Finance Research Letters 87 (2026) 109002

Contents lists available at ScienceDirect

Finance Research Letters

journal homepage: www.elsevier.com/locate/frl

Screening US mutual funds for social responsibility and irresponsibility[☆,☆☆]

Jakob Artinger[a], Christian Kreuzer[b,*]

[a] Deutsche Bundesbank, 80539 München, Germany
[b] Department of Finance, University of Regensburg, Germany

ARTICLE INFO

Keywords:
Mutual funds
Sustainable finance
Corporate scandals
CSR
ESG
SRI

ABSTRACT

This study examines 139 socially responsible and 4,898 conventional U.S. mutual funds from 2013–2020, analyzing social responsibility and irresponsibility by screening fund holdings. SR funds show higher weighted ESG ratings and subscores than conventional funds but underperform regarding corporate scandals, driven by social and governance issues, with no significant differences in environmental scandals. Passive SR funds tend to show higher weighted ESG scores than active SR funds. The findings offer valuable insights for ethical investors and fund managers to enhance investment strategies and decision-making.

1. Introduction

Socially responsible investing (SRI), which includes environmental, social, and governance (ESG) aspects into investment decisions (Eurosif, 2016), has gained significant importance in recent years. In mutual funds, incorporating ESG factors reflects ethical and social responsibility considerations (US SIF, 2018). Growing awareness of environmental and social risks has fueled demand for such funds, encouraging fund managers to emphasize ESG and market their funds as sustainable.

In 2021, the SEC's ESG Risk Alert[1] raised concerns about greenwashing in ESG funds, highlighting inconsistencies between ESG disclosures and actual practices, as well as overreliance on composite ESG scores as a sign of weak due diligence and compliance (Raghunandan and Rajgopal, 2022).

The SEC and DOL are undertaking initiatives in response to growing concerns surrounding ESG investing. To address the risk of greenwashing, particularly through fund names using terms such as 'ESG', 'green', or 'sustainable' that may misleadingly suggest compliance with ESG standards, the SEC has proposed amendments to the Investment Company Act's "Names Rule". As in the original 2001 rule, the 2023 version requires funds with names suggesting a focus on certain investments, industries, or regions to invest at least 80% of their assets in line with their names.[2] The amendments broaden this requirement to include names implying that a fund invests in assets with specific characteristics, such as ESG-related traits, thereby addressing misleading fund labeling practices and aiming to enhance consistency between fund portfolios and investor expectations (Fisch and Robertson, 2022). Besides, the DOL has implemented a rule that introduces potential legal risks for retirement plans that include ESG funds (Curtis et al., 2021).

☆ This article is part of a Special issue entitled: 'YSBC Sustainable Finance 2024' published in Finance Research Letters.

☆☆ The contributions to this study represent the authors' personal opinions and do not necessarily reflect the views of the Deutsche Bundesbank.

* Corresponding author.
E-mail addresses: jakob.artinger@bundesbank.de (J. Artinger), christian.kreuzer@wiwi.uni-regensburg.de (C. Kreuzer).

1 For more detailed information see https://www.sec.gov/newsroom/speeches-statements/peirce-statement-staff-esg-risk-alert.

2 For more detailed information see https://www.sec.gov/rules-regulations/staff-guidance/division-investment-management-frequently-asked-questions/2025-names-rule-faqs.

https://doi.org/10.1016/j.frl.2025.109002
Received 11 June 2024; Received in revised form 6 November 2025; Accepted 17 November 2025

Available online 24 November 2025
1544-6123/© 2025 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Unlike the European Union, the United States currently lacks a regulatory framework comparable to the EU's Sustainable Finance Disclosure Regulation (SFDR) and taxonomy. The absence of a regulatory framework increases the need for ethically motivated (passive) investors to rely on alternative metrics and scores to accurately assess the sustainability of mutual funds and to ensure alignment with their investment objectives. In doing so, CSR actions, as typically measured by ESG ratings, play an important role in this evaluation process. While Peng et al. (2023) find that SR mutual funds positively affect the ESG performance of their holdings, they also demonstrate that firms with higher ESG performance are more attractive to SR funds, underscoring the importance of ESG ratings for these funds.

However, a comprehensive evaluation of corporate responsibility must include both CSR activities (measured by ESG scores) and corporate social irresponsibility (CSI), such as harmful actions or scandals. Notably, CSI is not necessarily synonymous with poor CSR efforts, which may result in low ESG ratings. Moreover, the concepts of CSR and CSI are not perfectly inverse (Arora and Dharwadkar, 2011; Aouadi and Marsat, 2018). In this work, we investigate various dimensions of CSR as well as CSI of SR mutual funds compared to conventional funds. While various studies have examined the social performance of SR funds compared to conventional funds, some authors, (e.g. Kempf and Osthoff (2008), Muñoz et al. (2022) and Muñoz (2021)) claim that SR funds demonstrate a higher level of social responsibility, while others like Wimmer (2012), Utz et al. (2014) and Gangi and Varrone (2018) argue against sustained higher social performance of SR funds. A growing strand of academic literature examines aspects of social responsibility and irresponsibility in the context of mutual funds (Borgers et al., 2015; Dorfleitner et al., 2021; Petridis et al., 2023). Besides, other authors examine greenwashing in the context of mutual funds (Raghunandan and Rajgopal, 2022; Kim and Yoon, 2023; Kaustia and Yu, 2021; Abouarab et al., 2024; Guidolin and Magnani, 2024). While some authors investigate single aspects of corporate scandals or incidents (Raghunandan and Rajgopal, 2022; Kim and Yoon, 2023), a deeper insight into corporate scandals in the field of mutual SR funds is still lacking. In particular, the key question of how the CSI performance of socially responsible (SR) mutual funds differs from that of conventional funds remains unanswered. Our work contributes to the CSR-based academic literature in the context of mutual funds by shifting the focus toward CSR-based corporate scandals and also contributes to the growing literature on ESG fund greenwashing. In this paper, we show that SR mutual funds demonstrate a higher ESG performance than conventional mutual funds but perform worse in the context of CSI. While no differences are observed using a scoring-based CSI methodology, significant differences emerge when examining the absolute number of corporate scandals, with SR funds showing higher scandal counts. Additionally, passive SR funds demonstrate higher ESG performance compared to active SR funds.

1.1. Hypothesis development

Following Utz et al. (2014), SR mutual funds usually perform a two-stage process to identify equity holdings. First, a screening process, which a priori sorts out companies that do not meet the social responsibility requirements. Second, an asset allocation that allocates fund holdings based on the remaining companies (e.g., "best in class" approach). Academic literature monitoring SRI or ESG funds mostly shows that these funds exhibit higher ethical standards than conventional funds (Kempf and Osthoff, 2008; Curtis et al., 2021). Moreover, Peng et al. (2023) find a positive effect of SR mutual funds on the ESG performance of their holdings.

While this study directly and quantitatively compares overall portfolio-level, holdings-based weighted ESG scores between US-focused socially responsible (SR) and conventional mutual funds, existing literature addressed related, yet distinct, research questions. Durán-Santomil et al. (2019) examine the effects of SRI investments, measured by Morningstar sustainability scores, which include ESG and controversy dimensions, on the financial performance of European equity funds. They find that higher sustainability scores positively affect performance and emphasize that mutual funds tend to invest in firms with better ratings. Nitsche and Schröder (2018) compare average ESG rankings across European and global funds, concluding that SRI funds tend to exhibit higher ESG ratings than conventional counterparts. It is important to note that their calculations focus on top holdings only and do not assess full portfolio-level holdings-weighted ESG scores.

In summary, we therefore expect SR mutual funds to invest in companies with higher levels of social responsibility and, consequently, to exhibit higher overall ESG scores.

Hypothesis 1. SR mutual funds exhibit higher levels of weighted ESG scores than conventional mutual funds.

To date, no empirical study has directly and comprehensively compared the incidence of corporate scandals within the actual holdings of socially responsible (SR) and conventional mutual funds. Raghunandan and Rajgopal (2022) demonstrate that, compared to other funds managed by the same asset managers during the same time period, ESG funds are more likely to include stocks of firms that voluntarily disclose their carbon emissions performance. Paradoxically, these firms also tend to exhibit higher carbon emissions per unit of revenue. Moreover, the authors find that, on average, ESG funds select firms with poorer records on employee treatment and environmental practices compared to non-ESG funds. In line with the so-called *Janus phenomenon* (Dorfleitner et al., 2022), firms associated with high ESG scores even tend to be more likely to become involved in corporate scandals. Moreover, Dorfleitner et al. (2021) show, when focusing exclusively on SR funds, that SR funds have difficulty becoming leaders in both categories, ESG and corporate scandals.

All in all and consistent with our first hypothesis, we expect conventional mutual funds to exhibit lower levels of corporate scandals than SR mutual funds.

Hypothesis 2. Conventional mutual funds exhibit lower levels of corporate scandals in their holdings than SR mutual funds.

Table 1
Descriptive statistics.

Category	Obs.	Numb.of funds	Mean	Q1	Median	Q3	Std	Min	Max	Skewness
ESG score										
SR-funds	589	139	64.59	53.64	65.54	67.53	7.66	32.73	77.83	−0.98
Conventional funds	25 759	4898	59.84	53.64	61.61	67.53	10.21	20.44	86.96	−0.59
Controversies score										
SR-funds	589	139	71.85	59.42	70.95	84.60	13.43	33.45	100.00	0.05
Conventional funds	25 759	4898	71.36	59.42	69.89	84.60	16.59	2.29	100.00	−0.04
Combined score										
SR-funds	589	139	56.24	46.96	56.67	56.81	6.21	32.72	71.29	−0.36
Conventional funds	25 759	4898	51.62	46.96	52.15	56.81	7.31	20.44	86.66	−0.29
Absolute number of scandals										
SR-funds	589	139	5.08	1.11	4.15	6.09	4.32	0.00	23.07	1.21
Conventional funds	25 759	4898	4.42	1.12	3.30	6.09	4.53	0.00	44.39	2.04

This table presents the mean, first quartile, median, third quartile, standard deviation, minimum, maximum, and skewness values of the ESG, controversies, combined scores, and absolute number of scandals as well as the number of funds and the number of observations of the full dataset.

2. Data & methodology

This study analyzes mutual funds from the "CRSP Survivor-Bias-Free US Mutual Funds" database for the period 2013 to 2020. Following Borgers et al. (2015), index funds and ETFs are excluded.[3] ESG-related issues and corporate scandals are quantified using *ESG, ESG Controversies,* and *ESG Combined* from the Refinitiv Eikon database. The Controversies score assesses negative global media coverage of issues like tax fraud or human rights violations, while the Combined score integrates ESG and Controversies metrics. Scores are annual, ranging from 0 to 100, with higher values indicating better performance. A Controversies score of 100 denotes no involvement in scandals (Refinitiv, 2022).

To quantify the total score of a fund's holdings, we calculate a weighted sum of the latest available scores of the respective holdings prior this date and the individual security weighting (Wimmer, 2012; Dorfleitner et al., 2021): $Overall.Score = \sum_{i=1}^{N} x_i \cdot Score_i$.

Where, $Score_i$ is the score of asset i and $x_1, \ldots, x_N$ denotes the portfolio weights of asset i with $\sum_{i=1}^{N} x_i = 1$. To ensure reliable explanatory power, only funds with at least 70% score coverage are included in the dataset. In addition to scores, the *Absolute number of scandals* is used as an evaluation metric, representing the weighted sum of 23 scandal topics provided by Refinitiv, based on yearly raw counts of firm-specific negative news. Severe scandals, such as the Facebook-Cambridge Analytica or VW diesel scandals, can still be measured for years due to legal and reputational impacts. A fund is classified as sustainable if it is listed by US SIF (2022) or includes keywords like "green", "sustainable", "ESG", or "responsible" in its name, indicating a commitment to social responsibility. Table 1 shows the descriptive statistics of our full dataset.

3. Results

3.1. Conventional vs. SR funds

To test for differences between conventional and SR fund ratings, we use the Wilcoxon rank sum test (Wilcoxon, 1945). To address the unequal size of the fund categories, we follow the procedure of Benson et al. (2006) and create 100 subsamples from the conventional fund subcategory with the same size as the SR fund subcategory using the bootstrap procedure.[4] The results are presented in Table 2.

The mean or median of the conventional fund samples is the mean or median of the 100 independent subsamples. Thus, the statistical tests that use individual samples instead of the universe of conventional funds are not biased by the formation of smaller samples. The median Z-statistic, which is the median of Z-statistics of the 100 different subsamples, is significant at the 1% significance level for the ESG and Combined scores. The significant negative value indicates that the ESG and Combined scores of conventional funds are significantly lower than those of the SR funds. The number of significant Z-statistics measures the absolute frequency with which significant results can be examined from the 100 different Wilcoxon rank sum tests at the 5% significance level. For both ESG and Combined scores, all tests show a significant result. Moreover, the median effect size shows a medium effect for the ESG and Combined scores (Cohen, 1992). In sum, the results are in favor of Hypothesis 1, indicating that SR funds exhibit higher levels of weighted ESG scores than conventional funds. Thus, our results are in line with the findings of other authors, such as Kempf and Osthoff (2008) and Joliet and Titova (2018). Surprisingly, the results differ from Utz et al. (2014), who use a similar

[3] As a robustness test, we use the dataset, without excluding ETF and index funds, and re-run all calculations. The results remain materially unchanged.

[4] As a robustness check, we implement a permutation test (Strasser and Weber, 1999) by randomizing the subsamples and rerunning all calculations. The results remain consistent and are reported in the respective tables.

Table 2
Wilcoxon rank sum test and permutation test.

	ESG	Controv.	Combined	Numb. of scandals
Wilcoxon rank sum test				
Mean				
SR-funds	64.59	71.85	56.24	5.08
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.81	71.29	51.63	4.40
Median				
SR-funds	65.54	70.95	56.67	4.15
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.64	69.89	52.19	3.27
Median Z-statistics	−8.34***	−0.99	−11.01***	−3.63***
Numb. of significant Z-statistics (**)	100	16	100	100
Median effect size	0.34	0.04	0.45	0.15
Permutation test				
Z-statistics	−11.18***	−0.72	−15.12***	−3.48***

This table displays the mean and median of the scores of the SR funds, the conventional funds, and the 100 samples of conventional funds. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. The table displays the Z-statistics of the permutation test. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

methodology and find no significant differences between the two categories of funds. One reason for this may be that the number of SR funds in this study (27) is profoundly lower than the number of SR funds examined in our study (139), and the different time periods may also play a role. Another possible explanation is that socially responsible funds have become especially sensitive to care for ESG coverage due to various greenwashing scandals in recent years (e.g., JP Morgan, Barclays, Natixis, DWS).

Besides, the Controversies score results indicate insignificant differences between SR and conventional funds. Thus, contrary to ESG ratings, SR funds do not achieve a better Controversies rating than conventional funds. Following Dorfleitner et al. (2021), SR funds show difficulties in becoming leaders in ESG and controversies ratings. In consideration of this, one possible explanation may be that SR funds put special emphasis on achieving good ESG ratings and at least aim not to underperform in terms of controversies ratings compared to conventional funds. The results of the Absolute number of scandals show that SR funds exhibit a significantly higher number of aggregated corporate scandals than conventional funds. One aspect may be that the asset allocation of SR funds is mostly focused on selecting high ESG stocks ("best in class"), yet high ESG ratings alone may also indicate potential past or future corporate misconduct, i.e. penance/insurance mechanism (Dorfleitner et al., 2022).

The results presented in this chapter remain robust after controlling for industry sector allocation through an initial, straightforward approach. Therefore, the higher ESG and Combined scores, along with the simultaneously greater number of absolute scandals observed for SR funds relative to conventional funds, cannot be solely attributed to the funds' sectoral investment patterns.

Additionally, we complemented our analysis by classifying SR funds through additionally considering strategy types. The results for positive screening, negative screening, best-in-class, impact investing, and religious screening are presented in the Appendix (Tables 6–10). However, since all SR funds inherently apply ESG screening and responsible investment practices, these strategies are not considered separately. The results indicate that SR funds employing a particular investment style exhibit, apart from one single outlier, significantly higher ESG and Combined ratings compared to conventional funds. Furthermore, analyses of positive screening, best-in-class, and impact investing strategy subgroups reveal either a significant improvement in controversy ratings or a reduction in the number of scandals. Overall, we interpret these results as supporting our previous findings and conclude that the influence of investment strategy on our results is very limited.

3.2. Persistence analyses

To analyze the persistence of fund ratings, we divided the analyses into a short (one year), mid (1–3 years), and long-term (more than 3 years) horizon, inspired by Wimmer (2012) and Dorfleitner et al. (2021).[5] Table 11 shows that funds in the top and bottom quartiles are most likely to remain in their respective quartiles in the short term, with conventional funds exhibiting greater robustness than SR funds. Mid- and long-term analysis (Fig. 1) reveals consistent ratings for both SR and conventional funds across ESG, Controversies, and combined scores, without changes in initial quartile formation. However, the number of scandals in SR portfolios 2–4 tends to increase after 3–4 years (see (d)), while conventional funds show relatively stable scandal counts. Spearman correlation tests confirm significant consistency in rankings ($t+1, \dots, t+4$) over time for both fund types.

3.3. Active vs. passive SR funds

Passive SR funds were identified based on fund names (e.g., containing the terms "ETF" or "index") as well as fund characteristics from CRSP such as index affiliation and classification as ETFs.

[5] We have refrained from printing these tables due to space limitations. The results can be found in Appendix.

Table 3
Wilcoxon rank sum test and permutation test (active and passive funds).

	ESG	Controv.	Combined	Numb. of scandals
Wilcoxon rank sum test				
Mean				
Active funds	64.59	71.85	56.24	5.08
Passive funds	66.06	69.77	56.92	5.33
Subsamples of active funds	64.49	71.95	56.16	5.06
Median				
Active funds	65.54	70.95	56.67	4.15
Passive funds	67.88	68.91	57.05	3.28
Subsamples of active funds	65.43	70.89	56.63	4.14
Median Z-statistics	3.39	−1.99**	1.50	0.32
Numb. of significant Z-statistics (**)	0	68	0	0
Median effect size	0.14	0.08	0.06	0.02
Permutation test				
Z-statistics	2.60	−2.10**	1.51	0.78
Wilcoxon rank sum test				
Median Z-statistics (redirecting)	−3.20***	1.90	−1.40*	−0.40
Numb. of significant Z-statistics (**)	100	0	32	1
Median effect size (redirecting)	0.13	0.08	0.06	0.02
Permutation test				
Z-statistics	2.60***	−2.10	1.51*	0.78

This table displays the mean and median of the scores of the active funds, the passive funds, and the 100 samples of active funds. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. The table displays the Z-statistics of the permutation test. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 4
Wilcoxon rank sum test and permutation test (categories of numb. of scandals).

Scandals	Environment	Social	Governance
Wilcoxon rank sum test			
Mean			
SR-funds	0.16	3.11	1.89
Conventional funds	0.15	2.84	1.46
Subsamples of conventional funds	0.15	2.85	1.45
Median			
SR-funds	0.06	2.61	1.09
Conventional funds	0.07	2.16	0.89
Subsamples of conventional funds	0.07	2.17	0.88
Median Z-statistics	0.23	−3.32***	−5.44***
Numb. of significant Z-statistics (**)	0	100	100
Median effect size	0.02	0.11	0.18
Permutation test			
Z-statistics	−1.04	−2.05**	−5.04***

This table displays the mean and median of the scores of the SR funds, the conventional funds, and the 100 samples of conventional funds. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. The table displays the Z-statistics of the permutation test. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 3 reveals that active SR funds exhibit significantly higher weighted Controversies scores than passive SR funds. However, when assuming that passive SR funds have higher weighted scores (ESG, Combined, Controversies) and fewer scandals, calculations show that passive SR funds indeed have significantly higher ESG and Combined scores than active funds. No significant relationship is found for the absolute number of scandals. This may reflect SR fund managers prioritizing either ESG or Controversies scores, as suggested by Dorfleitner et al. (2021), and the potential negative correlation between these metrics.

3.4. Scandal and ESG subcategories

For a deeper insight, we divide the 23 scandal topics into the three pillar subcategories environmental, social, and governance-related scandals, mirroring the well-established ESG pillars.

Table 4 presents Wilcoxon rank sum test results (Wilcoxon, 1945), revealing that conventional mutual funds have significantly fewer social- and governance-related corporate scandals in their holdings compared to SR mutual funds, with all tests in these subcategories showing statistical significance. Interestingly, no significant differences are observed between SR and conventional funds regarding environmental-related scandals. This may be explained by the positive impact of strong environmental performance on future financial performance (Moneva and Ortas, 2010).

Table 5
Wilcoxon rank sum test and permutation test (ESG subcategories).

ESG	Environment	Social	Governance
Wilcoxon rank sum test			
Mean			
SR-funds	61.64	68.50	63.98
Conventional funds	55.63	63.17	61.37
Subsamples of conventional funds	55.74	63.16	61.39
Median			
SR-funds	62.27	69.68	64.31
Conventional funds	58.65	64.54	62.48
Subsamples of conventional funds	58.85	64.50	62.46
Median Z-statistics	−6.53***	−8.92***	−5.94***
Numb. of significant Z-statistics (**)	100	100	100
Median effect size	0.28	0.38	0.25
Permutation test			
Z-statistics	−9.44***	−12.02***	−8.75***

This table displays the mean and median of the ESG subscores of the SR funds, the conventional funds, and the 100 samples of conventional funds. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. The table displays the Z-statistics of the permutation test. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 5 shows the results when investigating ESG subcategories. The three highly significant negative values indicate that conventional mutual funds exhibit lower levels of environmental-, social-, and governance-subscores when compared to SR mutual funds. These findings show that SR funds outperform conventional funds in all three ESG categories and support Hypothesis 1.

4. Conclusion

This paper empirically examines the social responsibility and irresponsibility of socially responsible (SR) and conventional mutual funds through an analysis of historical holdings. Dimensions of corporate social irresponsibility is measured using both Refinitiv controversies scores and the absolute number of corporate scandals, while social responsibility is measured with ESG ratings. Results show that SR funds have higher weighted ESG scores than conventional funds, but conventional funds display fewer corporate scandals in their holdings–particularly in social and governance-related issues. No significant differences are found for the overall controversies score or for environmental-related scandals. These observations remain stable over time for both fund types. Passive SR funds exhibit higher ESG scores and lower Controversies scores than active SR funds, with no differences in the number of scandals. A potential limitation of this study lies in the observed timespan of our datasample. Future research may focus on changes in fund holdings of SR as well as conventional mutual funds after firm-specific corporate controversies.

All in all, this paper contributes to a deeper understanding of the relationship between corporate social responsibility and irresponsibility in the context of SR and conventional mutual funds. Against the background of increasing requirements and regulations for sustainable funds, our findings highlight the importance of considering corporate scandals and unethical behavior alongside traditional ESG metrics. Rather than interpreting the results as direct evidence of greenwashing, we emphasize that they point to a potential misalignment between ESG scores and realized corporate behavior. Future research could further explore the timing of investments, responses to controversies, and the role of different ESG evaluation frameworks in shaping these patterns.

Our findings reveal that SR funds tend to select companies with higher overall ESG scores but may still be exposed to firms involved in corporate scandals. For ethically motivated investors, these findings highlight the importance of scrutinizing not only ESG ratings but also corporate scandals in their investment decision process. For legislator and policymakers, this work underscores the need for enhanced disclosure requirements and stricter enforcement mechanisms to ensure that funds' SRI promises align with their actual portfolio holdings. Fund managers may improve their stock selection process by integrating assessments of corporate scandals, thereby increasing transparency and helping investors make investment choices consistent with their ethical principles.

CRediT authorship contribution statement

Jakob Artinger: Writing – original draft, Methodology, Investigation, Data curation. **Christian Kreuzer:** Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix

See Tables 6–11 and Fig. 1.

Table 6
Wilcoxon rank sum test (positive screening).

N = 171	ESG	Controv.	Combined	Numb. of scandals
Mean				
SR-funds (pos. screening)	63.47	74.66	56.10	4.71
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.95	71.48	51.58	4.42
Median				
SR-funds (pos. screening)	64.66	72.81	56.90	3.83
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.69	70.23	52.06	3.30
Median Z-statistics	−3.29***	−1.80**	−5.78***	−0.59
Numb. of significant Z-statistics (**)	98	58	100	5
Median effect size	0.25	0.14	0.44	0.05

This table displays the mean and median of the scores of SR funds applying a positive screening strategy, as well as conventional funds, and 100 samples of conventional funds. To ensure comparability, the number of conventional fund observations was matched to the SR fund sample size (N) using bootstrap resampling. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 7
Wilcoxon rank sum test (negative screening).

N = 238	ESG	Controv.	Combined	Numb. of scandals
Mean				
SR-funds (neg. screening)	62.70	72.53	54.83	4.42
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.77	71.46	51.68	4.43
Median				
SR-funds (neg. screening)	63.47	70.82	55.44	3.89
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.55	69.99	52.29	3.31
Median Z-statistics	−2.99***	−0.92	−4.64***	−1.19
Numb. of significant Z-statistics (**)	100	21	100	28
Median effect size	0.19	0.06	0.30	0.08

This table displays the mean and median of the scores of SR funds applying a negative screening strategy, as well as conventional funds, and 100 samples of conventional funds. To ensure comparability, the number of conventional fund observations was matched to the SR fund sample size (N) using bootstrap resampling. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 8
Wilcoxon rank sum test (best-in-class screening).

N = 60	ESG	Controv.	Combined	Numb. of scandals
Mean				
SR-funds (best-in-class)	68.33	66.58	57.85	7.46
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.90	71.47	51.70	4.45
Median				
SR-funds (best-in-class)	69.49	66.29	58.24	6.90
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.84	69.87	52.14	3.32
Median Z-statistics	−5.13***	1.57	−4.71***	−3.97***
Numb. of significant Z-statistics (**)	100	0	100	100
Median effect size	0.66	0.20	0.61	0.51

This table displays the mean and median of the scores of SR funds applying a best-in-class screening strategy, as well as conventional funds, and 100 samples of conventional funds. To ensure comparability, the number of conventional fund observations was matched to the SR fund sample size (N) using bootstrap resampling. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 9
Wilcoxon rank sum test (impact investing).

N = 40	ESG	Controv.	Combined	Numb. of scandals
Mean				
SR-funds (impact investing)	62.36	79.51	56.63	3.45
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.94	71.42	51.73	4.36
Median				
SR-funds (impact investing)	62.90	78.62	56.88	2.02
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.86	70.00	51.99	3.23
Median Z-statistics	−1.10	−2.13**	−3.09***	1.39
Numb. of significant Z-statistics (**)	19	77	99	0
Median effect size	0.17	0.34	0.49	0.22

This table displays the mean and median of the scores of SR funds applying an impact investing strategy, as well as conventional funds, and 100 samples of conventional funds. To ensure comparability, the number of conventional fund observations was matched to the SR fund sample size (N) using bootstrap resampling. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 10
Wilcoxon rank sum test (religious screening).

N = 34	ESG	Controv.	Combined	Numb. of scandals
Mean				
SR-funds (religious screening)	62.53	75.41	55.44	4.00
Conventional funds	59.84	71.36	51.62	4.42
Subsamples of conventional funds	59.93	71.47	51.44	4.37
Median				
SR-funds (religious screening)	67.79	72.35	57.81	3.24
Conventional funds	61.61	69.89	52.15	3.30
Subsamples of conventional funds	61.99	70.78	52.23	3.47
Median Z-statistics	−1.55*	−1.02	−2.40***	−0.02
Numb. of significant Z-statistics (**)	46	22	92	1
Median effect size	0.27	0.17	0.41	0.09

This table displays the mean and median of the scores of SR funds applying a religious screening strategy, as well as conventional funds, and 100 samples of conventional funds. To ensure comparability, the number of conventional fund observations was matched to the SR fund sample size (N) using bootstrap resampling. The median Z-statistic is the median of the 100 different sample Z-statistics. Significant Z-statistics are defined by the 5% significance level. The median effect size is the median of the effect sizes of the 100 rank sum tests. ***, **, and * indicate a significance level of 1%, 5%, and 10%, respectively.

Table 11
Contingency table of ESG, controversies, combined ratings and absolute number of scandals.

		ESG rating persistence (SR funds)						Controversies rating persistence (SR funds)			
		Rank in t_1						Rank in t_1			
		1	2	3	4			1	2	3	4
Rank in t_0	1	0.8411	0.1121	0.0374	0.0093	Rank in t_0	1	0.7217	0.2435	0.0261	0.0087
	2	0.1228	0.6754	0.2018	0.0000		2	0.2655	0.4690	0.2389	0.0265
	3	0.0261	0.2348	0.5826	0.1565		3	0.0360	0.2703	0.5766	0.1171
	4	0.0000	0.0091	0.1818	0.8091		4	0.0093	0.0000	0.1776	0.8131
		ESG rating persistence (Conv. funds)						Controversies rating persistence (Conv. funds)			
		Rank in t_1						Rank in t_1			
Rank in t_0	1	0.8575	0.1399	0.0018	0.0008	Rank in t_0	1	0.7837	0.1985	0.0153	0.0025
	2	0.0723	0.7568	0.1662	0.0047		2	0.2325	0.6107	0.1533	0.0035
	3	0.0023	0.1218	0.7403	0.1356		3	0.0235	0.2188	0.6857	0.0720
	4	0.0000	0.0031	0.1236	0.8733		4	0.0028	0.0072	0.1375	0.8525
		Combined rating persistence (SR funds)						Abs. numb. of scandals (SR funds)			
		Rank in t_1						Rank in t_1			
		1	2	3	4			1	2	3	4
Rank in t_0	1	0.7522	0.1947	0.0442	0.0088	Rank in t_0	1	0.8532	0.1468	0.0000	0.0000
	2	0.1727	0.6182	0.2000	0.0091		2	0.1316	0.6316	0.1842	0.0526
	3	0.0265	0.2124	0.5752	0.1858		3	0.0000	0.1250	0.6429	0.2321
	4	0.0091	0.0273	0.2091	0.7545		4	0.0000	0.0270	0.2072	0.7658
		Combined rating persistence (Conv. funds)						Abs. numb. of scandals (Conv. funds)			
		Rank in t_1						Rank in t_1			
Rank in t_0	1	0.7933	0.1940	0.0110	0.0018	Rank in t_0	1	0.8710	0.1232	0.0054	0.0004
	2	0.1331	0.6190	0.2289	0.0190		2	0.0574	0.7236	0.1985	0.0205
	3	0.0085	0.1967	0.6112	0.1835		3	0.0029	0.1284	0.6792	0.1895
	4	0.0074	0.0183	0.1775	0.7968		4	0.0008	0.0070	0.1513	0.8410

This table displays the contingency table of the initial (t_0) and the subsequent (t_1) quartile rank rating. Initially, the observed funds are ranked in one of the four quartile rank portfolios based on their ratings. These rankings are connected to the subsequent fund quartile ranking.

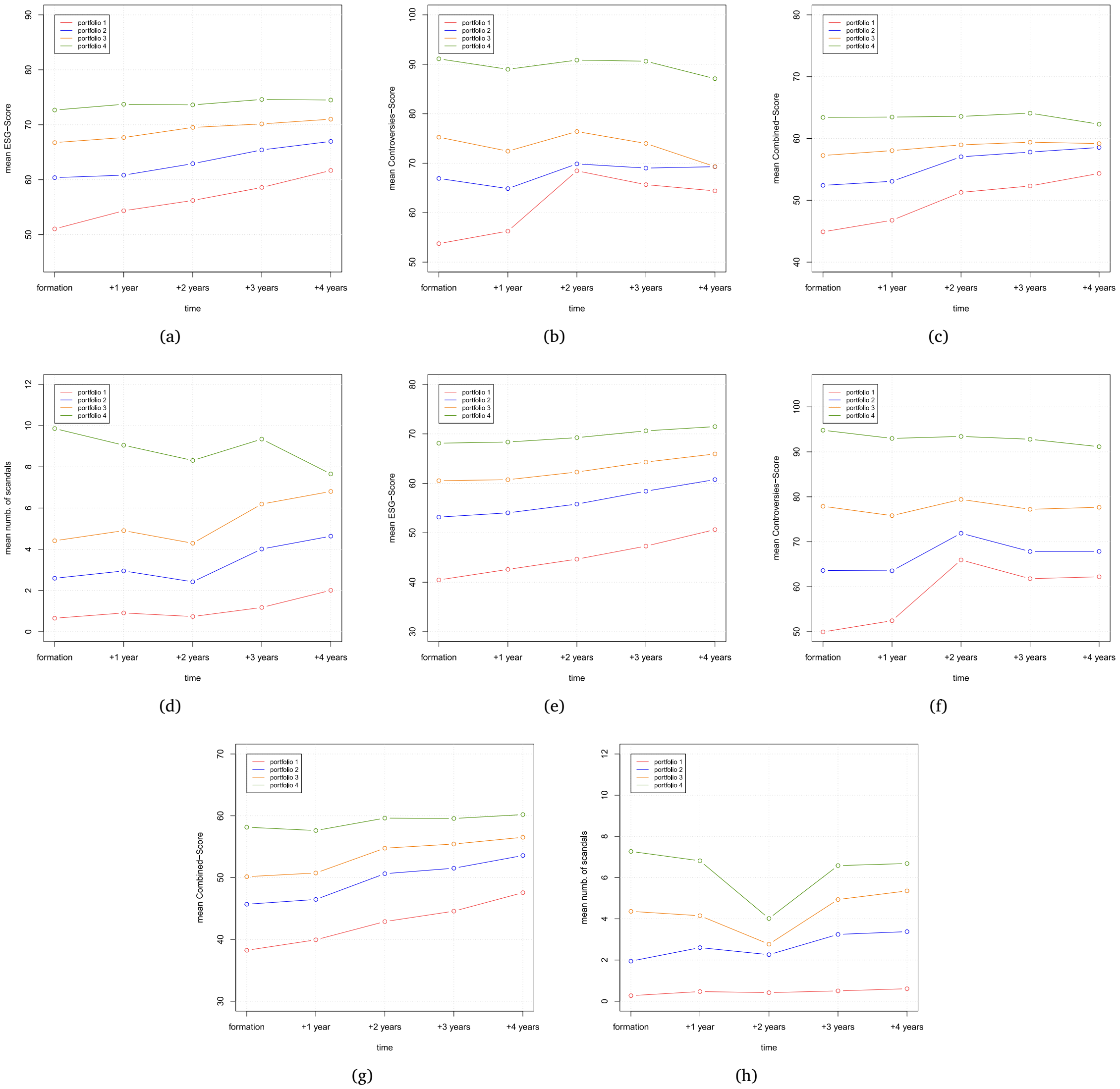

Fig. 1. We rank each SR fund in equal-weighted quartile SR portfolios based on their overall ESG (a), Controversies (b), and Combined scores (c) as well as absolute number of scandals (d). Similar, we rank each conventional fund in equal-weighted quartile portfolios based on their overall ESG (e), Controversies (f), and Combined scores (g) as well as absolute number of scandals (h). The lines represent the respective score of the four rank portfolios in their formation year (calculated by mean value) and the subsequent four years without any changes to their formation. Funds that initially achieve the highest ratings are contained in portfolio 4, and those with the lowest ratings are contained in portfolio 1.

Data availability

The authors do not have permission to share data.

References

Abouarab, R., Mishra, T., Wolfe, S., 2024. Spotting portfolio greenwashing in environmental funds. J. Bus. Ethics 1–29.

Aouadi, A., Marsat, S., 2018. Do ESG controversies matter for firm value? Evidence from international data. J. Bus. Ethics 151 (4), 1027–1047.

Arora, P., Dharwadkar, R., 2011. Corporate governance and corporate social responsibility (CSR): The moderating roles of attainment discrepancy and organization slack. Corp. Gov. Int. Rev. 19 (2), 136–152.

Benson, K.L., Brailsford, T.J., Humphrey, J.E., 2006. Do socially responsible fund managers really invest differently? J. Bus. Ethics 65 (4), 337–357. http://dx.doi.org/10.1007/s10551-006-0003-8.

Borgers, A., Derwall, J., Koedijk, K., Ter Horst, J., 2015. Do social factors influence investment behavior and performance? Evidence from mutual fund holdings. J. Bank. Finance 60, 112–126.

Cohen, J., 1992. A power primer. Psychol. Bull. 112 (1), 155–159.

Curtis, Q., Fisch, J., Robertson, A.Z., 2021. Do ESG mutual funds deliver on their promises? Mich. Law Rev. 120, 393.

Dorfleitner, G., Kreuzer, C., Laschinger, R., 2021. How socially irresponsible are socially responsible mutual funds? A persistence analysis. Finance Res. Lett. 43, 101990. http://dx.doi.org/10.1016/j.frl.2021.101990, URL: https://www.sciencedirect.com/science/article/pii/s1544612321000714.

Dorfleitner, G., Kreuzer, C., Sparrer, C., 2022. To sin in secret is no sin at all: On the linkage of policy, society, culture, and firm characteristics with corporate scandals. J. Econ. Behav. Organ. 202, 762–784.

Durán-Santomil, P., Otero-González, L., Correia-Domingues, R.H., Reboredo, J.C., 2019. Does sustainability score impact mutual fund performance? Sustainability 11 (10), 2972.

Eurosif, 2016. European SRI study 2016. URL: https://www.eurosif.org/wp-content/uploads/2022/03/Eurosif-SRI-study-2016.pdf.

Fisch, J.E., Robertson, A.Z., 2022. What's in a name? ESG mutual funds and the SEC's names rule. S. Cal. L. Rev. 96, 1417.

Gangi, F., Varrone, N., 2018. Screening activities by socially responsible funds: A matter of agency? J. Clean. Prod. 197, 842–855. http://dx.doi.org/10.1016/j.jclepro.2018.06.228.

Guidolin, M., Magnani, M., 2024. Do US active mutual funds make good of their ESG promises? Evidence from portfolio holdings. Risks 12 (2), 41.

Joliet, R., Titova, Y., 2018. Equity SRI funds vacillate between ethics and money: An analysis of the funds' stock holding decisions. J. Bank. Finance 97, 70–86. http://dx.doi.org/10.1016/j.jbankfin.2018.09.011.

Kaustia, M., Yu, W., 2021. Greenwashing in mutual funds. Available At SSRN 3934004.

Kempf, A., Osthoff, P., 2008. SRI funds: Nomen est Omen. J. Bus. Finance Account. 35 (9–10), 1276–1294. http://dx.doi.org/10.1111/j.1468-5957.2008.02107.x.

Kim, S., Yoon, A., 2023. Analyzing active fund managers' commitment to ESG: Evidence from the united nations principles for responsible investment. Manage. Sci. 69 (2), 741–758.

Moneva, J.M., Ortas, E., 2010. Corporate environmental and financial performance: A multivariate approach. Ind. Manage. Data Syst. 110 (2), 193–210.

Muñoz, F., 2021. Carbon-intensive industries in Socially Responsible mutual funds' portfolios. Int. Rev. Financial Anal. 75, 101740. http://dx.doi.org/10.1016/j.irfa.2021.101740.

Muñoz, F., Ortiz, C., Vicente, L., 2022. Ethical window dressing: SRI funds are as good as their word. Finance Res. Lett. 49, 103109. http://dx.doi.org/10.1016/j.frl.2022.103109.

Nitsche, C., Schröder, M., 2018. Are SRI funds conventional funds in disguise or do they live up to their name? In: Research Handbook of Investing in the Triple Bottom Line. Edward Elgar Publishing, pp. 414–446.

Peng, H., Zhang, Z., Goodell, J.W., Li, M., 2023. Socially responsible investing: Is it for real or just for show? Int. Rev. Financial Anal. 86, 102553.

Petridis, K., Kiosses, N., Tampakoudis, I., Ben Abdelaziz, F., 2023. Measuring the efficiency of mutual funds: Does ESG controversies score affect the mutual fund performance during the COVID-19 pandemic? Oper. Res. 23 (3), 54.

Raghunandan, A., Rajgopal, S., 2022. Do ESG funds make stakeholder-friendly investments? Rev. Account. Stud. 27 (3), 822–863.

Refinitiv, 2022. Environmental, social, and governance sores from refinitiv. URL: https://www.refinitiv.com/content/dam/marketing/en_us/documents/methodology/refinitiv-esg-scores-methodology.pdf.

Strasser, H., Weber, C., 1999. On the asymptotic theory of permutation statistics. Math. Methods Statist. 220–250.

US SIF, 2018. Report on US sustainable, responsible and impact investing trends.

US SIF, 2022. Sustainable investment mutual funds and ETFs chart. URL: https://charts.ussif.org/mfpc/.

Utz, S., Wimmer, M., Hirschberger, M., Steuer, R.E., 2014. Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. European J. Oper. Res. 234 (2), 491–498. http://dx.doi.org/10.1016/j.ejor.2013.07.024.

Wilcoxon, F., 1945. Individual comparisons by ranking methods. Biom. Bull. 1 (6), 80–83.

Wimmer, M., 2012. ESG-persistence in socially responsible mutual funds. J. Manag. Sustain. 3 (1), http://dx.doi.org/10.5539/jms.v3n1p9.