Research article

Cost and performance of carbon risk in socially responsible mutual funds

Juan Carlos Matallín-Sáez* and Amparo Soler-Domínguez

Departamento de Finanzas y Contabilidad, Universitat Jaume I, E-12080 Castellón, Spain

*   Correspondence: Email: matallin@uji.es; Tel: +34964387147; Fax: +34 964728565.

Abstract: Investors and other financial actors are attracted by the role of socially responsible (SR) mutual funds in the transition to a low-carbon economy. In response to the demand for more information, Morningstar reported the level of carbon risk of funds by using the following indicators: Carbon Risk, Carbon Management, Carbon Operations risk and Carbon Exposure. Dealing with a sample of 3370 equity SR mutual funds worldwide from 2017 to 2021, this study analyzes the relationships between these indicators and the expense ratio and performance of the funds. In general, the results point to funds with lower carbon scores that have lower fees and perform better than those with higher scores. Considering the effects of the COVID-19 crisis, this evidence holds true for most of the sample period analyzed. With a spatial analysis, although the evidence generally holds, regional differences are found. Thus, funds that invest in the USA and Canada are on average cheaper and show lower carbon scores, while funds that are oriented to other areas, such as emerging markets, are more expensive and show higher scores. In summary, there is good news for the utility function of the investor and the planet: Green investing is cheaper and better.

1. Introduction

Socially responsible (SR) investments are playing a central role in prompting more and more investors and industry participants alike to apply sustainable and environmental mandates to investment activity. The global sustainable investment market has undergone an extraordinary evolution as a result, thus shaping a new investment landscape centered on standards of best practice in which all parties involved seek to firmly engage with sustainable development and the environment. The 2020 Global Sustainable Investment report[1] states that sustainable investment had reached USD 35.3 trillion by the beginning of 2020, amounting to a growth of 15% in the last two years, and that it makes up 36% of the total assets under management, estimated at USD 98.4 trillion. These figures illustrate the popularity and relevance of sustainable investment as a preferred investment vehicle for investors who are particularly sensitive to issues such as sustainability and environmental protection.

The question of how environmental contamination impacts our world is far from trivial, and climate change is now an important topic in all spheres. Global warming, clean energy ecosystems, low carbon emissions and other environmental issues are increasingly dominating the interest of scientists, politicians, academics, managers, advisors and society in general. The United Nations Sustainable Development Goals, which are the pillars of the 2030 Agenda for Sustainable Development, were conceived to reflect a universal and integrated focus on social, environmental and economic concerns to reorient societal progress toward a better and more sustainable future. The business environment has responded to this demand by embedding corporate social responsibility (CSR) practices through a wide spectrum of company initiatives and solid management strategies aligned with sustainable development. A vast body of literature examines the core reasons why engagement with sustainability is so widespread in the business context. Fourati and Dammak (2021) claim that CSR has a direct impact on corporate reputation and, consequently, a significant and positive influence on corporate financial performance. Along the same lines, other authors look in more depth at the relationship between CSR and organizational performance. Hejase et al. (2012) identified enhanced business risk management and improved competitiveness together with higher operational efficiency and cost saving. Lu et al. (2021) provide further details on the positive economic aspects of CSR for companies' financial prosperity, including improving the company image, leveraging brand equity, flattening stock volatility and boosting long-term profitability. Barauskaite and Streimikiene (2021) undertook a comprehensive review of CSR and its relationship with financial performance.

Of particular importance are the efforts made since the 2015 Paris Agreement to develop global initiatives in support of a more sustainable future and improve the regulatory system. Frameworks for action have been developed in line with the regulations in each geographical area. For instance, the Taskforce on Climate-Related Financial Disclosures (TCFD) was created in 2017 by the Financial Stability Board with the aim of improving the quality of information in reporting activity, and to deal with the impact of climate change issues. The launch of these recommendations by the TCFD has influenced policymakers and regulations globally and, furthermore, has greatly altered the expectations of financial intermediaries and investors. Specifically, the TCFD recommendations are designed to foster transparency in organizations' processes by providing information on climate-

---

[1]Available at www.gsi-alliance.org/.

related risks, but also on the opportunities and the disclosures structured around governance, strategy, risk management metrics and targets. In 2021, the Taskforce on Nature-Related Financial Disclosures (TNFD) was endorsed by major financial institutions, corporations and governments with the primary goal of delivering an honest picture of their environmental risks. In this context, global initiatives that help to make nature-related risks more visible are appreciated by interest groups—namely, investors and managers, and this has an impact on their decision-making processes. In summary, the TNFD's recommendations promote awareness of environmental degradation and pave the way toward a shift from nature-negative to nature-positive in global finance.

Given the value of sustainability in business, companies all over the world are beginning to explore the added value of best practices. For instance, investment institutions have made significant efforts to adapt their businesses and incorporate sustainable metric schemes into the market. This is the case of Morningstar, a reputable financial information enterprise. In collaboration with Sustainalytics, Morningstar has recently introduced Portfolio Carbon Metrics in an attempt to help investors to evaluate portfolio exposure to carbon risk and provide essential information that allows them to achieve a positive environmental impact through their investments (Morningstar, 2018a). For the purpose of our analysis, we selected four Morningstar scores—Carbon Risk, Carbon Management, Carbon Operations risk and Carbon Exposure—that consider actions undertaken by management to mitigate a firm's carbon risk. Notably, these indicators provide a more accurate representation of the portfolio-level carbon risk information and, in so doing, address the limitation of the traditional carbon footprint assessment (Morningstar, 2018b). The introduction of these indicators represents a significant milestone in the sustainable investment field since sustainability-themed investment advocates embedding climate information in investment decisions. The literature to date is sparse since these scores have only been available in recent years, and, to the best of our knowledge, very few studies have specifically scrutinized them. In one such study, Nofsinger and Varma (2022) examined Morningstar's carbon risk exposure for a sample of 1474 USA mutual funds and identified a set of 98 funds that are designated as sustainable. The authors focus on the implications of carbon-related disclosures for fund flows in their analysis of two subperiods (pre-disclosure vs. post-disclosure) covering the period from March 2017 to March 2019. They found that these sustainable funds had significantly lower carbon risk scores than conventional funds before the disclosure, but, during the post-disclosure period, the carbon risk scores for the sustainable funds decreased significantly more than the conventional funds. In the same vein, Reboredo and Otero (2021) used the carbon risk score for 1280 USA funds from June 2018 to December 2019; they found a negative impact in which the lower (higher) the climate-related risk, the higher (lower) the fund flows.

However, apart from Carbon Risk, the other scores mentioned above have not yet been fully assessed. Our paper presents a wider and more complete picture by demonstrating that these Morningstar indicators are highly informative not only for evaluating a portfolio's environmental risks, but also for comparing them through performance diagnosis and by running the data for a longer period (i.e., from January 2017 to May 2021). Thus, this study expands on several other aspects in the analysis of SR mutual fund behavior by examining the implications of managerial decisions on financial performance, and by looking at score levels according to the geographical area of the investments. Like other recent studies, Reboredo and Otero (2022) addressed the impact of climate change in the USA investment context; in addition, these authors identified a lack of research that analyzes whether the conclusions hold for a broader spectrum of geographical areas. To address this gap, we use a sample of 3370 SR mutual funds from across the world to examine the following: (i) the behavior of the mutual

fund carbon scores over time, with particular attention to their persistence, and the analysis of the relationships between these scores; (ii) the relationship between the mutual funds' carbon scores and their financial performance; and (iii) the connection between carbon scores and the mutual fund fees or expense ratio. We also carry out additional analyses to provide a robust assessment: first, we consider the impact of the COVID-19 crisis; second, we perform a spatial analysis. We then group the SR funds according to the Morningstar carbon scores (Carbon Risk, Carbon Management, Carbon Operations Risk and Carbon Exposure) by establishing levels of scores (Low, Mid and High); finally, we undertake a multi-level spatial analysis for comparative purposes.

The main objective of the study is to assess the performance of SR funds. In this context, one line of particular interest in the mutual fund performance literature is the comparison between SR and conventional funds. The general consensus seems to be that there are no substantial differences in performance using this categorization, as supported by abundant objective evidence (Renneboog et al., 2008; Mallett and Michelson, 2010; Climent and Soriano, 2011; Leite et al., 2018; Matallín-Sáez et al., 2019). Nevertheless, the research is increasingly examining responsible investing, and the characteristics of SR funds in particular, to identify whether different levels of commitment to the environment ultimately have an impact on the funds' results. In fact, for conventional mutual funds, Kacperczyk et al. (2005) found that, on average, funds that concentrate their holdings in industries where they have informational advantages perform better. The evidence shows that the less harmful a company's impact on the environment, the better its performance (Durán-Santomil et al., 2019; Matallín-Sáez et al., 2019).

Our second main objective is to analyze the relationship between the characteristics of the SR fund and its cost. The interconnection between performance and expense ratio is another relevant aspect when analyzing a mutual fund scheme, and it has attracted much academic attention. Broadly, an inverse relationship between expense ratio and performance is identified for conventional funds (Carhart, 1997; Elton et al., 1993; Gil-Bazo and Ruiz-Verdú, 2008; Matallín-Sáez et al., 2021), but scant evidence exists in the literature for the case of SR. For instance, Gil-Bazo et al. (2010) found that SR investors do not pay a price for having a restricted portfolio; on the contrary, a performance premium is identified for those SR funds operated by specialized management. In the same vein, Chang et al. (2019) identified a negative link between the expenses of socially conscious funds and performance; this allows investors to do good both socially and economically.

Our results show that funds with lower carbon scores are cheaper and perform better than those with higher scores. We have found an inverse relationship between the levels of carbon scores and performance, and that lower scores perform better. Note that, in the case of the Carbon Risk score, the lower (higher) the scores, the higher (lower) the intensity in terms of environmental sustainability. Consequently, funds that are ranked more highly by their carbon score achieve, in general, poorer abnormal performance. This effect is clearly identified for the USA and Canada, and it could be one reason for the recent strong growth in sustainable investment assets under management in these regions (Global Sustainable Investment Review, 2020).

Our study is based on the novel Morningstar carbon scores and contributes to the literature in several ways. First, we provide new evidence and rich perspectives on the added value that environmental investments may offer in the context of SR mutual funds. This is a pressing issue that calls for additional research: the environmental debate is ongoing and contributions from a rigorous financial perspective are needed to expand information for investors on the strategies involved in the specific SR scores, but also on the short- and long-term impacts of their investments. In our view,

this study makes a significant contribution by showing that SR scoring using Morningstar carbon metrics matters for mutual fund assessment, with a particular interest in the effects of a multi-level analysis of geographical areas. Second, the study broadens the scope of the literature by providing an international meta-analysis of the financial performance of SR funds. This contribution is particularly relevant, since knowledge of the specific risks for each constructed portfolio is highly valuable information, especially in the case of sustainable investments. Third, the analysis highlights previously documented investors' motives and pro-environmental preferences in an objective way (see Kleimeier and Viehs, 2021; Zerbib, 2019). In addition to financial preferences (see Otero-González and Durán-Santomil, 2021), we offer a set of arguments referring to the features of the SR funds worldwide that investors may want to consider in their decisions. Previous research has called for new contributions to the global commitment to sustainable development and the transition to a lower carbon economy; our study responds to this call, demonstrating that the financial market offers a premium for green investors. Fourth, this empirical study expands the idea that new sustainability practices and regulatory instruments should be aligned with sustainable development. Therefore, issues such as information transparency and legal framework deserve special mention for the way they are shaping sustainable development. International policymakers should be aware of the potential impact of environmental policies and the importance of committing to additional initiatives that embrace the major challenges ahead in the transition toward a low-carbon economy and a sustainable future.

The remainder of the paper is organized as follows. Section 2 defines the performance methodology and data used. Section 3 presents the overall results derived from the empirical analyses. Section 4 provides some concluding remarks.

2. Data and methodology

2.1. Definition of mutual fund carbon scores

The SR mutual fund label can cover different levels of social commitment. For a conscientious investor who is looking for green investment, knowing the level of carbon risk of the SR funds is valuable information. To help investors, advisors and managers, Morningstar recently introduced its Portfolio Carbon Metrics, which assess a mutual fund's carbon risk. These are computed as the asset-weighted average of the scores of the companies that the fund holds. The lower (higher) the score, the lower (higher) the carbon risk of the fund's portfolio. It should be noted that, although carbon risk is connected to the concept of carbon footprint, they are not synonymous, since carbon risk assesses how companies manage that risk. According to Morningstar (2018b), carbon risk depends on two dimensions: exposure to carbon risk and how that risk is managed.

Table 1 shows the Morningstar definitions for the carbon scores used in this study. The first, i.e., the Carbon Risk score, is one of the main and best known indicators. According to Morningstar (2018a), "[t]he Carbon Risk of a company is the evaluation by Sustainalytics of the degree to which a firm's activities and products are aligned with the transition to a low-carbon economy. The assessment includes carbon intensity, fossil-fuel involvement, stranded assets exposure, climate risk mitigation strategies, and green-solutions involvement. The portfolio Carbon Risk score is displayed as a number between 0 and 100 (a lower score is better)". The second indicator, the Carbon

Management score, assesses how a company manages risks related to carbon operations, products and services that are considered manageable. Third, the Carbon Operation score assesses the carbon risk only for the operations in the company's value chain. Finally, the Carbon Exposure score is linked to the company's value chain by operations, products and services. For some companies, a significant portion of their Carbon Exposure risk is intrinsic to their industry and cannot be effectively managed away (Morningstar, 2018b).

Table 1 shows Morningstar's definitions of the different mutual fund carbon scores.

Table 1. Mutual fund carbon scores definitions.

Carbon risk
The asset-weighted carbon risk score of the equity or corporate-bond holdings in a fund (long positions only). To calculate the portfolio carbon-risk scores, Morningstar uses Sustainalytics' company carbon-risk ratings, which indicate the risk that companies face from the transition to a low-carbon economy. At least 67% of portfolio assets must have a carbon-risk rating from Sustainalytics in order for a score to be calculated. The percentage of assets covered is rescaled to 100% before calculating the score.
Carbon management
The asset-weighted average of the carbon management score for the fund's portfolio holdings with Sustainalytics carbon research. The carbon management score is one of the key components of the carbon risk score. It evaluates a company's preparedness and track record in managing carbon operations, products and services risks that are considered manageable.
Carbon operations risk
The carbon operations risk is the asset-weighted average of the carbon operations risk score for the portfolio's holdings with Sustainalytics carbon research. A lower score is better. The carbon operations score is part of the value chain evaluation and determines carbon-related operations risks among companies held in a portfolio.
Carbon exposure
The carbon exposure score is the asset-weighted average of the carbon exposure score for the holdings with Sustainalytics carbon research in a portfolio. A lower score is better. The carbon exposure score is one of the key components of the carbon risk score. It evaluates the level of carbon-related risk that exists in the value chain among companies held in the portfolio.

Source: Morningstar Direct database.

2.2. Mutual fund performance methodology

The abnormal performance of the SR funds is measured by using a multifactor model that adjusts returns to several risk factors or benchmarks. This methodology has been widely used in previous studies assessing mutual funds (Gruber 1996; Carhart, 1997; Fama and French, 2015). An advantage of these models is that they incorporate different risk factors or benchmarks according to the type of assets in which the mutual fund invests. A multifactor model, therefore, has a greater capacity to avoid the bias caused by omitting relevant benchmarks (Pástor and Stambaugh, 2002; Matallín-Sáez, 2006). Considering the nature of the mutual funds sample in this study (all SR funds worldwide), we apply a model that includes the following factors: the FTSE World Index as a global stock market benchmark; the DJ Sustainability World, a benchmark for sustainable investment; and the FTSE Emerging Index, a benchmark for investment in emerging markets. Thus, abnormal performance is measured by the intercept or alpha (α_p) of the model (1),

$$r_{p,t} = \alpha_p + \beta_{p,w}r_{w,t} + \beta_{p,s}r_{s,t} + \beta_{p,e}r_{e,t} + \varepsilon_{p,t} \qquad (1)$$

where the returns of portfolio p in period t are represented by $r_{p,t}$, and the returns of the benchmark's global stock market, SR investment and emerging market investment are represented, respectively, by $r_{w,t}$, $r_{s,t}$ and $r_{e,t}$. All returns are computed in excess of the risk-free asset in period t.

2.3. Data

This study considers all SR equity funds worldwide according to the records of the Morningstar Direct database. The sample is free of survivorship bias, as we consider all funds, even those that have not survived. In order to provide a minimum level of robustness in the results, funds with less than one quarter of daily return data were not considered. Thus, the final sample was made up of 3,370 funds. However, the number of funds analyzed in each part of the study may vary depending on the availability of information on the funds.

Conditioned by the start date of Morningstar's information on the funds' carbon scores, the sample period analyzed ran from January 2017 to May 2021. The following information was obtained from the Morningstar Direct database for each fund: daily return index, from which the daily returns of each fund, net of management expenses and fees are calculated; geographical investment area; Carbon Risk score; Carbon Management score; Carbon Operations Risk score; Carbon Exposure score and expense ratio. The daily returns of the benchmarks used in Model (1) were also calculated from the data obtained from the Morningstar Direct database. Data for the control variables used in the robustness analysis, namely, the age, size, turnover ratio and manager tenure, were also obtained from this database. Finally, the data for the risk-free asset were obtained from Kenneth French's website.[2]

Panel A of Table 2 shows some statistics for the mutual fund sample. The mean and standard deviation (s.d.) of daily returns were calculated for each fund. Then, the SR mutual funds were grouped according to the Morningstar carbon scores. The group Low (High) was formed from 30% of the funds with the lowest (highest) scores. The remaining funds belong to the Mid group. The table reports the cross-sectional mean and s.d. of the annualized mean and the s.d. of the daily returns of the funds in each group.

In all cases, the mean returns of the funds were higher for Low funds. Furthermore, the cross-sectional s.d. of this mean was lower for funds with lower scores. For instance, for the Carbon Risk score, the annualized mean return was 18% for Low funds and 13.07% for High funds, and for the cross-sectional s.d., it was 7.93% and 9.75%, respectively. Thus, High funds show a lower mean return and higher cross-sectional dispersion (higher management risk, Matallín-Sáez et al., 2021). Regarding the risk from return volatility of the mutual funds, the last four columns of Table 2 show the cross-sectional mean and s.d. of the returns of the funds in each group. In all cases, the average risk for Low funds is slightly higher than for High funds. For instance, in the Carbon Risk score, the mean of the s.d. was 17.99% for Low funds and 17.26% for High funds. Therefore, if we compare the returns and volatility within each group, we observe that funds with low scores show a higher mean and risk than funds with higher scores.

---

[2]See http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data library.html.

However, the difference in returns in favor of Low funds is proportionally greater than the increase in risk, which is already a first indicator that Low funds perform better. Finally, Panel B of Table 2 reports the annualized mean and s.d. from the daily returns of the benchmarks used in Model (1). Comparing this with values for SR funds in Panel A, we see that the s.d. of the returns is higher in all groups of funds than for benchmarks, and that funds with low scores achieve higher mean returns.

Panel A of Table 2 shows the cross-sectional mean and s.d. (in percentage and annualized terms) of the mutual funds grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. Panel B reports the statistics (in percentage and annualized terms) from the daily returns of the benchmarks used in Model (1).

Table 2. Summary statistics of the mutual funds in the sample and benchmarks.

Panel A						
			Annualized mean of daily returns		Annualized s.d. of daily returns	
Morningstar scores		Number of funds	Mean	s.d.	Mean	s.d.
Carbon risk	Low	1,016	18.00	7.93	17.99	4.30
	Mid	1,346	14.78	7.56	17.02	3.34
	High	1,008	13.07	9.75	17.26	3.87
Carbon management	Low	1,018	16.69	8.74	18.18	4.24
	Mid	1,355	15.27	8.50	16.86	3.73
	High	997	13.70	8.33	17.28	3.37
Carbon operations risk	Low	1,016	17.53	7.90	17.95	4.22
	Mid	1,348	14.11	6.69	16.82	3.42
	High	1,006	14.44	10.82	17.56	3.85
Carbon exposure	Low	1,024	18.06	7.94	18.03	4.38
	Mid	1,338	14.90	7.60	16.86	3.36
	High	1,008	12.82	9.61	17.42	3.73
	Total	3,370				
Panel B						
Benchmark		Average annualized return		Average annualized s.d.		
FTSE World		15.23%		15.92%		
DJ Sustain World		15.33%		15.33%		
FTSE Emerging		13.42%		15.52%		

3. Results

3.1. Mutual fund carbon scores

In this section, we analyze the time behavior of the carbon scores of the SR funds in the sample. First, we are interested in evaluating the level of persistence of the scores. To do this, we consider the

quarterly data of the scores from their inception in 2018 to 2021. We calculate the cross-sectional correlation between the fund scores for each pair of consecutive quarters. Panel A of Table 3 shows the average of these correlations, taking a value of around 0.98, which implies a relatively high persistence in the scores of the funds. In the following sections, we use the time mean of the scores to analyze the relationship with performance and the expense ratio. We then explore the relationship between the carbon scores by estimating the matrix of correlations. Panel B of Table 3 shows, in general, high correlations between the variables Carbon Risk, Carbon Operations Risk and Carbon Exposure. However, the Carbon Management variable shows the lowest correlation with the other variables, probably because it differs from the others in that it only measures the part of the carbon risk managed by the company.

Panel A of Table 3 measures the persistence of the mutual fund carbon scores. For each variable, the panel shows the average of the correlation between two consecutive quarterly data sets. Panel B shows the average of the cross-sectional correlations between mutual funds' carbon scores.

Table 3. Relations between mutual fund carbon scores, 2018–2020.

Panel A			
Carbon risk		0.981	
Carbon management		0.976	
Carbon operations risk		0.982	
Carbon exposure		0.982	
Panel B			
	Carbon management	Carbon operations risk	Carbon exposure
Carbon risk	0.119	0.895	0.889
Carbon management		0.164	0.522
Carbon operations risk			0.835

3.2. Mutual fund carbon scores and financial performance

We analyze the relationship between the performance of the SR funds estimated by using their carbon scores and Model (1). The results are shown in Table 4, with the funds grouped according to their carbon scores. The Low (High) group comprises 30% of the funds with the lowest (highest) scores. The rest of the funds are in the Mid group. The central part of Panel A shows the percentage of funds in each group that achieved a negative or positive abnormal performance, and also for which funds this is significant, with a level less than or equal to 5%. For all carbon scores, the funds in the Low (High) group show a higher percentage of funds with positive (negative) alphas. The last two columns of the table present the mean and s.d. of the performance of each group of funds. In general, the mean performance decreased as the value of the carbon score increased. For instance, in the first rows of annualized performance in Panel A of Table 4, the mean abnormal performance is 3.56% for funds with a lower Carbon Risk score, while, for the higher scores, the value is 1.27%. Therefore, an inverse relationship is found between abnormal performance and the levels of carbon scores of SR funds. In this sense, Panel B of Table 4 shows how the differences between the mean performance of the funds in the Low and High groups are positive and significant in all cases, except for the Carbon Management score. For instance, for the

Carbon Risk score, the Low group funds obtained a risk-adjusted return that was 2.28% higher than the High group funds.

As the last column of Panel A in Table 4 shows, the dispersion of the cross-sectional performance within each group is greater for funds with higher carbon scores. For instance, for the Carbon Risk score, the s.d. is 7.45% for the Low group and 8.31% for the High group. In other words, funds with higher carbon risk components present a higher management risk (Matallín-Sáez et al., 2021) because the value added by managers is more dispersed than in the case of funds with lower carbon risk.

Panel A shows the performance (in percentage and annualized terms) of the mutual funds grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. Abnormal performance is obtained by regressing the excess risk-free net returns of each fund on the factors included in Model (1). The performance results are split into alpha (positive/negative) and significance. The cross-sectional mean and s.d. of the alphas in each group are also included. Panel B reports the performance differences (in percentage and annualized terms) between the different levels of carbon scores (Low minus High). The table also reports the results for statistical significance, as obtained by bootstrapping one-sided p-values.

Table 4. Mutual fund performance and carbon scores, 2017–2021.

Panel A								
			Percentage of funds in group				Annualized performance	
Morningstar scores		Number of funds	<0	p-value <=0.05	>0	p-value <=0.05	Mean	s.d.
Carbon risk	Low	1,012	27.17	0.00	72.73	4.45	3.56	7.45
	Mid	1,346	44.95	1.71	55.05	1.63	1.75	6.61
	High	1,005	46.67	3.38	53.33	1.99	1.27	8.31
Carbon management	Low	1,016	37.70	1.38	62.30	2.56	1.98	7.24
	Mid	1,353	39.69	1.40	60.24	3.40	2.51	7.31
	High	994	43.16	2.41	56.84	1.51	1.83	7.87
Carbon operations risk	Low	1,015	28.97	0.69	70.94	3.55	3.35	6.81
	Mid	1,346	44.13	1.49	55.87	1.11	1.54	6.23
	High	1,002	46.01	2.99	53.99	3.59	1.77	9.27
Carbon exposure	Low	1,013	29.91	0.69	69.99	4.44	3.19	7.59
	Mid	1,346	41.08	1.26	58.92	1.71	2.22	6.19
	High	1,004	49.10	3.29	50.90	1.89	1.01	8.65
	Total	3,363						

Panel B				
Morningstar scores			Annualized performance	p-value
Carbon risk		Low-High	2.28	(0.000)
Carbon management		Low-High	0.15	(0.314)
Carbon operations risk		Low-High	1.58	(0.000)
Carbon exposure		Low-High	2.18	(0.000)

3.3. Carbon scores and mutual fund cost

Fund expenses are an important component of the abnormal performance, as they reduce the gross return of the fund portfolio. Expenses are the price that the investor pays for the active management of the mutual fund. Sometimes, expenses are computed on the mutual fund assets or size, so they could be considered quasi-fixed expenses. Thus, over the years, they significantly reduce the final returns achieved by the investor in the long term (Matallín-Sáez et al., 2021). It is pertinent, therefore, to analyze the fees that the investor pays in SR mutual funds in relation to the carbon scores. Table 5 shows the results. The funds are again ordered according to their carbon scores and in three groups: Low, Mid and High. Panel A shows the mean expense ratio of the funds in each group. The expenses applied to funds with low carbon scores are lower than those for funds with higher carbon scores. For instance, the mean cost of the funds in the Low Carbon Risk group is 1.14% per annum, compared to the 1.35% for the High group. Panel B shows that the −0.21% difference between the two values is significant. In fact, except for the Carbon Management score, all of the differences are significant. It is worth remembering that, as we saw in Table 3, the variables Carbon Risk, Carbon Operations Risk and Carbon Exposure are highly correlated, so similar results are to be expected with these variables.

The last column of Panel A shows that, for all carbon scores, the cross-sectional s.d. of expenses in the Low group is less than in the High group. In other words, funds with higher carbon scores show greater price dispersion. In summary, the results in Table 5 show that investors in funds with higher levels of carbon risk pay a higher and more dispersed price. If we consider the quasi-fixed nature of this cost and its cumulative effect on long-term investment, this evidence implies a higher utility for investors in SR funds with lower carbon risk. Considering this result, the evidence from Table 4 and the utility in terms of social responsibility, it can be said that low carbon SR funds provide higher utility to investors. Hence, going greener is cheaper and better.

In Table 5, mutual funds are grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. Panel A shows, in percentages, the cross-sectional mean and s.d. of the annual expense ratio in each group. Panel B reports the differences in expense ratio between the different levels of carbon scores (Low minus High). The panel also reports the results for statistical significance, as obtained by bootstrapping one-sided p-values.

Table 5. Mutual fund expense ratios and carbon scores, 2017–2021.

Panel A			
		Expense ratio	
Morningstar scores		Mean	s.d.
Carbon risk	Low	1.14	0.58
	Mid	1.11	0.58
	High	1.35	0.64
Carbon management	Low	1.18	0.55
	Mid	1.18	0.66
	High	1.21	0.59
Carbon operations risk	Low	1.11	0.55
	Mid	1.15	0.61
	High	1.32	0.64
Carbon exposure	Low	1.13	0.54
	Mid	1.16	0.62
	High	1.30	0.65
Panel B			
Morningstar scores		Expense ratio difference	p-value
Carbon risk	Low-High	−0.21	(0.000)
Carbon management	Low-High	−0.03	(0.227)
Carbon operations risk	Low-High	−0.20	(0.000)
Carbon exposure	Low-High	−0.17	(0.000)

3.4. Additional and robustness analyses

The previous section reported the main results of the study. We present some complementary analyses below in order to explore this issue more widely, as well as, in some cases, a robustness analysis.

3.4.1.   COVID-19 effect analysis

The period analyzed in the previous sections includes the 2020 COVID-19 crisis. In line with Pástor and Vorsatz (2020), we split our sample period into the following subsamples: pre-crisis period (January 1, 2017 to January 31, 2020); crisis crash period (February 20 to March 23, 2020); crisis recovery period (March 24 to April 30, 2020) and post-crisis period (May 1, 2020 to May 31, 2021). For each fund and subperiod, abnormal performance was estimated by using Model (1).

Panel A of Table 6 shows the cross-sectional mean and s.d. of each mutual funds group. The results for the pre-crisis period are similar to those shown in Table 4 for the whole of the sample period. For the crash period linked to the COVID-19 crisis, the performance is notably negative. It should be noted that this subperiod is approximately one month only, so its extrapolation over the year is not comparable with the abnormal performance achieved for longer periods of time. However, in this period, the funds clearly achieved very poor results compared to the benchmarks in Model (1). In contrast, during the following month, i.e., the crisis recovery period, the abnormal performance is

notably positive, although, in most cases, it had a lower value, in absolute terms, than that of the preceding subperiod. These results suggest that, compared with the benchmarks, the capacity of the funds to provide added value is greater in a bull market than in a bear market.

Panel B of Table 6 shows the difference in performance between the Low and High SR funds according to the carbon scores. In the pre-crisis subperiod, this difference was positive and significant in all cases, similar to Panel B of Table 4. For the following two subperiods, i.e., crash and recovery, these differences were also positive and significant in most cases, notably so for the crash subperiod. These results reflect a better performance of SR funds with lower carbon scores in times of great turbulence in the markets. This evidence is in line with the study by Pástor and Vorsatz (2020). However, for the post-crisis period, i.e., from May 2020 to May 2021, this difference has a significant negative sign; that is, funds with higher carbon scores perform better. This last subperiod was characterized by an increase in global economic activity and demand for fossil fuels, which generates an increase in prices; thus, investments with higher carbon scores are likely to perform better. In this line, Bolton and Kacperczyk (2021) found that stocks of firms with higher total carbon dioxide emissions earn higher returns.

In Table 6, mutual funds are grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. Panel A shows the cross-sectional mean and s.d. of the annualized performance using Model (1) in each group and for different subperiods: pre-crisis period (January 1, 2017 to January 31, 2020); crisis crash period (February 20 to March 23, 2020); crisis recovery period (March 24 to April 30, 2020) and post-crisis period (May 1, 2020 to May 31, 2021). Panel B shows the differences in annualized performance between the different levels of carbon scores (Low minus High). The panel also reports the results for statistical significance, as obtained by bootstrapping one-sided p-values.

Table 6. Mutual fund performance, carbon scores and COVID-19 crisis.

Panel A									
		Pre-crisis		Crisis crash		Crisis recovery		Post-crisis	
Morningstar scores		Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Carbon risk	Low	5.82	5.78	−52.76	196.83	39.35	63.47	3.41	9.89
	Mid	2.51	5.99	−51.89	160.58	22.79	59.81	4.72	9.24
	High	2.78	6.72	−65.43	170.87	27.85	58.67	6.50	12.63
Carbon management	Low	3.89	5.77	−18.18	140.92	28.82	52.71	2.40	9.35
	Mid	4.02	6.33	−33.53	161.44	30.34	58.15	4.59	10.99
	High	2.69	6.77	−124.76	202.77	28.34	71.57	7.69	10.64
Carbon operations risk	Low	4.95	5.45	−43.31	178.57	30.17	60.04	3.85	9.34
	Mid	3.11	5.93	−66.07	184.42	27.45	62.76	4.06	9.33
	High	2.84	7.38	−56.07	157.80	30.84	59.48	6.94	12.92
Carbon exposure	Low	5.29	5.67	−22.69	171.23	32.59	57.24	2.78	9.60
	Mid	3.49	6.39	−53.53	178.28	30.70	63.46	4.41	9.33
	High	1.99	6.45	−93.58	168.11	24.06	60.98	7.55	12.48

Continued on next page

Panel B									
		Pre-crisis		Crisis crash		Crisis recovery		Post-crisis	
Morningstar scores		Mean	p-value	Mean	p-value	Mean	p-value	Mean	p-value
Carbon risk	Low-High	3.04	(0.000)	12.67	(0.089)	11.49	(0.000)	−3.09	(0.000)
Carbon management	Low-High	1.20	(0.000)	106.58	(0.000)	0.48	(0.456)	−5.29	(0.000)
Carbon operations risk	Low-High	2.12	(0.000)	12.76	(0.061)	−0.67	(0.411)	−3.09	(0.000)
Carbon exposure	Low-High	3.29	(0.000)	70.89	(0.000)	8.54	(0.002)	−4.77	(0.000)

3.4.2. Spatial analysis, carbon scores and performance

To avoid any local bias when comparing funds that invest in different geographical zones, the sample of SR funds was split according to area. The geographical investment areas are categorized as Global, Europe, USA and Canada, Other and, finally, Undefined when this data are not reported. Table 7 shows the distribution of the mutual funds in these categories.

All of the mutual funds are grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. The table shows the distribution of the mutual funds according to their geographical investment zone.

Table 7. Mutual fund carbon scores and spatial distribution.

Geographical investment zone	Number of funds		Percentage of funds			
			Carbon risk	Carbon management	Carbon operations risk	Carbon exposure
Global	1,333	Low	35.33	23.78	32.63	32.71
		Mid	43.96	46.51	41.41	43.21
		High	20.71	29.71	25.96	24.08
Europe	523	Low	22.94	8.41	31.55	11.09
		Mid	57.93	23.71	50.10	41.49
		High	19.12	67.88	18.36	47.42
USA and Canada	336	Low	50.89	50.00	46.43	58.93
		Mid	33.93	46.73	42.26	37.50
		High	15.18	3.27	11.31	3.57
Other	451	Low	1.11	50.11	0.44	13.08
		Mid	23.73	36.36	37.03	35.48
		High	75.17	13.53	62.53	51.44
Undefined	720	Low	34.03	36.25	35.69	36.39
		Mid	32.78	40.00	30.97	37.08
		High	33.19	23.75	33.33	26.53
Total	3,363					

From the 3,363 SR funds in the sample, the main group was made up of 1,333 funds in the Global category, that is, those funds whose investment allocation is spread worldwide. The next set of columns to the right show the percentage of funds that, within each area, were classified in Table 4 as High, Mid or Low according to the Morningstar carbon scores. If no dependency existed between the investment area and the carbon scores, the approximate distribution of the percentage of funds within each area would, theoretically, be 30% High, 40% Mid and 30% Low. However, there were deviations in all areas, two of which are most striking. First, SRI funds that invest predominantly in the USA and Canada have lower carbon scores. As an example, 50.89% of the 336 funds in this category were classified as Low Carbon Risk out of all of the 3,363 funds, whereas only 15.18% of the funds in this category had higher scores. Second, in contrast, SRI funds categorized as Other present a negligible (high) percentage of funds with lower (higher) Carbon Risk and Carbon Operations Risk scores. In the remaining areas, the deviations were smaller, although they generally tended to have a relatively lower percentage of funds with higher scores. Also, in line with the results shown in the previous tables, the Carbon Management variable is shown to behave differently from the other carbon scores in some cases.

Table 7 shows a heterogeneous distribution of funds according to their carbon scores and their geographical investment area. To analyze the relationship between the scores and performance within each area, two alternative approaches are proposed. In the first approach (Panel A of Table 8), funds are grouped in Low, Mid and High groups according to their carbon score relative to the whole sample of 3,363 SR funds. The advantage of this approach is that it maintains the grouping used in Table 4. The disadvantage is that the number of funds in the subgroups according to carbon scores may be very uneven. For instance, in the Other investment area, the percentage of funds in the Low Carbon Risk group was only 1.11% of 451 funds, i.e., this subgroup comprised just five funds, while 75.17%, that is, 339 funds, belong to the High subgroup. Comparing such unequal subgroups may lead to less robust economic interpretations of the results. For this reason, a second approach is proposed (Panel B of Table 8) in which the Low (30%), Mid (40%) or High (30%) classification is conducted separately within each geographical investment area.

Table 8 shows the difference in the means of the annualized abnormal performance between the SR funds of the Low and High carbon score groups. For the Carbon Risk score, the differences are positive and significant in all cases for both panels, except for the Other zone category. The most notable result is for the USA and Canada, where the difference between the performance of the Low and High groups was an annualized 7.85% (in Panel B). Therefore, the evidence shown in Table 4 is robust and remains within the subgroups formed by geographical area, except for the case of those funds categorized in the Other area. There may be two causes for this result: (i) as already shown in Table 7, most of the funds belonging to the Other category had higher carbon risk scores; the distance between the Low and High funds may therefore be insufficient to show differences in their performance; and (ii) the heterogeneous nature of these funds, which invest in very different areas of the world, can distort the comparison within this category.

Table 8 reports the performance differences (in percentage and annualized terms) between the different levels of carbon scores (Low minus High). Abnormal performance was obtained by regressing the excess risk-free net returns of each fund on the factors included in Model (1). The table also shows the results for statistical significance, as obtained by bootstrapping one-sided p-values. In Panel A, all mutual funds are grouped according to the Morningstar carbon scores. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds

belong to the Mid group. In Panel B, the mutual funds are also grouped according to the Morningstar carbon scores, but separately within each geographical investment zone.

Table 8. Spatial distribution, carbon scores and mutual fund performance.

Geographical investment zone		Low-High			
		Carbon risk	Carbon management	Carbon operations risk	Carbon exposure
Panel A					
Global	Annu. performance	2.38	−0.59	1.02	1.31
	p-value	(0.000)	(0.196)	(0.062)	(0.022)
Europe	Annu. performance	2.70	0.70	1.10	2.26
	p-value	(0.001)	(0.085)	(0.053)	(0.000)
USA and Canada	Annu. performance	6.82	2.19	4.91	6.23
	p-value	(0.000)	(0.025)	(0.000)	(0.000)
Other	Annu. performance	−0.67	1.01	7.40	−0.73
	p-value	(0.255)	(0.191)	(0.000)	(0.255)
Undefined	Annu. performance	1.05	0.79	0.73	1.39
	p-value	(0.016)	(0.085)	(0.092)	(0.006)
Panel B					
Global	Annu. performance	1.87	−0.41	1.26	1.06
	p-value	(0.002)	(0.268)	(0.025)	(0.049)
Europe	Annu. performance	2.53	0.55	0.94	1.89
	p-value	(0.000)	(0.191)	(0.091)	(0.001)
USA and Canada	Annu. performance	7.85	−0.96	6.00	6.58
	p-value	(0.000)	(0.217)	(0.000)	(0.000)
Other	Annu. performance	0.33	0.98	2.45	0.24
	p-value	(0.359)	(0.167)	(0.016)	(0.404)
Undefined	Annu. performance	1.02	1.30	0.67	1.89
	p-value	(0.026)	(0.009)	(0.112)	(0.000)

The differences in the Low-High performance shown in Table 8 for the Carbon Operations Risk and Carbon Exposure are positive and significant in most of the cases. Again, the greatest Low-High differences are seen in the USA and Canada category, with an annualized 6% for the Carbon Operations Risk and 6.58% for the Carbon Exposure (in Panel B). Finally, the evidence for Carbon Management is not so clear. Only the difference for the Undefined category funds is significant, as shown in Panel B. This result is in line with that shown in Table 4. Indeed, Table 3 already showed that the variable Carbon Management had the lowest correlation with the other carbon scores; therefore, it is not surprising to find differences in the results of the analysis carried out.

In summary, the spatial analysis showed that, in general, the evidence found in Table 4 is not driven by any local bias. Thus, in most cases, and in aggregate terms, funds with low carbon scores outperform those with higher scores. Also, this effect is greater for funds whose geographical area of investment is the USA and Canada, which amounts to a broad, deep, mature and well-defined market. In contrast, this effect is not so noticeable for the funds that invest in the geographical zones

categorized as Other, possibly due to the multinational heterogeneity of this category and the presence of emerging markets.

3.4.3. Spatial analysis, carbon scores and mutual fund cost

In this subsection, as in Subsection 3.3 and Table 5, the relationships between the carbon scores and the cost of the SR funds are analyzed with regard to the funds' geographical investment area. Funds are grouped according to the Morningstar carbon scores separately within each investment zone. Table 9 reports the results of this spatial analysis. Panel A shows, in percentages, the cross-sectional mean of the annual expense ratio for each fund group. The more expensive funds are seen to be in the geographical investment area Other (1.53%, 1.38% and 1.63% for Low, Mid and High, respectively), and the least expensive are in the USA and Canada area (1%, 1.06% and 1.04%, respectively). However, in Table 7, we see that the funds of the Other (USA and Canada) area have predominantly higher (lower) carbon scores. Thus, part of the evidence for a positive relationship between the level of carbon scores and the cost shown in Table 5 could be driven by the spatial distribution of the SR funds. To test this, similar to Panel B of Table 5, we computed the difference in the expense ratio between the Low and High carbon score groups. The results are shown in Panel B of Table 9. The differences for the Carbon Risk score were negative, but only significant for the Undefined case.

In Table 9, mutual funds are grouped according to the Morningstar carbon scores separately within each geographical investment zone. The group Low (High) was formed from the 30% of funds with the lowest (highest) scores. The remaining funds belong to the Mid group. Panel A shows, in percentages, the cross-sectional mean of the annual expense ratio in each group. Panel B reports the differences in expense ratio between the different levels of carbon scores (Low minus High). The panel also reports the results for statistical significance, as obtained by bootstrapping one-sided p-values.

For Carbon Operations Risk and Carbon Exposure, the differences were negative (except for the USA and Canada), but, in general, only significant for Europe, Other and Undefined areas. There is less evidence supporting a positive relationship between carbon scores and cost than that shown in Table 5. Therefore, the evidence holds for some areas, but not for others.

In summary, and in general, comparing the results of Tables 5 and 9 and considering the evidence of Table 8, the following results can be highlighted. The differences in performance according to the carbon score level were much greater than the differences with respect to the expense ratio. Investors of funds in the USA and Canada have the cheapest funds on average, and, although there is no difference in the cost according to the level of the carbon score, it is evident that funds with low carbon scores perform better than those with high scores. In contrast, in general, investors in funds of the other group pay higher commissions, and, although there are some slightly cheaper funds with low carbon scores, there is no evidence of a significant relationship between the carbon score level and the fund's overall performance.

Table 9. Spatial distribution, carbon scores and mutual fund cost.

Panel A					
		Mean of expense ratio			
Geographical investment zone		Carbon risk	Carbon management	Carbon risk operations	Carbon exposure
Global	Low	1.30	1.20	1.24	1.31
	Mid	1.10	1.21	1.19	1.10
	High	1.35	1.30	1.28	1.32
Europe	Low	1.26	1.31	1.08	1.16
	Mid	1.27	1.27	1.39	1.43
	High	1.41	1.35	1.43	1.30
USA and Canada	Low	1.00	1.22	1.02	1.05
	Mid	1.06	0.89	1.07	1.05
	High	1.04	1.02	1.01	1.01
Other	Low	1.53	1.39	1.31	1.07
	Mid	1.38	1.55	1.55	1.72
	High	1.63	1.53	1.64	1.69
Undefined	Low	1.01	1.14	0.97	0.96
	Mid	0.97	1.04	1.00	1.07
	High	1.23	1.00	1.23	1.13
Panel B					
		Low-High			
Geographical investment zone		Carbon risk	Cabon management	Carbon operations risk	Carbon exposure
Global	Expense ratio	−0.05	−0.10	−0.04	−0.01
	p-value	(0.264)	(0.114)	(0.324)	(0.480)
Europe	Expense ratio	−0.15	−0.04	−0.35	−0.14
	p-value	(0.179)	(0.405)	(0.012)	(0.165)
USA and Canada	Expense ratio	−0.04	0.20	0.01	0.04
	p-value	(0.348)	(0.020)	(0.455)	(0.358)
Other	Expense ratio	−0.10	−0.14	−0.32	−0.62
	p-value	(0.324)	(0.153)	(0.047)	(0.000)
Undefined	Expense ratio	−0.22	0.15	−0.26	−0.17
	p-value	(0.005)	(0.019)	(0.000)	(0.021)

3.4.4. Econometric analysis with control variables

Finally, in this subsection, we propose a robustness analysis by using a different methodology and considering control variables that could influence the results. In the previous sections, we examined the differences between groups to analyze the relationship between performance and carbon scores; however, in what follows, we propose an alternative approach in which we consider the following regression model:

$$\alpha_p = b_1 cs_p + b_2 age_p + b_3 p + b_4 tur_p + b_5 mg_p + \sum b_i D_{p,i} + \varepsilon_p \tag{2}$$

where the dependent variable is the annualized abnormal performance of the fund p, measured as the intercept (α_p) of Model (1), and the first independent variable is the carbon score of the fund (cs_p).

The rest of the independent variables are control variables: the natural logarithm of the years of the fund since inception (age_p), the natural logarithm of the total net assets managed by the fund ($size_p$), the turnover ratio of the fund (tur_p), the natural logarithm of the manager tenure (mg_p) and five dummy variables ($D_{p,i}$) for each geographical investment zone that capture any domestic bias in the performance of the mutual fund.

Panels A and B of Table 10 show the results of the estimation of Model (2). The dependent variable is the annualized abnormal performance of fund p, measured as the intercept (α_p) of Model (1), and the first independent variable is the carbon score of the fund (cs_p). The rest of the independent variables are control variables: the natural logarithm of the years of the fund since inception (age_p), the natural logarithm of the total net assets managed by the fund ($size_p$), the turnover ratio of the fund (tur_p), the natural logarithm of the manager tenure (mg_p) and five dummy variables ($D_{p,i}$) for each geographical investment zone. P-values (in parentheses) are from the White heteroskedasticity-consistent standard error covariance. Panel A shows the results for net abnormal performance, and Panel B shows those for the gross abnormal performance. Panel C shows the differences between the coefficients of the carbon scores from Panels A and B. P-values (in parentheses) are from the equality of means test of the bootstrap distribution of the estimated coefficients.

This robustness analysis has several advantages. First, it evaluates the relationship between performance and carbon risk without having to divide the sample and group the funds into Low, Mid and High; second, it controls for the possible relationship between performance and other explanatory variables.

Panel A of Table 10 presents the estimation results for Model (2) for each carbon score. The first column shows the results for the Carbon Risk score. The coefficient takes a negative value, i.e., −0.005, and it is statistically significant. From this, we infer a negative relationship between the Carbon Risk score of the funds and their abnormal performance, that is, the same result as shown in Panel B of Table 4. In fact, the sign and significance of the coefficients of the carbon scores are in line with those shown in Table 4. Thus, we find a negative and significant relationship between the performance and the Carbon Operations Risk score, and also with respect to the Carbon Exposure score. And, as in Table 4, there is no evidence of this relationship in the case of the Carbon Management score.

In summary, the previous results shown in Table 4 are strong, as they hold even though the control variables are sufficiently significant to explain the performance of the funds. Thus, a negative relationship is found between performance and mutual fund age. The positive relationships with respect to size, turnover and management tenure are also significant. Regarding the dummies for geographic investment zone, evidence of a domestic bias only appears in some cases. For instance, in the case of the Carbon Risk score, the funds in the Global and Other areas achieved better performance on average. But, in short, what is relevant in terms of robustness is that considering these explanatory variables does not affect the previous evidence of a negative relationship between performance and carbon scores.

The second main result, shown in Table 5, is the evidence of lower expense ratios for mutual funds with lower scores in Carbon Risk, Carbon Operations Risk and Carbon Exposure. As in these cases, lower scores also imply better performance; a consequence of using the returns net of expenses to assess performance would be to strengthen the inverse relationship between abnormal performance and carbon scores. To analyze this question, we apply Model (2) but take the annualized gross abnormal performance as the dependent variable. This procedure is a robustness analysis for

the previous results of the study and of Panel A in Table 10, thus demonstrating that they are not driven by the positive relationship between carbon scores and expenses. On the other hand, if the evidence in Table 5 is true, the negative coefficients that measure the relationship between gross performance and carbon scores should be lower (in absolute value) than those obtained by using the net abnormal performance in Panel A.

Table 10. Econometric analysis with control variables.

Panel A								
Variable	Carbon risk	p-value	Carbon management	p-value	Carbon oper. risk	p-value	Carbon exposure	p-value
Carbon score	−0.005	(0.000)	0.000	(0.095)	−0.005	(0.000)	−0.003	(0.000)
Age	−0.005	(0.004)	−0.006	(0.002)	−0.006	(0.001)	−0.005	(0.006)
Size	0.002	(0.000)	0.002	(0.000)	0.002	(0.000)	0.003	(0.000)
Turnover	0.012	(0.000)	0.012	(0.000)	0.012	(0.000)	0.012	(0.000)
Management	0.007	(0.000)	0.005	(0.003)	0.006	(0.000)	0.006	(0.000)
Global	0.019	(0.030)	−0.001	(0.882)	0.010	(0.257)	0.020	(0.022)
Europe	−0.001	(0.936)	−0.022	(0.026)	−0.013	(0.129)	0.006	(0.529)
USA and Canada	0.004	(0.609)	−0.013	(0.160)	−0.003	(0.731)	0.003	(0.719)
Other	0.021	(0.024)	−0.015	(0.117)	0.004	(0.685)	0.013	(0.180)
Undefined	−0.005	(0.472)	−0.031	(0.000)	−0.017	(0.015)	−0.006	(0.433)
R^2	0.169		0.113		0.140		0.155	
Panel B								
Carbon score	−0.004	(0.000)	0.000	(0.079)	−0.003	(0.000)	−0.002	(0.000)
Age	−0.001	(0.671)	0.000	(0.856)	−0.001	(0.564)	−0.001	(0.690)
Size	0.001	(0.009)	0.002	(0.005)	0.001	(0.007)	0.001	(0.008)
Turnover	0.005	(0.087)	0.003	(0.292)	0.004	(0.189)	0.004	(0.172)
Management	0.012	(0.000)	0.010	(0.000)	0.010	(0.000)	0.011	(0.000)
Global	0.014	(0.270)	−0.006	(0.660)	−0.004	(0.725)	0.015	(0.258)
Europe	0.000	(0.982)	−0.015	(0.290)	−0.017	(0.151)	0.006	(0.652)
USA and Canada	0.006	(0.649)	−0.014	(0.294)	−0.011	(0.363)	0.004	(0.749)
Other	0.040	(0.006)	0.007	(0.630)	0.013	(0.366)	0.031	(0.031)
Undefined	0.015	(0.173)	−0.008	(0.521)	−0.005	(0.614)	0.014	(0.216)
R^2	0.147		0.081		0.091		0.128	

Panel C				
Variable	Carbon risk	Carbon management	Carbon operations risk	Carbon exposure
Panel A	−0.005034	−0.000269	−0.005400	−0.002711
Panel B	−0.004493	−0.000367	−0.002853	−0.002228
Difference	−0.000541	0.000098	−0.002547	−0.000483
p-value	(0.0000)	(0.0002)	(0.0000)	(0.0000)

The results are shown in Panel B of Table 10. When compared with Panel A, we first find that the evidence for the relationship between scores and performance holds and is not driven by expenses. Thus, in the first row, the coefficient for the carbon scores remains negative and significant for Carbon Risk, Carbon Operations Risk and Carbon Exposure. Second, as we predicted, these coefficients take lower absolute values. Accordingly, Panel C shows the differences between the coefficients of the carbon scores from Panels A and B, where p-values represent the equality of means test of the bootstrap distribution of the estimated coefficients. This panel shows how these differences are negative and significant for Carbon Risk, Carbon Operations Risk and Carbon Exposure. On the other hand, in Panel B, there are changes to the estimates of some control variables; thus, the variables age and turnover are no longer significant.

In summary, in this section, we conducted a robustness analysis with a different methodology than the one used in the previous sections of the study, thereby avoiding the presence of biases due to the thresholds for grouping funds due to the geographical zones of investment, or due to other control variables with explanatory capacity for performance. This robustness analysis confirms the results obtained previously.

4. Conclusions

Interest in SR investment from investors, advisors, intermediaries and managers is increasing, as reflected in both the growth of SR funds and the information on the subject. In addition to the SR label, some companies are now offering investors a wide range of data that enhance and expand knowledge about these funds. One such case is Morningstar, which provides several variables to measure the level of carbon risk of the funds, namely, the Carbon Risk score, the Carbon Management score, the Carbon Operations score and the Carbon Exposure score. These variables respond to the demand for information related to the widespread growing concern about the effects of climate change and the need for a transition to a low-carbon economy. For a mutual fund, these scores are calculated by weighting the carbon scores of the companies in each fund's portfolio.

In this context, the objective of the study was to analyze the relationship between the funds' carbon scores and their performance and cost. To this end, we selected a sample of SR funds from across the world. First, the results show that the carbon scores are highly correlated, with the exception of Carbon Management risk, which only measures the level of carbon risk that the company managers have control over, and it does not consider the systematic carbon risk intrinsic to the company's activity in its sector. Second, we found that, on average, funds with low carbon scores outperform those with high scores. The difference between the Low and High performance was an annualized 2.28% for the Carbon Risk score, 1.58% for the Carbon Operations Risk score and 2.18% for the Carbon Exposure score. However, the difference was not significant for Carbon Management risk.

We tested robustness with three additional analyses. First, we evaluated the effect of the COVID-19 crisis, finding that the evidence holds for the subperiods before the crisis, as well as the subperiods of the crisis itself and the subsequent recovery period. Only in the case of the last post-crisis subperiod linked to a reactivation of economic activity and a higher demand for fossil fuels did the evidence not hold, as funds with higher carbon scores performed better. The second robustness test was a spatial analysis that split the sample of SR funds into their geographical areas of investment. In general, the previous evidence holds for all funds in each geographical area, except for the group of funds that invests outside of Europe, the USA and Canada. The

negative relationship between performance and carbon scores was also found to be stronger in those SR funds that invest in the USA and Canada. Third, we performed a regression analysis that avoids biases due to the thresholds for grouping funds due to the geographical zones of investment, or due to other control variables with explanatory capacity for performance. This robustness analysis confirms the results obtained previously.

Third, we analyzed the relationship between carbon scores and the cost of the SR funds, as measured by using the expense ratio borne by investors. We found that, for the variables Carbon Risk, Carbon Operations Risk and Carbon Exposure, the cost of funds with lower carbon scores is lower than for those with higher scores. Although the differences in performance between Low and High scores due to cost show a lower value than the total amount of the differences in performance between Low and High funds, it is nevertheless important to highlight the effect of the cost borne by the investors, as it impacts the funds' long-term performance. The robustness of these results was also tested by means of a spatial analysis in which the funds were grouped according to the geographical area where the investments were made. The results show that, although the earlier evidence partially holds, the evidence found in the analysis of all of the funds could be driven by the heterogeneous spatial distribution of the SR funds. As can be seen, the SR funds from the USA and Canada area show lower cost and lower carbon scores, while the funds categorized in Other investment zones show the highest cost and the highest carbon scores.

In summary, this study shows that investors, advisors, intermediaries and managers should evaluate, aside from the SR label, the characteristics that define the SR investment. The evidence found in this study points to the fact that, in general, funds with lower carbon risk are cheaper and perform better. This is good news for the utility function of investors, and for the planet.

Acknowledgments

Conflict of interest

The authors declare no conflict of interest.

References

Barauskaite G, Streimikiene D (2021) Corporate social responsibility and financial performance of companies: The puzzle of concepts, definitions and assessment methods. *Corp Soc Resp Env Ma* 28: 278–287. https://doi.org/10.1002/csr.2048

Bolton P, Kacperczyk M (2021) Do investors care about carbon risk? *J Financ Econ* 142: 517–549. https://doi.org/10.1016/j.jfineco.2021.05.008

Carhart MM (1997) On persistence in mutual fund performance. *J Financ* 52: 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chang CE, Krueger TM, Witte HD (2019) Saving green while going green: The role of expenses in socially responsible equity mutual fund selection. *Manag Financ* 45: 21–35. https://doi.org/10.1108/MF-03-2018-0095

Climent F, Soriano P (2011) Green and good? The investment performance of US environmental mutual funds. *J Bus Ethics* 103: 275–287. https://doi.org/10.1007/s10551-011-0865-2

Durán-Santomil P, Otero-González L, Correia-Domingues RH, et al. (2019) Does sustainability score impact mutual fund performance? *Sustainability* 11: 2972. https://doi.org/10.3390/su11102972

Elton EJ, Gruber, MJ, Blake CR (1996) Survivor bias and mutual fund performance. *Rev Financ Stud* 9: 1097–1120. https://doi.org/10.1093/rfs/9.4.1097

Elton EJ, Gruber MJ, Das S, et al. (1993) Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Rev Financ Stud* 6: 1–22. https://doi.org/10.1093/rfs/6.1.1

Fama EF, French KR (2015) A five-factor asset pricing model. *J Financ Econ* 116: 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fourati YM, Dammak M (2021) Corporate social responsibility and financial performance: International evidence of the mediating role of reputation. *Corp Soc Resp Env Ma* 28: 1749–1759. https://doi.org/10.1002/csr.2143

Gil-Bazo J, Ruiz-Verdú P (2008) When cheaper is better: Fee determination in the market for equity mutual funds. *J Econ Behav Organ* 67: 871–885. https://doi.org/10.1016/j.jebo.2007.04.003

Gil-Bazo J, Ruiz-Verdú P, Santos AA (2010) The performance of socially responsible mutual funds: The role of fees and management companies. *J Bus Ethics* 94: 243–263. https://doi.org/10.1007/s10551-009-0260-4

Gruber MJ (1996) Another puzzle: The growth in actively managed mutual funds. *J Financ* 51: 783–810. https://doi.org/10.1111/j.1540-6261.1996.tb02707.x

Global Sustainable Investment Alliance (2020) *Global Sustainable Investment Review. Biennial Report.* Available from: https://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf.

Hejase H, Farha C, Haddad Z, et al. (2012) Exploring the multiple benefits of CSR on organizational performance: Case of Lebanon. *J Soc Sci* 1: 1–23. https://doi.org/10.25255/jss.2012.1.1.1.23

Kacperczyk M, Sialm C, Zheng L (2005) On the industry concentration of actively managed equity mutual funds. *J Financ* 60: 1983–2011. https://doi.org/10.1111/j.1540-6261.2005.00785.x

Kleimeier S, Viehs M (2021) Pricing carbon risk: Investor preferences or risk mitigation? *Econ Lett* 205: 109936. https://doi.org/10.1016/j.econlet.2021.109936

Leite C, Cortez MC, Silva F, et al. (2018) The performance of socially responsible equity mutual funds: Evidence from Sweden. *Bus Ethics* 27: 108–126. https://doi.org/10.1111/beer.12174

Lu J, Liang M, Zhang C, et al. (2021) Assessment of corporate social responsibility by addressing sustainable development goals. *Corp Soc Resp Env Ma* 28: 686–703. https://doi.org/10.1002/csr.2081

Mallett J, Michelson S (2010) Green investing: Is it different from socially responsible investing? *Intl J Bus* 15: 395–410.

Matallín-Sáez JC (2006) Seasonality, market timing and performance amongst bench- marks and mutual fund evaluation. *J Bus Financ Account* 33: 1484–1507. https://doi.org/10.1111/j.1468-5957.2006.00636.x

Matallín-Sáez JC, Soler-Domínguez A, de Mingo-López DV (2021) On management risk and price in the mutual fund industry: style and performance distribution analysis. *Risk Manag* 23: 150–171. https://doi.org/10.1057/s41283-021-00072-9

Matallín-Sáez JC, Soler-Domínguez A, de Mingo-López DV, et al. (2019) Does socially responsible mutual fund performance vary over the business cycle? New insights on the effect of idiosyncratic SR features. *Bus Ethics* 28: 71–98. https://doi.org/10.1111/beer.12196

Morningstar (2018a) Morningstar portfolio carbon metrics frequently asked questions. Hale, J and Koska, J, Morningstar Research.

Morningstar (2018b) Measuring transition risk in fund portfolios. The Morningstar portfolio carbon risk score. Hale, J, Morningstar Research.

Nofsinger JR, Varma A (2022) Keeping Promises? Mutual Funds' Investment Objectives and Impact of Carbon Risk Disclosures. *J Bus Ethics,* 1–24. https://doi.org/10.1007/s10551-022-05264-1

Otero-González L, Durán-Santomil P (2021) Is quantitative and qualitative information relevant for choosing mutual funds? *J Bus Res* 123: 476–488. https://doi.org/10.1016/j.jbusres.2020.10.015

Pástor L, Stambaugh R (2002) Mutual fund performance and seemingly unrelated assets. *J Financ Econ* 63: 315–349. https://doi.org/10.1016/S0304-405X(02)00064-8

Pástor L, Vorsatz MB (2020) Mutual fund performance and flows during the COVID-19 crisis. *Rev Asset Pricing Stud* 10: 791–833. https://doi.org/10.1093/rapstu/raaa015

Reboredo JC, Otero LA (2021) Are investors aware of climate-related transition risks? Evidence from mutual fund flows. *Ecol Econ* 189: 107148. https://doi.org/10.1016/j.ecolecon.2021.107148

Reboredo JC, Otero LA (2022) Low carbon transition risk in mutual fund portfolios: Managerial involvement and performance effects. *Bus Strateg Environ* 31: 950–968. https://doi.org/10.1002/bse.2928

Renneboog L, Horst JT, Zhang C (2008) The price of ethics and stakeholder governance of socially responsible funds. *J Corp Financ* 14: 302–322. https://doi.org/10.1016/j.jcorpfin.2008.03.009

Zerbib OD (2019) The effect of pro-environmental preferences on bond prices: Evidence from green bonds. *J Bank Financ* 98: 39–60. https://doi.org/10.1016/j.jbankfin.2018.10.012