustainable but they are based on the level of sustainability of the companies of their portfolio. Thus, by using scores, investors have at their disposal a large number of funds that are not declared sustainable but with comparable portfolios of

self-declared sustainable funds. This result is similar to Dolvin, Fulkerson and Krukover (2017) for the US market.

Table 2- Sustainability and ESG scores for different groups (B)

Sociallyconscious	Sustaineddummy		Total
	0	1	
0	671	790	1,461
1	19	113	132
Total	690	903	1,593

This table reports the number of mutual funds classified as sustainable using two different dummy variables. Sustaineddummy is based on low or high sustainable scores depending on whether their score is above or below the median. Sociallyconscious is for those mutual funds declared as socially conscous.

3.3. Performance variables

We considered different performance measures from the Morningstar Direct database. Given that we only have ESG data available for December 2016, we have analyzed the performance and risk effects using the performance and risk metrics for the last two years based on Wimmer (2012), who showed that ESG scores persisted for two years, and were motivated by the changes in the holdings of the SRI mutual funds. In particular, we used the raw return and Sharpe ratios. We also computed Carhart's alphas based on values provided on Kenneth French's website^v.

The differences in performance between the high and low ESG scored funds are negative when considering raw returns, Sharpe ratios and two years alphas (Table 3). That is, higher ESG scored funds show a poorer performance, except in the case of a one year alpha. Our results are consistent with those achieved by El Ghouli and Karoui (2017) and Dolvin, Fulkerson and Krukover (2017) for US mutual funds.

Table 3- Performance metrics for different groups

Metric	Low ESG score	High ESG score	t-statistic
Alpha-2y	0.030	-0.270	1.61*
Alpha-1y	-0.070	-0.200	0.49
Sharpe-2y	0.047	-0.004	4.34***
Sharpe-1y	0.009	-0.100	5.04***
Return-2y	7.450	6.370	5.27***

Return-1y	3.680	1.060	6.14***
------------------	-------	-------	---------

This table reports the values of the performance metrics indicated in the first column, Alpha (Carhart's alpha) and Sharpe (Sharpe ratio) are risk adjusted returns calculated for two and one years estimated at the end of 2016. Return is the raw measure of profitability. The data has been obtained from Morningstar Direct database and Kenneth French's website. The funds are classified into low or high groups depending on whether their ESG score is above or below the median. The t-statistic for difference of means is reported in the third column. *Significant at 10%; ** significant at 5% and *** significant at 1%.

3.4. Downside risk variables

We also assessed fund performance by considering downside risk. Tail risk is commonly taken by mutual funds and it has been shown to be useful in explaining fund performance (Kelly and Jiang, 2014). Specifically, we examined whether sustainable mutual funds are more or less exposed to tail risk by measuring mutual fund downside risk by using the Value at Risk (VaR). VaR measures the maximum loss that a fund i can obtain for a given time period and a given confidence level (1-p) as:

$$Pr(R_i \leq VarR_i)$$

which is the loss associated with the p-th percentile of the return distribution. It can be computed as $Var_i = F_i^{-1}(p)$, where F_i is the return distribution of the fund i .

Table 4 shows the difference of means for downside risk measured by the historical monthly VaR at a 99% confidence level. The evidence for VaR reveals that highly scored mutual funds display less tail risk but are only statistically significant for the two years measured.

Table 4- VaR for different groups

	Low ESG score	High ESG score	t-statistic
VaR2y	7.97	7.61	4.30***
VaR1y	6.96	6.83	1.31

This table reports the VaR values obtained from Morningstar Direct database. The funds are classified into low or high groups depending on whether their ESG score is above or below the median. The t-statistic for difference of means is reported in the third column. *Significant at 10%; ** significant at 5% and *** significant at 1%.

3.5. Flow of funds

We measure the flow of funds as:

$$Flow_i = \frac{TNA_{i,t} - TNA_{i,t-1} (1 + R_{i,t})}{TNA_{i,t-1}}$$

Where $TNA_{i,t}$ and $TNA_{i,t-1}$ are the total net assets for fund i at the end of year t and $t - 1$, respectively, and $R_{i,t}$ is the return of fund i in year t .

Table 5 displays the difference of means for the flow of funds showing positive differences for higher scored mutual funds.

Table 5- Flows of funds for different groups

	Low ESG score	High ESG score	t-statistic
Flowasset_2015	-0.01	0.02	-2.15**
Flowasset_2016	-0.10	-0.07	-1.75*

This table reports the values of flow of funds obtained from Morningstar Direct database. The funds are classified into low or high groups depending on whether their score is above or below the median. The t-statistic for difference of means is reported in the third column. *Significant at 10%; ** significant at 5% and *** significant at 1%.

3.6. Descriptive statistics

Table 6 shows the different variables considered in our work. As can be seen, the variables related to the level of sustainability have an average level close to 60 points, and the difference between the minimum and maximum is around 25 points. On average, the funds have a negative alpha despite yielding positive returns for the term of 1 and 2 years. The average flow has been negative and the percentage declared to be socially responsible is very small (8%). The size is very variable, the expense ratio is greater than 1% because the mutual funds included invest in equity, and in general the funds have a high average age.

Table 6- Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Sustscore	1,593	56.01	4.20	40.20	66.08
ESGscore	1,593	61.16	5.02	44.89	70.76
Envscore	1,593	59.89	4.63	45.78	68.48
Socscore	1,593	60.38	4.48	45.61	69.66
Govscore	1,593	58.19	4.07	44.59	65.98
Alpha-2y	1,766	-0.12	3.51	-23.57	51.67
Alpha-1y	1,766	-0.07	5.79	-46.98	103.11
Sharpe-2y	1,829	0.03	0.26	-1.49	1.33
Sharpe-1y	1,829	-0.04	0.44	-2.68	2.34

Return-2y	1,829	7.01	4.51	-14.10	52.66
Return-1y	1,829	2.42	8.47	-41.86	110.74
VaR-2y	1,829	7.69	1.64	2.15	18.77
VaR-1y	1,829	6.86	2.01	0.85	19.32
Size_2015	1,953	307.00	811	0.00	18,900.00
Size_2016	1,834	289.00	683	0.89	14,800.00
Flowasset_2015	1,856	-0.00	0.24	-0.98	0.62
Flujoasset_2016	1,757	-0.08	0.23	-0.99	0.62
ExpRatio	765	1.34	0.73	0.00	5.21
Age	2,08	13.02	7.10	1.50	59.93
Sociallyconciuous	2,08	0.08	0.27	0	1

This table reports the values of the variables considered in the analysis obtained from Morningstar Direct database. Sustscore is the level of sustainability of the mutual fund. ESGscore is the ESG score of a fund. Envscore, Socscore and Govscore are the mutual fund scores for the three dimensions (Environment, Social and Corporate Governance). Alpha and Sharpe are risk adjusted metrics. Return is the raw profitability and VaR a downside risk measure. Size is the amount invested in millions of euros. Flowasset is the percentage of new funds over total assets and ExpRatio the net expense ratio. The age are the years from inception date and Sociallyconciuous is a dummy that takes the value one when the fund declares itself socially responsible and zero otherwise.

3.7. Fund performance and Sustainability Scores

In this part, we test if the degree of sustainability measured through ESG scores has a positive or negative effect on performance. In addition, we consider ESG scores to evaluate the contribution of each dimension to the portfolio performance. We propose the following model:

$$Y_i = cons_i + \beta_1 Sustscore_i + \beta_2 Age_i + \beta_3 LossDev_i + \beta_4 logSIZE_i + \beta_5 ExpRat_i + \beta_6 Sociallyconciuous_i + \sum_j Category_j + \varepsilon_i$$

where:

Y_i = Alternative performance metrics for fund i .

i = 1 through N , where N is the total number of funds in the sample.

$Sustscore_i$ is the sustainability score provided by Morningstar.

Age = Years since inception date.

$LossDev$ = standard deviation of mutual funds returns.

$LogSize$ = logarithm of mutual fund market value.

$ExpRat$ = Net expense ratio of fund i .

Sociallyconciuous = dummy of SRI mutual funds.

Category= dummies of categories except small style.

$cons_i$ and $\beta_1, \beta_2, \beta_3,$ and β_4 are parameters of the regression and ε_i the term error.

Our results show that Sustscore is significant in explaining the level of performance for all the metrics and terms. If we use ESG scores instead of Sustscore, the results are mainly the same. Most of the models present a negative sign in line with El Ghoul and Karoui (2017) and Renneboog *et al.* (2008), who suggest that socially responsible mutual funds underperform other funds. The dummy variable is also significant, showing that considering the level of sustainability can help to better understand the relationship between performance and social responsibility. Our results support Statman and Glushkov (2016), who conclude that the lack of clearly defined criteria to distinguish mutual funds as socially responsible affects the results of previous research based on dichotomy variables. Among the control variables, Table 7 shows that volatility and the expense ratio, but only in some models, are negatively related to performance, while size and age are not significant.

Table 7- Sustainability Scores and fund performance

Variable	Return-2y	Return-1y	Sharpe-2y	Sharpe-1y	Alpha-2y	Alpha-1y
Sustscore	-0.2806***	-0.7726***	-0.0133***	-0.0351***	-0.0244***	-0.0384***
logSize	0.1083	-0.1633	0.007	-0.0078	0.0027	-0.0164
Sociallyconciuous	1.8588***	2.0932**	0.1110**	0.1055*	0.1943***	0.0118
Age	0.0073	0.0499	0.0003	0.0027	-0.0018	-0.0017
LossDev	-0.7611***	-1.3520***	-0.0335***	-0.0668***	-0.0097	-0.2148***
ExpRatio	-0.4003	-0.4921	-0.0265*	-0.0374*	-0.0293	-0.0714**
Largeblend	-8.4605***	-12.5254**	-0.4801***	-0.6001**	-0.3150	-0.9774*
Largegrowth	-7.4792***	-14.8297**	-0.4052***	-0.7483***	-0.3123	-0.9498*
Largevalue	-6.7732***	-6.835	-0.3787***	-0.2632	-0.0612	-0.6205
Midblend	-6.0349***	-10.8419*	-0.3143***	-0.4648**	-0.3255	-0.5232
Midgrowth	-4.5252***	-14.2118**	-0.2129***	-0.6378***	-0.4435**	-0.8585
Midvalue	-9.9228***	-9.611	-0.5346***	-0.3633	-0.5221	-0.8966
cons	36.2064***	73.0547***	1.4123***	3.2451***	2.7974***	5.2575***
N	571	570	571	570	571	571
r2	0.3552	0.5621	0.3131	0.4951	0.1168	0.6464

This table reports the coefficients for the regression models for different performance measures. Alpha is the Carhart's alpha measure; Sharpe is the yearly risk-adjusted return and Return is the total net return. Sustscore is the level of sustainability of the fund provided by Morningstar and Sociallyconciuous is the common dummy

variable used to analyse sociallyconscious mutual funds. N is the number of observations and r2 the R-squared fit measure. The dummies of categories have been included and compared with small mutual fund of Morningstar Style Box. *Significant at 10%; ** significant at 5% and *** significant at 1%.

Using the different elements in which ESG scores are subdivided, we have achieved similar results, finding in most models a negative relation between the dimensions of sustainability and performance (Table 8). Again, those mutual funds with higher environmental scores reduce the level of performance adjusted and non-adjusted in five of the six models estimated. For the other dimensions (social and governance) the results are quite similar, concluding that, in general, the effects of the different dimensions have a negative impact on alternative performance metrics.

Table 8- ESG pillars models and fund performance

Variable	Return2y	Return1y	Sharpe2y	Sharpe1y	Alphacat2y	Alphacat1y
Environment Score Models						
Envscore	-0.3373***	-0.7500***	-0.0169***	-0.0343***	-0.0243***	-0.0430***
logSize	0.1717**	-0.1821	0.0114**	-0.0086	0.0024	-0.0154
Sociallyconciuous	1.6921**	1.1247	0.1058**	0.0621	0.1670**	-0.0203
Age	0.0082	0.0492	0.0004	0.0027	-0.0019	-0.0017
LossDev	-0.6834***	-1.3802***	-0.0296***	-0.0680***	-0.0038	-0.2141***
ExpRatio	-0.3602	-0.4927	-0.0235	-0.0376*	-0.0302	-0.0733**
_cons	37.2674***	72.6572***	1.4858***	3.2394***	2.7454***	5.4902***
N	571	570	571	570	571	570
r2	0.3791	0.5443	0.3379	0.4834	0.1126	0.6510
Social Score Models						
Socscore	-0.3071***	-0.7310***	-0.0151***	-0.0335***	-0.0222***	-0.0378***
logSize	0.1559*	-0.2072	0.0106*	-0.0098	0.0012	-0.0164
Sociallyconciuous	1.7537**	1.4089	0.1081**	0.0751	0.1717**	-0.0152
Age	0.0094	0.053	0.0005	0.0029	-0.0018	-0.0016
LossDev	-0.7240***	-1.4174***	-0.0316***	-0.0697***	-0.0068	-0.2179***
ExpRatio	-0.3423	-0.4659	-0.0225	-0.0363	-0.0289	-0.0700*
_cons	36.9893***	73.9977***	1.4561***	3.3005***	2.7305***	5.3529***
N	571	570	571	570	571	570
r2	0.3644	0.5429	0.3252	0.4823	0.1075	0.6443
Government Score Models						
Socscore	-0.1133***	0.1867***	0.0012	-0.0318***	-0.0231***	-0.0335***
logSize	0.1907***	0.02	0.0139***	-0.011	0.0004	-0.0175
Sociallyconciuous	0.0874	-0.1171	0.0507	0.0665	0.1719**	-0.0313
Age	-0.0244**	0.0003	-0.0007	0.0033	-0.0015	-0.0011
LossDev	-0.5027***	-0.7989***	-0.0382***	-0.0712***	-0.0050	-0.2206***
ExpRatio	-	-	-0.0248*	-0.0423*	-0.0343	-0.0751**
_cons	10.3321***	7.5867	0.1834	3.1898***	2.7359***	5.1005***
N	723	729	571	570	571	570
r2	0.1128	0.2555	0.3358	0.4698	0.1067	0.6361

This table reports the coefficients for the regression models for different performance measures. Alpha is the Carhart's alpha measure; Sharpe is the yearly risk-adjusted return and Return is the total net return. Sociallyconciuous is a dummy variable used to analyse sociallyconscious mutual funds. N is the number of observations and r2 the R-squared fit measure. *Significant at 10%; ** significant at 5% and *** significant at 1%.

3.8. Downside risk and sustainability scores

In this part, we test if the degree of sustainability measured through ESG scores and their components has a positive or negative effect on the historical VaR of the portfolio.

We used the following model:

$$VaR_i = cons_i + \beta_1 Sustscore_i + \beta_2 Age_i + \beta_3 LossDev_i + \beta_4 logSIZE_i + \beta_5 ExpRat_i + \beta_6 SocRespDum + \sum_j Category_j + \varepsilon_i$$

As Table 9 shows, the downside risk of mutual funds is affected by the level of sustainability (ESG score). Specifically, we observed how the variable *Sustscore* is negatively and significantly related to the VaR of the fund at a 99% confidence level in both terms of one and two years. These results support that funds with a higher degree of sustainability better protect investors against extreme losses. As Kurtz (1997) or Goldreyer and Diltz (1999) explain, SRI mutual fund managers base their decisions on deeper, more complete, and higher quality information, resulting in a significant reduction in the risk of their investment decisions. On the other hand, the dichotomous variable commonly used has a positive sign, and opposite sign to that resulting from using a continuous variable. We also made the analysis for the different sub factors, observing again a negative and significant relationship for most of the estimated models. As can be seen in Table 9, the increase in the level of environmental, social, and governance sustainability reduces the level of extreme losses of investment funds. It is again observed that the dummy variable is significant and positively related to the level of risk. From this analysis, we observed that the results of evaluating the effect of sustainability based on dichotomous variables may yield contradictory results to those obtained when continuous variables are used.

Table 9- Sustainable score and downside risk

Variable	VaR-2y	VaR-1y	VaR-2y	VaR-1y	VaR-2y	VaR-1y	VaR-2y	VaR-1y
Sustscore	-0.0211*	-0.0277***	-	-	-	-	-	-
Envscore	-	-	-0.0193	-0.013	-	-	-	-
Socscore	-	-	-	-	-0.0256**	-0.016		
Govscore	-	-	-	-	-	-	-0.0289**	-0.0148
logSize	-0.0231	-0.0278	-0.0236	-0.0278	-0.025	-0.0285	-0.0258	-0.0291
Sociallyconciuous	0.1671*	0.3396**	0.1397	0.2730*	0.1637*	0.2872*	0.1698*	0.2820*
Age	0.0083**	0.0149***	0.0082**	0.0148***	0.0085**	0.0149***	0.0089**	0.0151***
LossDev	0.7757***	0.6817***	0.7805***	0.6745***	0.7777***	0.6753***	0.7798***	0.6744***
ExpRatio	0.0185	0.0472	0.018	0.0536	0.0163	0.0526	0.0094	0.05
Largeblend	-1.1074***	-0.5057	-1.0197***	-0.5082	-1.0188***	-0.5148	-0.9981***	-0.5246
Largegrowth	-1.0165***	-0.6747*	-0.9683***	-0.7100*	-0.9533***	-0.7043*	-0.9541***	-0.7232*
Largevalue	-1.2509***	-0.5573*	-1.1523***	-0.5113	-1.1715***	-0.5302	-1.1336***	-0.5225
Midblend	-0.7316***	-0.538	-0.7191***	-0.5515	-0.7328***	-0.5607	-0.7162***	-0.5559
Midgrowth	-0.6022***	-0.7343**	-0.6060**	-0.7415*	-0.6009***	-0.7383**	-0.5998***	-0.7388**
Midvalue	-1.0851***	-0.7785**	-1.0145***	-0.6938	-1.0971***	-0.7441*	-1.0475***	-0.7105*
cons	2.4437***	2.7915***	2.3131**	2.0917***	2.7512***	2.2913***	2.8750***	2.2177***
N	571	570	571	570	571	570	571	570
r2	0.72	0.85	0.72	0.85	0.72	0.85	0.72	0.85

This table reports the coefficients for the regression models. VaR is the maximum loss that a fund *i* can obtain for a given time period and a given confidence level. Sustscore is the level of sustainability of the mutual fund measured by Morningstar. Envscore, Socscore and Govscore are the mutual fund scores for the three dimensions (environment, social and corporate governance). Sociallyconciuous is a dummy variable used to analyse Sociallyconciuous mutual funds. N is the number of observations and r2 the R-squared fit measure. *Significant at 10%; ** significant at 5% and *** significant at 1%.

3.9. Flows and sustainability scores

In this section, we analyze the effect of sustainability on the flows of investment funds. In particular, flows of sustainable funds are generally considered to be less sensitive to changes in performance because investors value other elements in their utility function. Benson and Humphrey (2008) and Renneboog et al. (2011) obtained evidence in favor of greater stability in flows for sustainable funds, while Bollen (2007) found that SRI mutual funds are more sensitive to positive returns and less to negative ones. In line with Doven et al. (2017) and El Ghouli and Karoui (2017), we argue that funds with higher ESG scores attract more conscious investors, who are less worried about performance and therefore the flows are less sensitive to past performance. Thus, we estimate the following model to evaluate the effect of sustainability on the flow of funds using the different performance metrics (alpha, Sharpe, raw return), the sustainability score, and the interaction of the product (SustPerf: sustsharpe, sustalpha or sustreturn):

$$Flow_i = \alpha_i + \beta_1 Performance_i + \beta_2 Sustscore_i + \beta_3 SustPerf_i + \beta_4 Age_i + \beta_5 LossDev_i + \beta_6 logSIZE_i + \beta_7 ExpRatio_i + \beta_8 Sociallyconciuous + \sum_j Category_j + \varepsilon_i$$

Where:

SustPerf: is the product of Sustscore and sharpe (sustsharpe), alpha (sustalpha) or net return (sustreturn) depending on the model.

Table 10 shows that the model that takes the Sharpe ratio as the performance variable (Model 1) β_1 is not significant. This would indicate that unadjusted and adjusted returns measured by Carhart's alpha have had an influence on investment decisions. On the other hand, in Model 3, the sustainability score is also significant, so that higher-rated funds received a larger volume of funds than those with a lower score. This fact shows that the degree of sustainability stimulates fund raising and more when the degree of sustainability is higher. Also, when we analyzed the effect of the sustainability dummy variable (Sociallyconciuous), it is significant in all models, which confirms the importance of sustainability in attracting investors interested in funds that are declared sustainable. This fact can be related to both greater social awareness and expectations of greater profitability in SRIs. Finally, the negative sign of the interaction variable (sustreturn) shows the lower sensitivity of sustainable funds, supporting the results found by Doven et al. (2017) and El Ghoul and Karoui (2017) using alternative metrics and US funds.

Table 10- Sustainability score and flow of funds

Variable	Model 1	Model 2	Model 3
Sharpe-2y	0.3936		-
Alpha-2y	-	0.0772**	-
Return-2y	-		0.0224***
sustsharpe	-0.0029		-
sustalpha	-	0.0001	-
susreturn	-		-0.0002**
Sustscore	0.0028	0.0028	0.0090***

Sociallyconciuous	0.1014***	0.1050***	0.1045***
logSize	0.0243**	0.0226**	0.0295***
Age	0.0001	0.0007	-0.0004
LossDev	-0.0148**	-0.0138**	-0.0056
ExpRatio	0.04	0.0380	0.0438*
Largeblend	-0.2039**	-0.2952***	-0.1727
Largegrowth	-0.2269**	-0.2787***	-0.1537
Largevalue	-0.1957**	-0.2671***	-0.1977
Midblend	-0.2237**	-0.2327**	-0.1624
Midgrowth	-0.2316***	-0.2388**	-0.1222
Midvalue	-0.1913	-0.2642**	-0.1823
cons	-0.4622*	-0.4529*	-1.0624***
N	560	531	560
r2	0.084	0,0709	0.0995

This table reports the coefficients for the regression models. Alpha is the Carhart's alpha measure; Sharpe is the yearly risk-adjusted return and, Return is the total net return. Sustscore is the level of sustainability of the mutual fund measured by Morningstar. Sociallyconciuous is a dummy variable used to analyse socially conscious mutual funds. N is the number of observations and r2 the R-squared fit measure. *Significant at 10%; ** significant at 5% and *** significant at 1%.

5. Robustness

We conducted some additional robustness tests to check the consistency of our results and to provide other complementary analyses. We checked whether performance may differ according to the fund manager skills, considering the quantiles of different performance measures; differences in the quantiles would indicate differences in the fund manager's ability to deal with performance.

Quantile regression allowed us to capture information about the coefficients at different quantiles of the dependent variable given the set of endogenous variables. In addition, the conditional quantile regression developed by Koenker and Bassett (1978) successfully deals with skewed distributions of fund performance. In particular, we adopted the bootstrapping method proposed by Efron (1979) and implemented in the software Stata 12. Given y_i as the different performance metrics used in this paper (alpha, Sharpe and returns), and X_i as a vector of exogenous variables representing the sustainable score of each mutual fund and other controls, the quantile model can be written as:

$$y_i = X_i' \beta_\phi + u_{\phi i}$$

Assuming that:

$$Quant_{\phi}(y_i|X_i) = X_i'\beta_{\phi}$$

$$Quant_{\phi}(u_{\phi i}|X_i)=0$$

Table 11 reports quantile parameter estimates for three different adjusted risk-return performances. Our evidence for all quantiles confirms no differences in the results and sustainability seems to be important independent of the level of performance analysed.

We also calculated the models excluding the expense ratio because this variable has many blanks and reduces the sample. After the calculations, we again observed no differences with the models presented in the previous empirical analyses. Finally, we recalculated the models for each category and we obtained different results depending on the category, concluding that on average the effect is negative on performance but specific for each category.

Table 11- Quantile regression

	Return2y	Return1y	Sharpe2y	Sharpe1y	Alphacat2y	Alphacat1y
q25						
Sustscore	-0.1842***	-0.3622***	-0.0115***	-0.0205***	-0,0046	0.1326**
logSize	0.1722*	-0,1016	0,0086	-0,01	0.2409**	-0,0609
Sociallyconcious	0,999	0,4466	0,06	0,0575	0,2995	0,0173
Age	0,0149	0,0525	0,0008	0,0039*	0,0044	0,0675*
LossDev	-0.7367***	-1.1443***	-0.0234**	-0.0511***	-0.4590***	-0.6372***
ExpRatio	-0.8319***	-0.9671***	-0.0515**	-0.0706***	-1.1022***	-1.7055***
cons	30.6208***	39.3289***	1.0995***	2.0197***	3,7736	4,2163
q50						
Sustscore	-0.2361***	-0.6639***	-0.0131***	-0.0356***	-0.0687***	0,077
logSize	0.1673**	-0,0868	0,0089	-0,0079	0.2665***	-0,032
Sociallyconcious	1,0751	1.8932*	0,0815	0.1372**	0,7089	-0,1515
Age	0,0103	0,0626	0,0007	0,0043	-0,0109	0,0347
LossDev	-0.4361***	-1.2064***	-0.0204***	-0.0631***	-0.3271***	-0.6350***
ExpRatio	-0.6436**	-0.7434*	-0.0364**	-0.0661***	-0.4729**	-1.6412***
cons	30.4631***	57.1273***	1.3114***	3.0006***	4,6984	11,0889
q75						
Sustscore	-0.3204***	-0.8217***	-0.0157***	-0.0365***	-0.0992***	-0,0348
logSize	0,0408	-0,0436	-0,0024	0,0105	0,0465	-0,0804
Sociallyconcious	1.7210**	2,3816	0.1026*	0,0727	0,9631	-0,0929
Age	-0,0272	0.0826*	-0,0014	0,0042	-0,0083	0,0065
LossDev	-0.6449***	-1.5050***	-0.0331***	-0.0803***	-0.3354***	-0.4743***
ExpRatio	-0,442	0,3107	-0,0344	0,0134	-0.5127*	-1.1507***
cons	39.7659***	83.9232***	1.7920***	3.6249***	10.7921***	27.0078**

N	571	570	571	570	541	541
---	-----	-----	-----	-----	-----	-----

This table reports the coefficients for the quantile regression models (q25 or lower quartile, q50 or median and q75 or upper quartile). Sustscore is the level of sustainability of the mutual fund measured by Morningstar. Sociallyconciuous is a dummy variable used to analyse socially conscious mutual funds. N is the number of observations. *Significant at 10%; ** significant at 5% and *** significant at 1%.

6. Conclusion

In Europe, SRI strategies grew by 11.7% from 2014 to 2016 to reach \$12.04 trillion (GSIA, 2017). Traditional studies focus their work on mutual funds which declare themselves as funds that support an SRI approach. One important limitation of this approach is that results could be biased, because SRI mutual funds could have different levels of sustainability and differences with conventional funds may not be significant. Recently, Morningstar launched the Morningstar Sustainability Score to classify mutual. The use of sustainability scores in our work allows us to evaluate the effect of the degree of sustainability on performance, risk, or flows on European equity mutual funds.

Our results show that there are a large number of funds that are not declared sustainable but their portfolio is comparable to sustainable mutual funds. Furthermore, Sustainability Score is significant, explaining the level of performance for all the metrics analysed (alpha, Sharpe, and net return), and has negative sign in most models. Using a conventional dummy to declare social mutual funds, the results are significant but with the contrary sign, showing that considering the level of sustainability can help to better understand the link between performance and social responsibility. Our results are in accordance with Statman and Glushkov (2016), who concluded that the lack of clearly defined criteria to distinguish SRI mutual funds affected the results. Also, we obtained similar results to El Ghouli and Karoui (2017) for the US mutual funds market. Using the different pillars of ESG scores (environmental, social, and governance), we were able to achieve a negative link between the dimensions of sustainability and performance, showing that all the dimensions play an important role in explaining performance. Our

results are consistent with the idea that investors are paying a premium for investing in high scored mutual funds.

In terms of downside risk, the level of sustainability is negatively and significantly related to the VaR of the fund, supporting that higher scored mutual funds better protect against extreme losses. The opposite is found for the conventional dummy, showing the advantages of employing a quantitative measure of sustainability to evaluate assets' risk. This result could mean that SRI mutual fund managers base their decisions on a deeper analyses resulting in a significant reduction in the risk of their investment decisions. Our work shows that sustainability scores can be used by investors worried by extreme losses and not only by values-motivated investors.

Finally, we analyzed the effect of sustainability on the flows, confirming the importance of sustainability in attracting investors. The effect of the sustainability dummy variable is significant in all models. Unadjusted returns and Carhart's alphas have a positive influence on investment decisions. The sustainability score is significant on the flows, so higher-rated funds received a larger volume of funds. Finally, the negative sign of the interaction variable (product of sustainability and return) shows the lower sensitivity of sustainable funds. This shows the different sensitivity to performance of values-motivated investors.

Future research will benefit from the increasing amount of data to make empirical studies based on sustainability criteria. Unfortunately, due to data limitations, Morningstar Sustainability scores are only available from 2016, our sample assumes the score is constant prior to 2016. Another limitation of our work is that there may be some survivorship bias, but since our sample only includes two years, this bias must be very small.

7. References

Barnett, M.L. and R.M. Salomo, (2006). Beyond dichotomy: The curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal*, 27(11): 1101–1122.

Barreda-Tarrazona, I., Matallín-Saez, J., and M. Balaguer-Franc, (2011). Measuring Investors' Socially Responsible Preferences in Mutual Funds. *Journal of Business Ethics*, 103(2): 305- 330.

Bauer, R., Otten, R. and A.T. Rad (2006). Ethical investing in Australia: Is there a financial penalty?. *Pacific-Basin Finance Journal*, 14(1), 33-48.

Bauer, R., Derwall, J. and R. Otten (2007). The ethical mutual fund performance debate: new evidence from Canada. *Journal of Business Ethics*, 70(2): 111–124.

Bauer, R., Koedijk K. and R. Otten, (2005). International Evidence on Ethical Mutual Fund Performance and Investment Style. *Journal of Banking and Finance*, 29(7): 1751-1767.

Becchetti, L., Ciciretti, R., Dalo, A. and S. Herzel (2015). Socially Responsible and Conventional Investment Funds: Performance Comparison and the Global Financial Crisis. *Applied Economics*, 47(25): 2541-2562.

Bello, Z.Y. (2005). Socially Responsible Investing and Portfolio Diversification. *Journal of Financial Research*, 28(1): 41–57.

Benson, K.L., Brailsford, T.J. & Humphrey, J.E. (2006). Do Socially Responsible Fund Managers Really Invest Differently?. *Journal of Business Ethics*, 65(4), 337-357.

Benson, K.L. And J. Humphrey (2008) Socially Responsible Investment funds: Investor reaction to current and past returns. *Journal of Banking and Finance*, 32(9): 1850-1859.

Bilbao-Tero, A.; Álvarez-Otero, C.; Bilbao-Tero, C. and V. Cañal Fernández (2017). Hedonic evaluation of the SRI label of mutual funds using matching methodology. *International Review of Financial Analysis*, 52: 213-227.

Bollen, N.P.B. (2007). Mutual fund attributes and investor behavior. *Journal of Financial and Quantitative Analysis*, 42(3): 683-708.

Bos, J. (2017). Sustainability Scores for Investment Funds. *CFA Institute Magazine*, 28(1): 26- 28.

CFA Institute (2015). Environmental, Social, and Governance Issues in Investing: A Guide for Investment Professionals, Available at <http://www.cfapubs.org/doi/pdf/10.2469/ccb.v2015.n11.1>

Chegut, A., Schenk, H. and B. Scholtens (2011). Assessing SRI fund performance research: Best practices in empirical analysis. *Sustainable Development*, 19(2): 77–94.

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

Cortez, M.C., Silva, F. and N. Areal (2011). The performance of European socially responsible funds. *Journal of Business Ethics*, 87(4): 573–588.

Cortez, M.C., Silva, F. and N. Areal (2012). Socially responsible investing in the global market: The performance of US and European funds. *International Journal of Finance and Economics*, 17(3): 254–271.

Dolvin, S. D.; Fulkerson, J. A. and A. Krukover. (2017). Do 'Good Guys' Finish Last? The Relationship between Morningstar Sustainability Ratings and Mutual Fund Performance. Available at SSRN: <https://ssrn.com/abstract=3019403>

El Ghoul, S. and A. Karoui (2017). Does corporate social responsibility affect mutual fund performance and flows?. *Journal of Banking and Finance*, 77: 53-63.

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *Annals of Statistics*, 7(1): 1-26.

Flammer, C. (2015). Does Corporate Social Responsibility Lead to Superior Financial Performance? A Regression Discontinuity Approach. *Management Science*, 61(11): 2549-2568.

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

Gil-Bazo, J.; Ruiz-Verdú, P. and A. Santos. (2010). The performance of socially responsible mutual funds: The role of fees and management companies. *Journal of Business Ethics*, 94(2): 243—263.

Goldreyer, E. and D. Diltz (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 R. Luther (1997). Ethical until trust financial performance: Small company effects and fund size effects. *Journal of Business Finance and Accounting*, 24(5): 705–725.

Gregory A and J.M. Whittaker (2007). Performance and performance persistence of ethical unit trusts in the UK. *Journal of Business Finance and Accounting*, 34(7-8): 1327–1344.

GSIA (2017): 2016 Global sustainable Investment Review, Available at <https://www.msci.com/esg-fund-metrics>

Hamilton, S., Jo, H. and M. Statman (1993). Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds. *Financial Analysts Journal*, 49(6): 62-66.

Hill, R.P., Ainscough, T., Shank, T. and D. Manullang (2007). Corporate social responsibility and social responsible investing: A global perspective. *Journal of Business Ethics*, 70(2): 165-174.

Humphrey, J. and D.T. Tan (2014). Does it Really Hurt to be Responsible?. *Journal of Business Ethics*, 122(3): 375-386.

Humphrey, J.E., Warren, G.J. and J. Boon (2016). What is Different about Socially Responsible Funds? A Holdings-Based Analysis. *Journal of Business Ethics*, 138(2), 263-277.

Ibikunle, G. and T. Steffen (2017). European Green Mutual Fund Performance: A Comparative Analysis with their Conventional and Black Peers. *Journal of Business Ethics*, 145(2): 337-355.

Junkus, J. and T.D. Berry, (2015). Socially responsible investing: a review of the critical issues. *Managerial Finance*, 41(11): 1176-1201.

Kelly, B. and H. Jiang, (2014). Tail Risk and Asset Prices. *The Review of Financial Studies*, 27(10): 2841-2871.

Koenker, R., and G. Basset (1978). Regression Quantiles. *Econometrica*, 46(1): 33-50.

Kreander, N., Gray, R.H., Power, D.M. and C.D. Sinclair (2002). The Financial Performance of European Ethical Funds 1996-1998, *Journal of Accounting and Finance*, 1: 3-22.

Kreander, N., Gray, R.H., Power, D.M. and C.D. Sinclair (2005). Evaluating the Performance of Ethical and Non-ethical Funds: A Matched Pair Analysis. *Journal of Business Finance and Accounting*, 32(7-8): 1465-1493.

Kurtz, L. (1997). The Impact of Social Screening on Growth-Oriented Investment Strategies. *The Journal of Performance Measurement*, Spring, 65-71.

Leite, P. and M.C. Cortez (2014). Style and performance of international socially responsible funds in Europe. *Research in International Business and Finance*, 30(1): 248-267.

Leite, P. and M.C. Cortez (2015). Performance of European Socially Responsible Funds during Market Crises: Evidence from France. *International Review of Financial Analysis*, 40: 132-141

Lewis, A. and C. Mackenzie (2000). Morals, Motives and Money: The Case of U.K. Ethical Investing. *Human Relations*, 53(2): 179-191.

Luther, R.G., Matatko, J. and D.C. Corner (1992). The Investment Performance of UK 'Ethical' Unit Trusts'. *Accounting, Auditing and Accountability Journal*, 5(4): 57-70.

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

Margolis, J.D., Elfenbein, H.A. and J.P. Walsh (2009). Does it pay to be good? A meta-analysis and redirection of research on the relationship between corporate social and financial performance. Available at SSRN: <https://ssrn.com/abstract=1866371>.

Morningstar (2016a). The Morningstar Sustainable Investing Handbook. Available at <http://corporate1.morningstar.com/Morningstar-Sustainable-Investing-Handbook.pdf>

Morningstar (2016b). The Morningstar Sustainability Rating. Available at <http://corporate1.morningstar.com/Morningstar-Sustainable-Rating-Methodology.pdf>

Muñoz, F., Vargas, M. and I. Marco (2014). Environmental Mutual Funds: Financial Performance and Managerial Abilities. *Journal of Business Ethics*, 124(4): 551-569.

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

Pérez-Gladish, B.; Méndez Rodríguez, P.; M'Zali, B. and P. Lang (2013). Mutual Funds Efficiency Measurement under Financial and Social Responsibility Criteria. *Journal of Multi-Criteria Decision Analysis*, 20 (3-4): 109-125.

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

Renneboog, L, Horst, J and C. Zhang (2011). Is Ethical Money Financially Smart? Nonfinancial Attributes and Money Flows of Socially Responsible Investment Funds. *Journal of Financial Intermediation*, 20(4): 562-588.

Rivoli, P. (2003). Making a Difference or Making a Statement? Finance Research and Socially Responsible Investment. *Business Ethics Quarterly*, 13(3): 271-287.

Rudd, A. (1981). Social Responsibility and Portfolio Performance. *California Management Review*, 23(4): 55-61.

Scholtens, B. (2005). Style and performance of Dutch social responsible investment funds. *Journal of Investing*, 14(1): 63-72.

Statman, M. and D. Glushkov (2016). Classifying and Measuring the Performance of Socially Responsible Mutual Funds. *Journal of Portfolio Management*, 42(2): 140-151

Statman, M. (2000). Socially Responsible Mutual Funds. *Financial Analysts Journal*, 56(3): 30-39.

Sustainalytics (2016). Sustainalytics' Research Methodology. Company ESG Research.

Syed, A. M. (2017). Socially responsible: Are they profitable? *Research in International Business and Finance*, 42: 1504–1515.

Trinks, P.J. and B. Scholtens (2017). The Opportunity Cost of Negative Screening in Socially Responsible Investing. *Journal of Business Ethics*, 140 (2): 193–208.

Webley, P., Lewis, A. and C. Mackenzie (2001). Commitment Among Ethical Investors: An Experimental Approach. *Journal of Economic Psychology*, 22(1): 27-42.

Wimmer, M. (2012). ESG-Persistence in Socially Responsible Mutual Funds. Available at SSRN: <https://ssrn.com/abstract=2138882>

White, M. (1995). The Performance of Environmental Mutual Funds in the United States and Germany: Is there Economic Hope for “Green” Investors?. *Research in Corporate Social Performance and Policy*, Supplement 1: 323- 344.

Notes

ⁱ See <http://www.ussif.org>. US SFI is the Forum for Sustainable and Responsible Investment.

ⁱⁱ SRI funds and conventional funds have a different industry asset composition, so assets of SRI funds are highly concentrated within certain industries (renewable industries, et cetera.), while industries such as mining, oil, and gas are underweighted. Benson, Brailsford and Humphrey (2006) concluded that SRI funds have a different industry composition and that this affected performance.

ⁱⁱⁱ Sustainalytics is a global company leader in ESG research and analysis. Sustainalytics’ ESG methodology consists of approximately 150 ESG indicators to measure a company’s sustainable practices. Sustainalytics assess company performance based on several internal and external data sources (Sustainalytics, 2016): review of company reporting (annual reports, et cetera.), review of external sources (NGOs, publications, et cetera.), structural peer review, etc.

^{iv} Morningstar Portfolio Environmental Scores, Social Scores, and Governance Scores are calculated as an asset-weighted average of the scores of the individual companies.

^v See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.