Journal of Financial Economics

Sustainability or performance? Ratings and fund managers' incentives

Nickolay Gantchev, Mariassunta Giannetti *, Rachel Li

Prof Mariassunta Giannetti, Stockholm School of Economics Department of Finance, Sweden

ARTICLE INFO

JEL Classification:
G11
G12
G23
G2

Keywords:
Sustainability
ESG
Mutual funds
Fund flows
Ratings

ABSTRACT

We explore how mutual fund managers and investors react when the tradeoff between a fund's sustainability and performance becomes salient. Following the introduction of Morningstar's sustainability ratings (the "globe" ratings), mutual funds increased their holdings of sustainable stocks to attract flows. Such sustainability-driven trades, however, underperformed, impairing the funds' overall performance. Consequently, a tradeoff between sustainability and performance emerged. In the new equilibrium, the globe ratings do not affect investor flows and funds no longer trade to improve their globe ratings.

Demand for sustainable investments has increased dramatically over the last two decades, and partially due to increased demand, sustainable investments have been performing well (Pastor et al., 2022). Hence, it is still a matter of contention whether investors select sustainable investments because of their nonpecuniary preferences for sustainability (Riedl and Smeets, 2017) or because they consider sustainability as a signal of future performance (Amel-Zadeh and Serafeim, 2017). As Starks (2023) highlights in her presidential address, it is also unclear how investors trade off sustainability and (financial) performance. Besides being important for understanding investors' preferences, evidence on how mutual fund investors approach the tradeoff between sustainability and performance would be useful for evaluating whether transparency about mutual funds' portfolios can increase the allocation of capital to sustainable investments.[1]

Morningstar's introduction of the globe ratings, which rank the sustainability of mutual funds' portfolios, offers a unique opportunity to

Nikolai Roussanov was the Editor for this article. Send correspondence to Mariassunta Giannetti (mariassunta.giannetti@hhs.se). Gantchev (nickolay.gantchev@wbs.ac.uk) is with the Warwick Business School at the University of Warwick, CEPR, and ECGI; Giannetti is with the Stockholm School of Economics, CEPR, and ECGI; Rachel Li (rachelli.research@gmail.com) is with the U.S. Securities and Exchange Commission (SEC). We thank the Editor, an anonymous referee, Darwin Choi, Diane Del Guercio, Chotibhak (Pab) Jotikasthira, Marcin Kacperczyk, Loriana Pelizzon, Anjana Rajamani, José Scheinkman, Patrick Verwijmeren, and Qifei Zhu for helpful discussions, and seminar participants at the MIT Financial Policy and Environment Conference, the 11th MSUFCU Conference on Financial Institutions and Investments, the FMA Consortium on Asset Management at the University of Cambridge, the Center for Economic Policy Research (CEPR) Endless Summer Conference, the Hanken School of Economics/Journal of Corporate Finance Conference on Ownership and Corporate Social and Sustainable Policies, the 2022 Hedge Fund Research Conference, the China International Conference in Finance, the Asian Bureau of Finance and Economic Research (ABFER) 8th Annual Conference, the 2023 HKU-TLV Finance Forum, American University, Baylor University, City University of Hong Kong, Durham University, ESSEC Business School, Hebrew University, Louisiana State University, Queen Mary University of London, the Securities and Exchange Commission, Stockholm School of Economics, University of Bristol, University of Cambridge, University of Exeter, University of Kansas, University of Liverpool, University of Oregon, University of Porto, University of Vienna, and University of Warwick. Giannetti acknowledges financial support from the Swedish House of Finance, the Karl-Adam Bonnier Foundation, the Nasdaq Nordic Foundation, and the Jan Wallander and Tom Hedelius Foundation. The Securities and Exchange Commission disclaims responsibility for any private publication or statement of any SEC employee or Commissioner. This article expresses the author's views and does not necessarily reflect those of the Commission, the Commissioners, or members of the staff. This paper was initially released prior to the author joining the Commission.

[1] With this objective, in 2021, the European Union introduced the Sustainable Finance Disclosure Regulation (SFDR), which pertains to all asset managers, regardless of whether they have an ESG or sustainability focus. The Securities and Exchange Commission is also ruling about disclosures to be made by investment funds that market themselves as sustainable (see https://www.ft.com/content/6fefdb2c-f72e-4e52-b95b-c0727aeb1a94).

Received 21 June 2023; Received in revised form 5 March 2024; Accepted 12 March 2024

Available online 27 March 2024

explore these critical issues. Morningstar rates mutual funds along a variety of dimensions and its ratings have been shown to affect flows (see, e.g., Del Guercio and Tkac, 2009; Ben-David et al., 2022; Heath et al., 2023). The sustainability ratings are no exception. In the aftermath of their introduction in March 2016, these easy-to-process and attention-grabbing signals significantly increased flows to the funds that received the highest sustainability ratings; in contrast, the funds with the lowest sustainability ratings experienced outflows (Hartzmark and Sussman, 2019).[2] Yet, these results cannot distinguish whether fund investors interpreted the globe ratings as a signal of future performance or whether they followed the ratings because of their preferences for sustainability.

We show that the aftermath of the globe ratings' introduction provides a laboratory to explore this important question. We document that a tradeoff between a fund's aspirations to achieve (maintain) a better globe rating and the fund's performance emerged because fund managers do not appear to be very skilled at trading the stocks that can improve their funds' globe ratings. As a result, sustainability (in terms of a better globe rating) became associated with bad performance, and the top globe ratings became unlikely to be perceived as a signal of superior future returns. In this context, we study whether investors continue to pursue funds with higher globe ratings and whether fund managers continue to tilt their portfolios towards high ESG stocks.

Fig. 1 illustrates the main result of our paper. Panel A shows that globe ratings' changes stopped affecting flows shortly after their introduction. When we distinguish between upgraded funds with bad and good performance in Panel B, it becomes evident that outflows from funds experiencing poor performance drive this finding.

In what follows, we document how the globe ratings affected fund managers' incentives and performance, and how this in turn led to outflows from the funds that succeeded in achieving (maintaining) a better globe rating. We show that ultimately the globe ratings became irrelevant, suggesting that investors initially cared about the globe ratings because they erroneously interpreted them as a signal of future performance.

We start by exploring how the globe ratings affected mutual funds' trading. Fund managers, whose compensation depends on assets under management (Geczy et al., 2021; Ibert et al., 2018; Ma et al., 2019), compete for flows. Their incentives to pursue different objectives depend on the relative weights that mutual fund investors in the aggregate put on performance versus sustainability. Naturally, observing that the globe ratings affected flows upon their introduction, managers with higher chances to achieve a better globe rating (or to be downgraded) should have changed their investment strategies to improve the sustainability rankings of their portfolios. Accordingly, we show that after the introduction of the globe ratings, mutual funds, whose current holdings placed them close to the cutoffs for the top and bottom ratings, increased (decreased) their investments in stocks with high (low) sustainability scores more than other funds.

However, we show that mutual funds that were striving to achieve better sustainability ratings experienced poor performance in the high sustainability stocks they purchased, but not in the remaining portions of their portfolios. In addition, these funds sold stocks with poor sustainability ratings that ended up performing well, creating profitable trading opportunities for other market participants. The poor performance experienced by funds that traded to a larger extent to increase their portfolios' sustainability is not explained by stock characteristics, such as value, size, or ESG rating, which may have been associated with negative shocks. Hence, these patterns are unlikely to be related to shifts in sustainability concerns, stemming from changes in the US administration, but rather are due to the fact that fund managers did not follow their strategies, skills, and information in their attempts to obtain better globe ratings. In particular, our findings are consistent with the idea that ESG information is complex and only few funds, even among those specialized in ESG, are able to incorporate it successfully in their investment process and generate alpha (Cremers et al., 2023).

Furthermore, we show that the funds that traded most to enhance their sustainability ratings experienced poor overall performance. As a result, the globe ratings stopped affecting flows, suggesting that investors favored performance over sustainability, and funds that were attempting to achieve better sustainability ratings ended up suffering net outflows. Unsurprisingly, experiencing costs in terms of performance and no benefits in terms of sustainability-driven flows, asset managers stopped tilting their portfolios to achieve better globe ratings.

Taken together, our results suggest that in the long term the globe ratings became ineffective because of the tradeoff between sustainability and performance and are in line with survey evidence that sustainability is viewed by some investors as positively predicting future performance (Amel-Zadeh and Serafeim, 2017) but are inconsistent with the idea that investors' nonpecuniary motives had a significant impact on flows (Hartzmark and Sussman, 2019). Our findings also indicate that fund managers, like the econometrician, were initially unable to discern why the globe ratings were associated with flows. When the globe ratings became irrelevant for flows, fund managers chose to pursue performance, which consistently leads to higher flows. Thus, our empirical investigation implies that given the preferences of US investors, transparency about asset managers' portfolios should not be presumed to increase flows to sustainable investments.

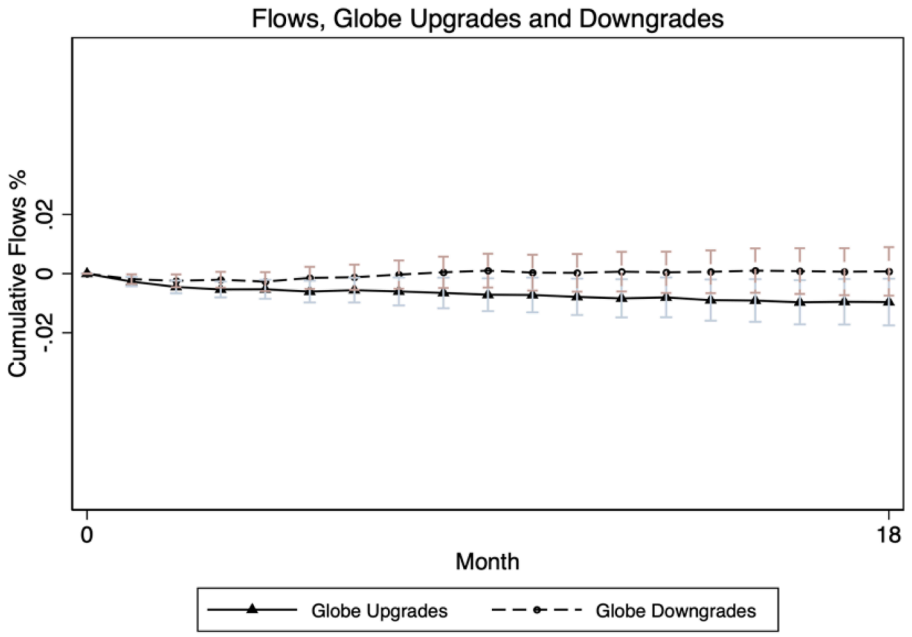
Different metrics to evaluate environmental and social performance are widely debated, and the globe ratings are no exception. However, our analysis shows that right after their introduction, the globe ratings affected fund flows and asset managers' portfolios, suggesting that market participants perceived the ratings as a valid sustainability indicator. Nevertheless, within less than a year after the ratings' introduction, fund flows stopped responding to globe rating upgrades and downgrades, despite the continued high interest in the ratings, as evidenced by Google Trends searches, and the high frequency of globe ratings' upgrades and downgrades. Morningstar's subsequent changes in the criteria for assigning the ratings, which should have attracted considerable investor attention, did not make the ratings more relevant for flows. Even for funds with an explicit sustainability focus as indicated in their prospectuses, the globe ratings do not affect flows after the initial period, suggesting that the average investor in these ESG-focused funds is unlikely to have genuine pro-social preferences.

This paper contributes to a growing literature that explores how sustainability affects investors' strategies and performance. Socially responsible investors are generally believed to put sustainability before performance (Riedl and Smeets, 2017; Barber et al., 2021; Bauer et al., 2021). Arguably for this reason, socially responsible mutual funds have been shown to have a lower flow-performance sensitivity (Bollen, 2007; Pastor and Vorsatz, 2020). However, there is no consensus on whether ESG investment is positively or negatively associated with performance, with a number of studies highlighting that sustainability improves performance and limits downside risk (see, e.g., Edmans, 2011; Lins et al., 2017; and Albuquerque et al., 2019), and others showing that these effects are only driven by temporary increases in demand (Pastor et al., 2022).[3] For these reasons, even ESG funds are believed to have conflicting objectives (Li et al., 2023). It is, therefore, important to examine a context in which the tradeoff between sustainability and performance becomes salient, as we do in this paper. We show that too few US mutual fund investors value sustainability over performance to generate any

[2] Ammann, Bauer, Fischer, and Müller (2018) and Ceccarelli, Ramelli, and Wagner (2024) also show that flows to funds with high sustainability ratings increase in the aftermath of the ratings' introduction.

[3] Confusion about the effects of ESG factors on financial performance is also frequently discussed in the press. See "ESG outperformance narrative 'is flawed', new research shows", Financial Times, May 3, 2021, available at https://www.ft.com/content/be140b1b-2249-4dd9-859c-3f8f12ce6036.

Panel A. Fund flows and globe upgrades/downgrades



Panel B. Fund flows, performance, and upgrades to globe 2 or 5

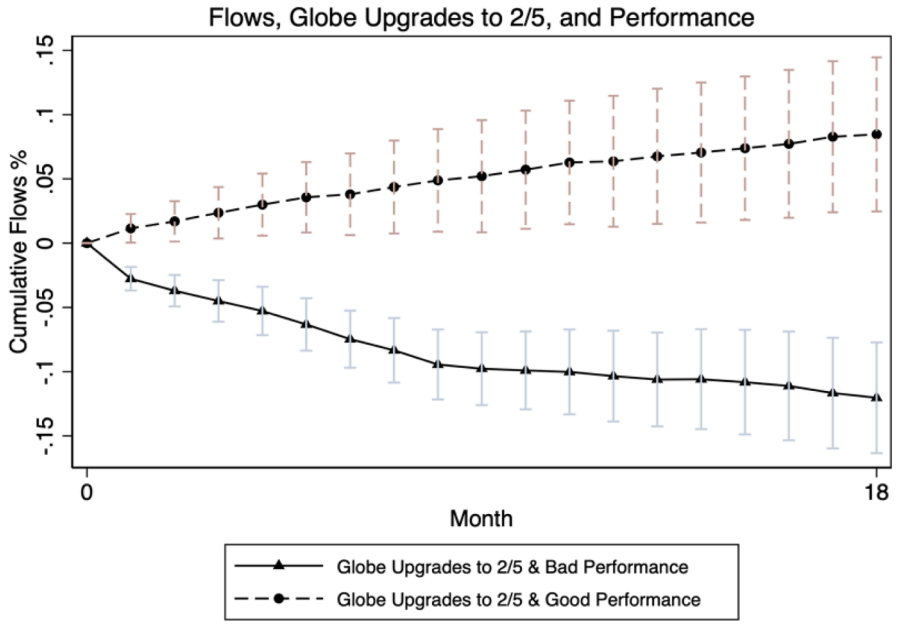


Fig. 1. Globe rating changes, fund flows, and performance
This figure compares the cumulative flows for funds that experienced a globe upgrade or downgrade (Panel A) and funds that were upgraded to globe 2 or globe 5 and had experienced good (bad) performance over the prior month (Panel B). We classify a fund as having good (bad) performance if the fund's performance between $t=-1$ and $t=0$ belongs to the top (bottom) quartile, compared to other funds during the same month. Fund flows are adjusted by the average fund flows within each Morningstar category during each month. The 90 % confidence intervals are also reported.

long-term effects of the globe ratings on the allocation of capital.

Another strand of the mutual fund literature studies how investor flows respond to attention-grabbing and easy-to-process signals, such as external rankings of the funds' performance (see, e.g., Del Guercio and Tkac, 2009; Evans and Sun, 2021; Ben-David et al., 2022; Kim, 2022; Reuter and Zitzewitz, 2021) or of the sustainability of the funds' portfolios (Hartzmark and Sussman, 2019; Ammann et al., 2018). Specifically, we build on the work of Hartzmark and Sussman (2019), who investigate the effects of the globe ratings on fund flows in a narrow time frame after the ratings' introduction, abstracting from general

equilibrium implications. We explore how asset managers respond to the ratings and how their response is driven by flows. In addition, while Hartzmark and Sussman (2019) conclude that both investors' expectations about the performance of funds with high sustainability ratings and nonpecuniary motives could explain the effects of the globe ratings on flows, our results imply that nonpecuniary motives did not play any role.

Prior work has shown that fund managers' pursuit of better star ratings affects stock demand and prices (Han et al., 2022; Kim, 2022). We are silent on whether the poor performance that fund managers experience when attempting to increase the sustainability of their portfolios arises because their behavior causes demand pressure or because they do not follow their strategies, skills, and information and execute poor trades. Regardless of the reasons driving poor performance, we highlight the tensions arising when funds are rated along two different dimensions that may create opposing incentives for fund managers. We show that in the long run, only ratings on the dimension that is followed by a larger proportion of investors matter.

1. Morningstar's sustainability ratings

The objective of Morningstar's globe ratings is to rank the sustainability of mutual funds' portfolios and to provide a way for investors to evaluate how a fund's investments meet environmental, social, and governance standards. The globe ratings and their methodology were publicly announced to mutual fund investors on March 1, 2016, when the sustainability ratings were first revealed. Since then, funds' globe ratings have been prominently displayed on Morningstar's website, along with the star ratings, which rank funds within a Morningstar category based on their performance over the previous three-, five-, and ten-year periods (if available). The globe ratings were and continue to be the subject of numerous press releases by Morningstar and are therefore widely covered by the media.[4] The sustained interest attracted by the globe ratings is evident from the time series of Google Trends searches for the term "globe rating", which as shown in Fig. 2, if anything, have increased in frequency since the ratings were first introduced.

A fund's globe rating is based on the fund's portfolio sustainability score, which is also available to Morningstar users, albeit less prominently displayed than the globe rating. A fund's portfolio sustainability score is computed as a weighted average of the ESG scores of the securities in the fund's portfolio, with the fund's portfolio shares as weights. The ESG scores of the securities are the ESG ratings of the issuers, obtained from Sustainalytics. Morningstar rates only funds that hold at least 50 % of their portfolios in securities with sustainability ratings.

A fund's globe rating is the percentile rank of its portfolio sustainability score relative to other mutual funds in the same Morningstar style category; thus, systematic differences in the ESG scores of the investment opportunities of funds with different specializations (e.g., growth vs. value) do not affect the initial version of the globe ratings we analyze in our main tests.[5] Only funds belonging to categories with at least ten funds are ranked.

Morningstar gives five globes and rates a fund as "High" sustainability if the fund is in the top 10 % of funds in its category. A fund is given four globes and rated as "Above Average" if it is ranked between 10 % and 32.5 %; it is given three globes and rated "Average" if it is ranked between 32.5 % and 67.5 %; and it is given two globes and rated "Below Average" if it is ranked between 67.5 % and 90 %. Finally, a fund is given one globe and rated "Low" sustainability if it is ranked in the bottom 10 % of its category.

The globe ratings exhibit a small positive correlation of 6.8 % with the star performance ratings, but as Table IA.1 in the Internet Appendix (IA) shows, star and globe ratings capture different fund characteristics with most globe five funds having star ratings below five.

Since the ESG scores of the securities typically change annually, the main determinant of the monthly changes in globe ratings is the fund's trading. Table 1 compares the frequency of globe rating upgrades and downgrades to that of the star ratings. Given that the star ratings depend on historical performance, it is unsurprising that the frequency of globe rating upgrades and downgrades is higher than that of the star ratings. A total of 277 (334) funds were upgraded (downgraded) to the top (bottom) rating in the first 18 months after the introduction of the globe ratings.

Based on the evidence presented in Table 1, changes in the globe ratings should have an effect on flows, just as star rating upgrades and downgrades do (see, e.g., Del Guercio and Tkac, 2009; Evans and Sun, 2021; Ben-David et al., 2022; Kim, 2022; Reuter and Zitzewitz, 2021). As we show, this does not seem to be the case in the data, even as the globe ratings continue to be frequently changed and prominently publicized. Our paper provides an explanation for why the globe ratings do not appear to affect flows in the long term.

2. Data and descriptive statistics

We obtain data on mutual funds' equity holdings from Morningstar and mutual funds' characteristics from Morningstar Direct. Our sample includes all US-domiciled funds, which invest in US equity and end up having globe ratings. This restriction ensures that we can explore changes in mutual funds' portfolios and performance in a relatively homogeneous sample. Since we focus on funds that invest in US equity, our sample is somewhat smaller than that in Hartzmark and Sussman (2019), who include all US-domiciled funds. Importantly, we confirm that flows increase (decrease) for funds with the top (bottom) globe rating in the aftermath of the ratings' introduction (Hartzmark and Sussman, 2019), indicating that before the tradeoff between sustainability and performance becomes apparent, mutual fund investors care about sustainability, and the globe ratings in particular.

As is common in the literature (Chevalier and Ellison, 1997), we include funds that have at least $10 million in assets under management. We also require funds to have information about their returns, age, expense ratio, TNA, and Morningstar category. Since in our tests we compare the effects of the sustainability and performance ratings on fund flows, we also require that funds have star ratings, which are assigned only to funds that are at least three years old.

Our main sample period ranges between July 2015 and September 2017 and includes 1959 unique funds. Among these, 1761 are active. Since most of our tests aim to capture the effects of funds' strategic behavior, we focus on active funds, unless noted otherwise.

We are unable to extend the sample before July 2015 because the availability of funds' portfolio sustainability scores is limited, preventing our analysis. However, we perform robustness tests on more recent periods (up to September 2020), which we introduce later in the paper.

The sample funds belong to the following Morningstar categories: US Fund Large Blend; US Fund Large Growth; US Fund Large Value; US Fund Mid-Cap Blend; US Fund Mid-Cap Growth; US Fund Mid-Cap Value; US Fund Small Blend; US Fund Small Growth; and US Fund Small Value.

Similar to Albuquerque et al. (2023), we also use fund prospectuses to identify funds with an explicit sustainability objective. We find 118 funds that mention words associated with social and environmental

[4] See, e.g., https://www.morningstar.co.uk/uk/news/227541/morningstar-globes-top-rated-sustainable-funds-in-2022.aspx.

[5] This feature of the globe ratings changed in a subsequent revision of the methodology. We show in Table 11 that this change does not affect our findings.

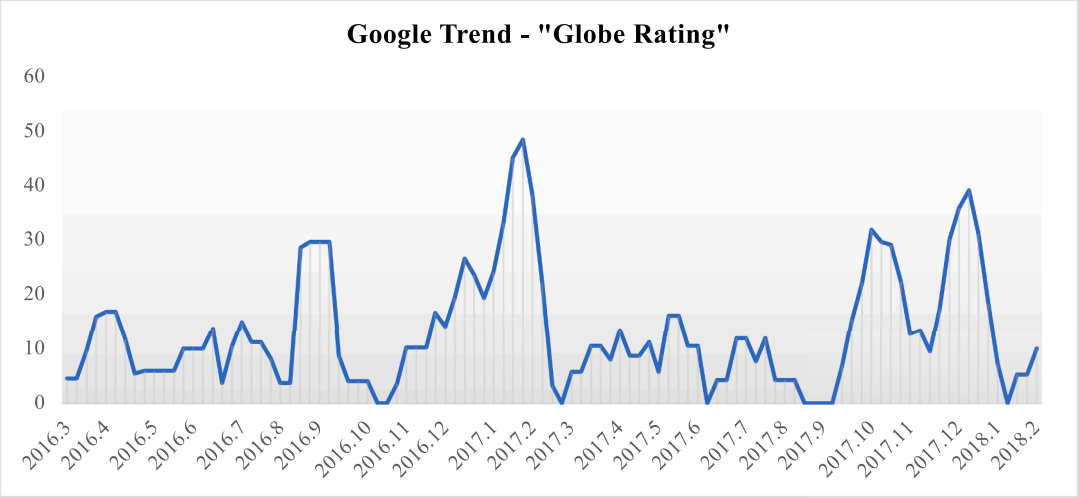


Fig. 2. Google Trends searches for "Globe rating"
This figure presents the search volume of the term "Globe rating" from Google Trends between March 2016 and February 2018. The monthly search volume is the four-week moving average of the weekly measure.

Table 1
Morningstar's star and globe rating upgrades and downgrades
This table shows the frequency of globe and star rating upgrades and downgrades in the first half of the sample period (from March to December 2016) and the second half of the sample period (from January to September 2017). Panel A includes all globe/star rating upgrades and downgrades, whereas Panel B focuses on upgrades from globe/star 1 to 2 and 4 to 5 and downgrades from globe/star 5 to 4 and 2 to 1 (i.e., changes from/to the bottom/top ratings).

	Globes		Stars	
	Upgrade	*Downgrade*	*Upgrade*	*Downgrade*
Panel A: all rating changes				
2016.3 - 2016.12	8.07 %	7.19 %	6.29 %	6.56 %
2017.1 - 2017.9	9.55 %	9.42 %	5.88 %	6.25 %
Panel B: change to/from top/bottom rating				
2016.3 - 2016.12	2.46 %	2.21 %	1.85 %	2.05 %
2017.1 - 2017.9	3.00 %	2.99 %	1.82 %	1.85 %

objectives in their prospectuses and we define these funds as having an explicit sustainability focus.[6] Thus, the vast majority of our sample consists of managers without a definite ESG focus. Interestingly, as shown in Fig. 3, most ESG funds have above average globe ratings, confirming that the globe ratings are informative.

Table 2 summarizes the main variables, distinguishing between the period before and the period after the introduction of the globe ratings. Detailed variable definitions are in the Appendix. For each fund, we aggregate fund size (TNA) and flows across share classes and calculate the fund's mean expense ratio and returns. On average, the sample funds have around $2500 million in assets under management and experience outflows equivalent to 0.4 % (0.6 %) of their TNA in the quarters preceding (following) the introduction of the globe ratings. Both characteristics indicate that the sample is very similar to that of Hartzmark and Sussman (2019). Other fund characteristics, such as expenses, equal to 1.1 % of TNA, are comparable to those in other studies of US mutual funds specialized in US equity (see, e.g., Han, Roussanov, and Ruan, 2022). The sample funds are around 18 years old, which is somewhat older than the average US-domiciled mutual fund investing in US equity because we restrict the sample to funds that have at least three years of historical performance by requiring the availability of star ratings.

Consistent with the globe rating definition, the median fund has a rating of 3, while the top (bottom) decile is 5 (1). As noted earlier, the globe ratings change more often than the widely studied star ratings, which rank funds based on their historical performance.

Panel C of Table 2 also summarizes stock characteristics, which we obtain from Compustat and CRSP, and the stocks' effective ESG scores, which are provided by Sustainalytics. We use this information to evaluate the performance of different portions of the mutual funds' portfolios and to explore how funds trade in stocks with different characteristics. In most empirical tests, we use monthly fund information because all funds report flows and performance at the monthly frequency, except in the tests exploring funds' trading in different types of stocks, where we use quarterly information because approximately 30 % of the funds report their positions only at the quarterly frequency.

3. The introduction of the sustainability ratings and funds' demand for high ESG stocks

We explore how the introduction of the sustainability ratings affects funds' trading behavior. Fund managers should have incentives to improve their funds' globe ratings if they expect better globe ratings to increase assets under management.

As shown for corporations that attempt to manipulate their credit ratings by changing their capital structure (Kisgen, 2006), these incentives should be particularly strong for funds close to the rating cutoffs because they are more likely than other funds to achieve a better rating, or equivalently, to avoid a downgrade. Thus, we expect such mutual funds, on average, to rebalance their portfolios towards stocks with high ESG ratings more than other funds.

To evaluate how the globe ratings affect fund managers' incentives, we construct a quarterly fund-stock-level panel and investigate the change in the position of fund f in stock i in quarter t, defined as:

[6] We define funds to have an explicit sustainability focus if their prospectuses include the following words/phrases: ESG, carbon, carbon neutral, clean energy, clean fuel, climate, climate impact, climate initiative, climate pledge, climate risk, CO2, conserve environment, CSR, data security, DEI, efficient energy, electric vehicle, emission, energy efficiency, energy reform, environmental, equality, fossil fuel, GHG, global warm, green, green business, green economy, green energy, greenhouse gas, less fossil, low carbon, mitigate carbon, new energy, Paris Accord, pollution, reduce carbon, reduce fossil, renewable, social impact, social issue, solar, SRI, stakeholder, sustainability impact, sustainability need, sustainability outcome, sustainability reference, sustainability report, wind energy, wind power, woman/women.

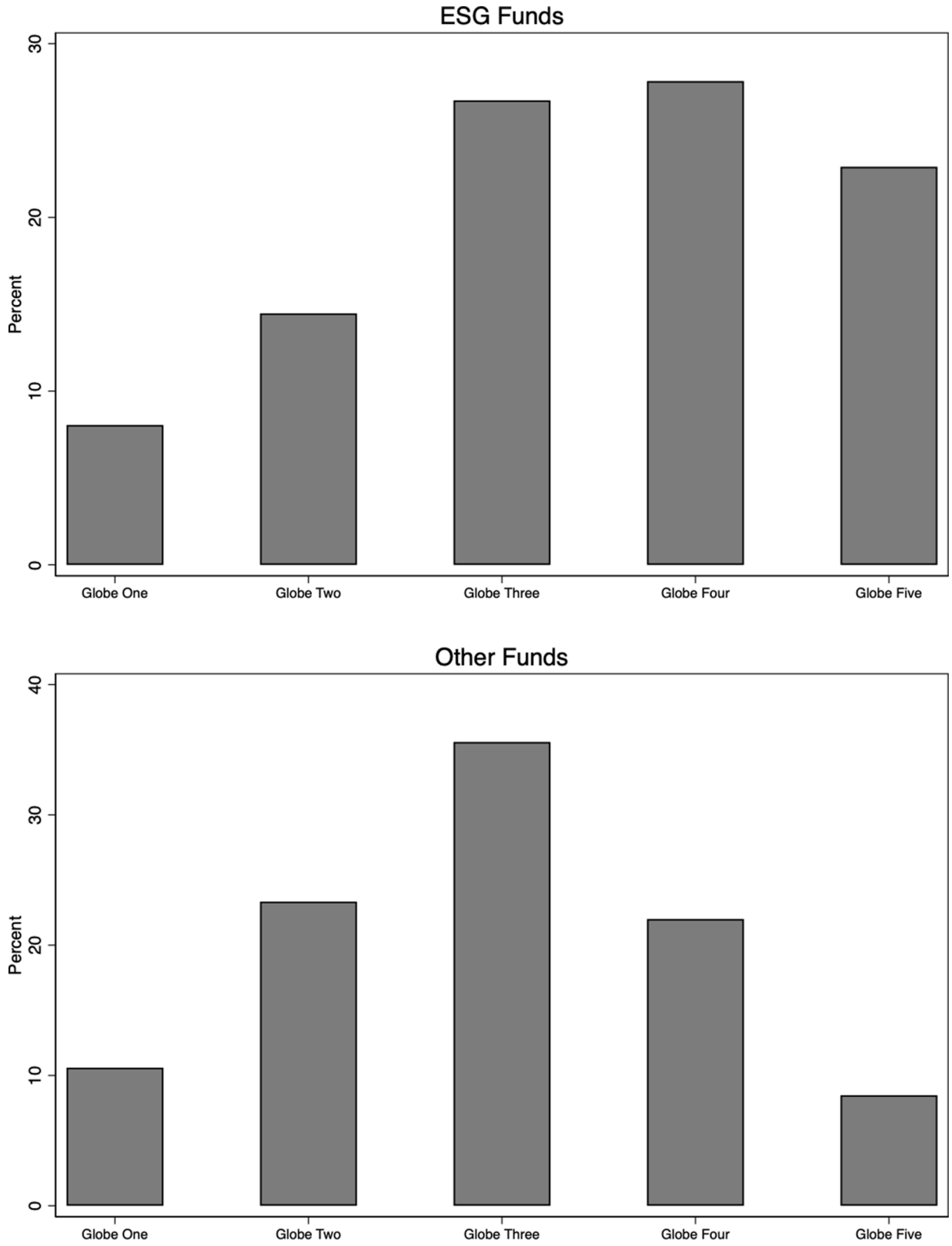


Fig. 3. Globe rating distribution for ESG-focused funds and other funds
This figure presents the distribution of globe ratings in the sample of ESG-focused funds and other funds. ESG funds are identified by searching for words associated with social and environmental objectives in the funds' prospectuses. Morningstar gives five globes and rates a fund as "High" sustainability if the fund is in the top 10 % of funds in its category. Similarly, Morningstar assigns four globes ("Above Average") if a fund is ranked between 10 % and 32.5 %; three globes ("Average") if a fund is ranked between 32.5 % and 67.5 %; and two globes ("Below Average") if a fund is ranked between 67.5 % and 90 %. A fund is given one globe and rated "Low" sustainability if it is ranked in the bottom 10 % of its category.

Table 2
Summary statistics
This table reports summary statistics of monthly mutual fund characteristics from July 2015 to February 2016 in Panel A (*Pre-globes*) and from March 2016 (when the globe ratings were first published) to September 2017 in Panel B (*Post-globes*) as well as quarterly stock characteristics from July 2015 to September 2017 in Panel C. The sample includes US-domiciled funds that invest in US equities and have at least $10 million in assets under management. All variables are defined in the Appendix.

	Num obs	Mean	Std dev	10th pctl	Median	90th pctl
Panel A: Fund (Monthly) – Pre-globes						
Flow (% TNA)	14,636	−0.004	0.035	−0.031	−0.004	0.023
TNA ($ million)	14,636	2219.14	5502.92	36.484	541.924	4956.70
Fund Age (Years)	14,636	18.155	12.038	5.75	16.25	29.917
Expense Ratio (%)	14,636	1.102	0.417	0.56	1.13	1.567
Star Rating	14,636	3.231	1.016	2	3	5
Fund Turnover (% TNA)	14,636	0.644	0.446	0.141	0.561	1.243
Position Change (Fund-Stock-Qtr)	426,240	−0.004	0.337	−0.189	0	0.171
Panel B: Fund (Monthly) – Post-globes						
Flow (% TNA)	29,556	−0.006	0.032	−0.03	−0.006	0.016
TNA ($ million)	29,556	2386.54	5799.00	38.467	579.495	5406.51
Fund Age (Years)	29,556	18.789	12.294	5.75	17	31
Expense Ratio (%)	29,556	1.077	0.416	0.543	1.106	1.52
Star Rating	29,556	3.214	1.014	2	3	4
Globe Rating	29,556	2.983	1.118	1	3	4
Fund Turnover (% TNA)	29,267	0.648	0.445	0.156	0.56	1.244
Abnormal ESG Trading	29,151	0.129	0.132	−0.02	0.119	0.303
Abnormal ESG Turnover	29,151	0.029	0.078	−0.038	0.013	0.114
Position Change (Fund-Stock-Qtr)	1,427,023	0.001	0.274	−0.106	0	0.101
Fund return	29,556	1.535	2.643	−1.367	1.242	4.938
FF4-Adj return	29,499	−0.159	1.211	−1.507	−0.143	1.187
DGTW-adj return	27,652	0.014	1.068	−1.204	−0.024	1.313
Buy high ESG return	27,652	0.007	3.056	−3.217	0	3.265
Sell low ESG return	27,652	0.022	2.313	−2.54	0	2.644
Buy other return	27,652	0.051	3.061	−3.164	0	3.394
Sell other return	27,652	0.003	2.095	−2.176	0	2.286
No-trade high ESG return	27,652	−0.043	1.959	−2.118	−0.041	2.042
No-trade low ESG return	27,652	0.051	2.144	−2.366	−0.018	2.583
Panel C: Stock (Qtrly) – Pre-/Post-globes						
Effective ESG Score	16,907	44.647	7.088	37.233	43.592	54.166
Ln Market Cap	36,349	6.867	1.797	4.314	6.78	9.294
Book to Market	36,317	0.531	0.423	0.093	0.449	1.095
ROA	35,434	0.012	0.047	−0.051	0.019	0.054
Ret	36,198	0.023	0.195	−0.213	0.017	0.265
Leverage	34,949	0.228	0.219	0	0.177	0.541
Sales Growth Rate	35,399	0.035	0.223	−0.161	0.019	0.224

$$Position\ Change(f,i,t) = \frac{Price(i,t-1)*[(NumShares(f,i,t) - NumShares(f,i,t-1)]}{TNA(f,t-1)}.$$

We normalize fund *f*'s change in the holdings of stock *i* by the fund's TNA at the beginning of the quarter and value the position using the beginning-of-quarter price of stock *i* (*Price*$(i,t-1)$).[7]

We consider funds whose portfolio sustainability scores in quarter *t-1* are within +/- 2.5 % from the top and bottom globe ratings as those with the strongest incentives to purchase (sell) stocks with high (low) sustainability scores. We label them *Border Funds*. This definition of border funds is not only consistent with theory (Bordalo et al., 2013) and evidence (Hartzmark, 2015) that ranking effects matter most for the best and the worst performers, but also takes into account that presumably managers of funds without an explicit sustainability focus care mostly about (not) being singled out for their very high (very poor) portfolio sustainability with a top (bottom) rating. In what follows, we test the plausibility of this assumption.

Mutual fund managers may have become aware of the globe ratings' planned introduction and methodology after August 2015, when Morningstar purchased (a large stake in) Sustainalytics, the company whose firm-level sustainability ratings are used to compute the fund portfolios' sustainability scores. Therefore, the investment policies of asset managers could have started to change during the second half of 2015, that is, before the official publication of the ratings.

In Table 3, we explore how funds' trading of stocks with high sustainability scores changes starting from the third quarter of 2015.[8] To investigate whether funds trade preemptively to improve their portfolio sustainability scores, we define a pre-globes period from the third quarter of 2015, when asset managers may have learned about the impending introduction of the globe ratings, to the first quarter of 2016. We also subdivide the post-globes period, following the official introduction of the globe ratings, into a first half – from the second quarter to the fourth quarter of 2016 – and a second half – from the first quarter to the third quarter of 2017.

Since the globe ratings were not yet available at the end of 2015 and during the first quarter of 2016, we use funds' portfolio sustainability

[7] As we show in Table IA.2, results are invariant if we use the end-of-quarter stock price to evaluate the change in position.

[8] Our sample starts in the third quarter of 2015 because the availability of funds' sustainability scores is limited before that time, which prevents the analysis.

Table 3

Mutual fund trading and stocks' ESG scores

Panel A of this table estimates the relation between funds' position changes (× 100) and stocks' Sustainalytics ESG scores. We consider active funds in columns 1–3 and index funds in columns 4–6. We define an indicator *Border Funds*, which equals one for funds with portfolio sustainability scores within +/−2.5 % of the globe rating cutoffs for globe 1 and globe 5, and zero otherwise. The sample period is from the third quarter (Q3) of 2015 to the third quarter (Q3) of 2017, divided into a pre-globes period (Q3 2015 – Q1 2016, columns 1 and 4) and a post-globes main period of two equal subperiods – first half (Q2 2016 – Q4 2016, columns 2 and 5) and second half (Q1 2017 – Q3 2017, columns 3 and 6). Panel B includes only active funds and reports estimates using indicators for *High ESG* (*Low ESG*) stocks, defined as those with ESG scores in the top (bottom) tercile of the Sustainalytics ESG scores. All specifications include lagged firm-level controls, including firm size, book-to-market ratio, leverage, ROA, sales growth rate, and quarterly stock return, and interactions of fund and year-quarter fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

Panel A. Border funds' trading and stocks' ESG scores

	(1)	(2)	(3)	(4)	(5)	(6)
	Position Change *(f,i,t)*					
	Active Funds			Index Funds		
	Pre-globes	Post-globes		Pre-globes	Post-globes	
Border Fund definition: Within 2.5 %, Globes 1/5	2015Q3–2016Q1	2016Q2–2016Q4	2017Q1–2017Q3	2015Q3–2016Q1	2016Q2–2016Q4	2017Q1–2017Q3
ESG Score	−0.014*	−0.005	0.014**	−0.011	−0.009*	−0.001
	(−1.946)	(−0.791)	(2.507)	(−1.245)	(−1.735)	(−0.198)
ESG Score × Border Funds	−0.004	0.049**	0.016	0.026	0.046	0.050
	(−0.123)	(2.003)	(0.846)	(0.560)	(1.277)	(0.965)
Observations	299,967	441,014	515,780	126,246	200,791	269,389
Adjusted R-squared	0.147	0.183	0.224	0.408	0.554	0.610
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ

Panel B. Indicators for high and low sustainability stocks

	(1)	(2)	(3)	(4)	(5)	(6)
	Position Change *(f,i,t)*					
	Active Funds					
	Pre-globes	Post-globes		Pre-globes	Post-globes	
Border Fund definition: Within 2.5 %, Globes 1/5	2015Q3–2016Q1	2016Q2–2016Q4	2017Q1–2017Q3	2015Q3–2016Q1	2016Q2–2016Q4	2017Q1–2017Q3
High ESG	−0.020	−0.001	−0.026			
	(−0.149)	(−0.012)	(−0.235)			
High ESG × Border Funds	−0.938	1.051**	0.513			
	(−1.264)	(2.572)	(1.424)			
Low ESG				−0.232*	0.102	−0.230**
				(−1.665)	(0.958)	(−2.351)
Low ESG × Border Funds				−0.511	−0.936**	−0.147
				(−0.783)	(−2.056)	(−0.459)
Observations	391,253	488,192	570,563	391,253	488,192	570,563
Adjusted R-squared	0.152	0.182	0.223	0.152	0.182	0.223
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ	Fund*YQ

scores to compute the cutoffs for the globe ratings that would eventually be introduced. Throughout the analysis, we control for various stock characteristics, including market capitalization, stock returns, book-to-market ratio, etc., which could be correlated with a stock's ESG score. We also include interactions of fund and quarter fixed effects, which capture the propensity of different funds to trade in a given quarter, including changes in the funds' assets under management.[9]

Panel A investigates funds' purchases of stocks with different Sustainalytics effective ESG scores. Column 1 shows that on average, active funds are not inclined to purchase stocks with high ESG scores, as captured by the negative and statistically significant coefficient on *ESG Score*. The funds that would eventually become *Border Funds* because of their portfolio sustainability scores are no different. Thus, there is no evidence that border funds tried to preemptively improve the sustainability of their portfolios. This is not entirely surprising: Engaging in a preemptive attempt to tilt the sustainability of fund portfolios in expectation of a higher globe rating (or to avoid being downgraded) requires considerable effort. Since the globes are based on a relative ranking, fund managers would need up-to-date information for all funds within the same category (a variable that is itself changing).

While on average fund managers avoided high ESG-rated stocks before the introduction of the globe ratings, possibly because they believed that their valuations were too high (Pastor et al., 2022), managers' incentives changed after March 2016, when they started observing that the globe ratings actually mattered for flows and they could use reported percentile rankings and information about their closest rivals within their category as a predictor of future rankings.

In the aftermath of the globe ratings' introduction, border funds engaged in trading to improve their portfolios' sustainability scores. The positive coefficient on the interaction term *ESG Score* × *Border Funds* in column 2 indicates that *Border Funds* rebalanced their portfolios towards stocks with high ESG scores. In terms of economic magnitudes, a one-standard-deviation increase in a stock's ESG score is associated with an increase in border active funds' positions in the stock of 24.4 % of the interquartile variation in our sample (calculated as 7.92×0.049/(0.52 - (- 1.07)). Notably, this behavior of *Border Funds* is observed only until the fourth quarter of 2016. As seen in column 3, the interaction term is not statistically significant in the second half of the post-globes period, indicating that these funds had on average the same trading behavior as other funds.

[9] Following the introduction of the globe ratings, border funds experienced net flows similar to those of other funds. As seen in Table IA.3, there are also no statistically significant differences in fund size and turnover in the first half of the post-globes period, but border funds appear to have marginally higher expense ratios and marginally lower performance ratings.

Since the globe ratings continue to be updated throughout the sample period, as shown in Table 1, it is unlikely that the lack of portfolio reallocation in the later part of the sample is due to the fact that funds had already achieved their desired sustainability ratings. Funds continue to be upgraded and downgraded, but the aspiration to achieve a better globe rating does not seem to affect their trading any longer. This evidence casts doubt on the presence of long-term effects of the globe ratings on fund managers' incentives. In the rest of the table, we scrutinize whether this finding is robust.

In the remaining columns of Panel A, we consider index funds, which we identify using the Morningstar flag. While active funds can strategically increase their holdings of stocks with high ESG scores, index funds must passively follow their benchmark indexes. Therefore, we should not observe that index funds whose portfolio ESG score is in a neighborhood of the cutoffs for the top and bottom globe ratings attempt to increase their holdings of stocks with high ESG scores. Consistent with this conjecture, in columns 4–6, we do not find any evidence that border index funds increase their holdings of stocks with high ESG scores. On average, index funds sell stocks with higher ESG ratings during the first three quarters after the introduction of the globe ratings (column 5). These findings support our interpretation that in the aftermath of the globe rating introduction, the trading behavior of active border funds is driven by strategic considerations.

Panel A considers as *Border Funds* only the funds within a narrow +/−2.5 % neighborhood of the cutoffs for the top and bottom ratings. These funds should have particularly strong incentives to trade to improve or maintain their globe ratings because being categorized as low (high) sustainability is expected to be particularly consequential for flows. Unable to stand out in terms of their funds' sustainability, fund managers whose portfolio sustainability scores are close to the cutoffs for the intermediate globe ratings are less likely to care about obtaining a higher or lower globe rating, especially because most sample funds do not have an explicit sustainability mandate.

To investigate this conjecture, in Table IA.4, we broaden the definition of border funds. As mentioned above, we continue to focus only on active funds. In columns 1–3, we define *Border Funds* as funds within +/−2.5 % from the cutoffs of all globe ratings. We expect this broader definition of border funds to include fund managers with weaker incentives to purchase stocks with high ESG scores. As expected, we do not find evidence that such funds trade to improve their globe ratings.

In columns 4–6 of Table IA.4, we extend the definition of *Border Funds* by considering a +/−5 % neighborhood around the cutoffs for the bottom and top ratings. As they are not as close to being upgraded/downgraded, these funds are less likely to be able to improve their portfolio sustainability scores relative to their peers. Therefore, we expect this broader definition of border funds to include fund managers with weaker incentives to purchase stocks with high ESG scores. The parameter estimates in column 5 are indeed smaller in magnitude, compared to column 2 of Panel A. Importantly, as in Panel A, it still appears that the aspiration to improve the fund's globe rating or to avoid a downgrade affects border fund behavior only up to three quarters after the ratings' introduction. Even though as shown in Table 1, the turnover in all globe ratings, and the bottom and top globe ratings in particular, continues to be high, we find no evidence of a differential effect in the trading of border funds in the last three quarters of the sample.

Since the managers of funds with sustainability scores close to the bottom and top ratings appear to have stronger incentives to improve the sustainability of their portfolios, in what follows, we consider as border funds only the funds whose portfolio sustainability scores are in a +/−2.5 % neighborhood of the cutoffs for the bottom and top ratings.

So far, we have explored how the trading of border funds varies depending on the stock's continuous sustainability score. Funds that attempt to achieve better globe ratings should not only purchase stocks with high sustainability scores but also sell stocks with low sustainability scores. To distinguish between stock purchases and sales, in Panel B of Table 3, we replace the continuous Sustainalytics *ESG Score* with indicators for *High/Low ESG* stocks, defined as those with ESG scores in the top/ bottom tercile of the Sustainalytics ESG scores. As seen in columns 2 and 5, managers of *Border Funds* purchase relatively more stocks with high ESG scores and sell more stocks with low ESG scores only in the first three quarters after the introduction of the globe ratings (up to the end of 2016). The effect is not only statistically, but also economically significant. For example, border funds reduce their positions in *Low ESG* stocks by 58.9 % of the interquartile variation during the first half of our sample period (−0.936/(0.52 - (−1.07)).

Interestingly, we observe that all active funds exhibit a tendency to purchase stocks with high ESG scores in the last part of the sample period (column 3 of Panel A and columns 3 and 6 of Panel B), when differences between border funds and other funds are no longer statistically significant. This tendency appears to be driven by the propensity to sell low ESG stocks (column 6 of Panel B). One possibility is that the sales of border funds may have driven down the returns of these stocks, and the funds that had purchased low ESG stocks on the cheap subsequently sell them after having realized the profits from their positions.

We also explore whether funds with an explicit ESG objective as disclosed in their prospectuses may have continued to trade to improve their ESG scores. Urging caution in interpretation due to the fact that we have a small sample of ESG funds, we show in Table IA.5 that ESG funds in general, and border ESG funds in particular, trade in a way that is not statistically different from other funds. In particular, we find no evidence that border ESG funds continued to trade to improve their ESG scores after the initial period.

Overall, the evidence described in this section shows that the introduction of the globe ratings initially influenced funds' portfolio allocations, but also raises the question why funds stopped pursuing better globe ratings only nine months after the ratings' publication. Since globe rating upgrades and downgrades continued to occur during the sample period, the lack of portfolio reallocation cannot be explained by the fact that funds had achieved their target rating. For this reason, to understand the tradeoffs managers face, in the next section, we explore the effects of ESG trading on fund performance.

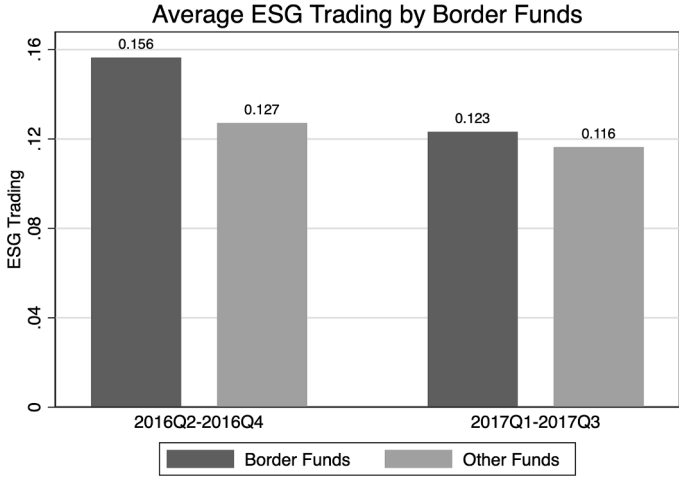


Fig. 4. Differences in ESG trading across funds
This figure compares the *Abnormal ESG Trading* (as defined in the Appendix) of border funds and other funds after the official publication of the globe ratings. We consider only active mutual funds and separately present the average ESG trading during March to December 2016, when border funds appear to have incentives to improve their globe ratings, and from January to September 2017, when border funds do not appear to trade in a way to improve their globe ratings. Border funds are funds with portfolio sustainability scores within +/−2.5 % of the rating cutoffs for globe 1 and globe 5.

4. Tradeoff between sustainability and performance

4.1. Abnormal ESG trading

In this subsection, we consider the consequences of the funds' trading strategies on their portfolios' composition and ratings. Our ultimate goal is to test whether funds that tilted more their portfolios to improve their ESG scores and achieve a better globe rating (or avoid being downgraded) experienced worse performance.

We conjecture that funds' performance suffers if managers deviate from the funds' usual trading strategies and do not rely on their information and skills to select high ESG stocks to purchase and low ESG stocks to sell. To evaluate this conjecture, we define a fund to have *Abnormal ESG Trading* if it purchased a large amount of stocks with high sustainability scores and/or sold a large amount of stocks with low sustainability scores, relative to its overall turnover and in comparison to the fund's trading in the period prior to the introduction of the globe ratings.

Specifically, we construct *Abnormal ESG Trading* as a fund-month variable that we relate to the fund's monthly performance:[10]

$$Abnormal\ ESG\ Trading(f,t) = ESG\ Trading(f,t) - \frac{1}{24} \times \sum_{\tau=March2016-36}^{March2016-12} ESG\ Trading(f,\tau)$$

$$\text{where } ESG\ Trading(f,t) = \frac{\sum_{j=1}^{gabs}(NumShares(f,j,t) - NumShares(f,j,t-1)) \times Price(j,t-1)}{\sum_{i=1}^{nabs}(NumShares(f,i,t) - NumShares(f,i,t-1)) \times Price(i,t-1)},$$

i is any stock held by fund f and

$$j \in \{High\ ESG\ stocks | NumShares(f,j,t) - NumShares(f,j,t-1) > 0\}$$
$$U\ \{Low\ ESG\ stocks | NumShares(f,j,t) - NumShares(f,j,t-1) < 0\}$$

That is, the numerator of $ESG\ Trading(f,t)$ captures fund f's purchases of high ESG stocks, valued using the stock price at *t-1*, plus the fund's sales of low ESG stocks, also valued using the stock price at *t-1*. High (Low) ESG stocks are defined as stocks in the top (bottom) tercile of the Sustainalytics effective ESG score. The denominator is the absolute value of the total trading of the fund (i.e., the change in the number of shares in *any* traded stock *i*, multiplied by the price of stock *i* at *t-1*). To capture deviations from the fund's usual trading strategy, we subtract the average ESG trading in the two years prior to the introduction of the globe ratings, excluding the 12 months closest to the introduction when the fund's behavior may have started to change.[11]

Consistent with the evidence in Table 3, mutual funds' ESG trading is larger in the first nine months after the introduction of the globe ratings (0.139 vs. 0.116, respectively). Importantly, Fig. 4 shows that in the first three quarters after the introduction of the globe ratings, the average abnormal ESG trading of border funds (as defined in Table 3, Panel A) is 0.156, compared to 0.127 for other funds. This difference is statistically significant, with a *t*-statistic of 7.30. In the second half, we do not see statistically different trading between the two groups; moreover, the abnormal ESG trading of all funds decreases.

We validate our ESG trading proxy using actual globe rating changes. Table 4 shows that funds that tilt their portfolios towards stocks with high sustainability ratings are more likely to experience a globe rating upgrade and less likely to experience a downgrade. All specifications in Table 4 and the following tables, in which we explore the effects of *Abnormal ESG Trading*, control for the funds' portfolio turnover as well as the turnover in ESG stocks, alleviating the concern that funds with abnormally high ESG trading simply trade more stocks with high ESG scores. Specifically, the variable *Abnormal ESG Turnover* controls for any trading in high and low ESG stocks, including sales of high ESG stocks and purchases of low ESG stocks that would result in a decrease of the funds' portfolio ESG scores.

The estimates confirm that our proxy captures the extent to which funds trade to improve their sustainability ratings. The effect is not only statistically, but also economically significant: An interquartile change in *Abnormal ESG Trading* (0.208 – 0.033 = 0.175) is associated with a 1.79 % (=0.102×0.175) higher probability of a globe rating upgrade, which is equivalent to a 20 % increase, compared to the average probability of a globe rating upgrade of 8.97 %. While this effect may appear small, it is important to consider that all funds have incentives to trade to improve their portfolio sustainability scores to be upgraded or avoid being downgraded. The actual outcome depends on factors that are not entirely under managerial control, such as stock prices affecting the portfolio shares and peer funds' actions. The mechanism resembles that of career concern models in which managers exert suboptimally high effort (Holmstrom, 1982), even though this has small effects on their reputation and compensation because all managers that they are competing with are also exerting suboptimally high effort.

Overall, these findings validate our interpretation that some funds tilt their portfolios towards stocks with high ESG scores to improve their globe ratings in the aftermath of the globe rating introduction. We can thus explore how pursuing a strategy that aims to improve a fund's sustainability rating affects the fund's performance.

4.2. ESG trading and fund performance

We test how a fund's performance depends on its abnormal ESG trading. We relate the abnormal ESG trading of all (active) funds in our sample to various measures of performance because any fund may have incentives to improve or maintain its portfolio sustainability score and globe rating, even though these incentives are particularly strong for border funds that are closest to the cutoffs and have a higher probability of succeeding in being upgraded or not being downgraded. In addition, not all border funds would have the same incentives, and hence, the dichotomous variable *Border Funds* is too noisy to demonstrate the

[10] Whenever possible, we use funds' monthly portfolio holdings, which are available for roughly two thirds of the funds in our sample. For the remaining funds, we use quarterly holdings to construct a quarterly ESG trading. In the regressions that rely on samples with monthly frequency, we input the (quarterly) Abnormal ESG Trading for all the months in a quarter. We proceed in the same way with the construction of Abnormal ESG Turnover. Our results are qualitatively invariant if we restrict the sample to funds that report monthly holdings (see Table IA.6 and Table IA.7).

[11] In this way, we make sure that we compare a fund's trading after the introduction of the globe ratings to that in a period before the globe ratings were published.

Table 4
Funds' ESG trading and globe rating upgrades and downgrades
This table studies the relation between the likelihood of an active fund experiencing a globe rating upgrade or downgrade and the fund's *Abnormal ESG Trading* (as defined in the Appendix). The sample period is from March 2016 to September 2017. In column 1 (column 2), the dependent variable is an indicator equal to one if the fund experiences an upgrade (downgrade) in its globe rating in month $t + 1$, and zero otherwise. All specifications include *Abnormal ESG Turnover* and lagged fund-level controls as well as interactions of Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

	(1) Globe Upgrade	(2) Globe Downgrade
Abnormal ESG Trading	0.102***	−0.132***
	(6.464)	(−8.642)
Abnormal ESG Turnover	0.090***	0.157***
	(3.178)	(5.987)
Fund Turnover (% TNA)	−0.003	0.016***
	(−0.690)	(3.136)
One Star	−0.004	0.008
	(−0.387)	(0.809)
Two Stars	0.004	0.006
	(0.767)	(1.053)
Four Stars	−0.004	−0.004
	(−0.763)	(−0.856)
Five Stars	0.010	−0.003
	(1.387)	(−0.449)
One Globe	0.067***	−0.102***
	(6.008)	(−22.722)
Two Globes	0.052***	−0.023***
	(6.547)	(−3.458)
Four Globes	−0.051***	0.051***
	(−8.450)	(6.799)
Five Globes	−0.113***	0.044***
	(−25.314)	(4.163)
Ln TNA	−0.003**	−0.003**
	(−2.007)	(−2.249)
Age	0.000	−0.004
	(0.115)	(−1.164)
Flow	0.062	0.101
	(0.851)	(1.537)
Expense Ratio	−0.000	0.003
	(−0.034)	(0.369)
Constant	0.155***	0.173***
	(5.044)	(5.734)
Observations	24,696	24,696
Adjusted R-squared	0.062	0.041
Fixed effects	Cat*YM	Cat*YM

mechanism we aim to study.[12] Thus, we recognize that *Abnormal ESG Trading* is more likely to capture that funds, and to a larger extent border funds, are not following their information and skills when trading to improve their portfolio ESG scores and provide evidence on how ESG trading affects fund performance.

Table 5 shows that the funds that attempt to improve their sustainability ratings suffer worse performance. In Panel A, we measure performance using the fund's portfolio monthly return in excess of the risk-free rate at $t + 1$ and control for fund characteristics, including the fund's past flows and TNA, both computed over the previous month, which capture any effects of changes in size on performance.[13] It appears that ESG trading, that is, abnormal purchases of high ESG stocks and sales of low ESG stocks relative to the fund's usual trading strategy, are negatively related to the fund's performance and that this negative effect emerges only in the first nine months after the introduction of the globe ratings, that is, when some funds actively tried to achieve better globe ratings.[14] Interestingly, in columns 1 and 2, *Abnormal ESG Turnover* is also negatively related to fund performance, indicating that the higher ESG turnover of mutual funds during the period immediately following the introduction of the globe ratings was largely driven by the funds' purchases of high ESG stocks and sales of low ESG stocks. Thus, while *Abnormal ESG Turnover* has a larger coefficient (its average and standard deviation are lower than those of *Abnormal ESG Trading*) and higher statistical significance than *Abnormal ESG Trading,* its sign indicates that in the first half of the sample it largely captures the same effect as *Abnormal ESG Trading*. However, the sign of *Abnormal ESG Turnover* changes in the second half of the sample, when funds stopped trading to improve their globe ratings and funds with ESG expertise could trade high and low ESG stocks without worrying about their portfolios' ESG scores. This suggests that funds that traded stocks with ESG ratings without specifically attempting to improve the funds' ESG scores experienced better performance in the second half of the sample.

Panel B provides more direct evidence on our conjecture that *Abnormal ESG Trading* is negatively associated with fund performance after the introduction of the globe ratings because fund managers had incentives to improve their portfolio sustainability scores without following their skills and information when picking stocks. As discussed before, these incentives should be disproportionately stronger for border funds that are closer to the cutoffs, and ceteris paribus, have a higher likelihood of being upgraded or not downgraded. We define the indicator *Border Funds* as equal to one for the funds with portfolio sustainability scores within +/−2.5 % from the cutoffs for the top and bottom ratings.

Consistent with our conjecture, funds with stronger incentives to tilt their portfolios towards stocks with high ESG scores and to deviate from their usual trading strategy appear to drive the negative effect of abnormal ESG trading on performance, as the direct effect of *Abnormal ESG Trading* is not statistically significant. However, for border funds, a one-standard-deviation increase in *Abnormal ESG Trading* results in a −2.37 % annualized return (=0.129×(−0.35–1.184)×12). Since for border funds *Abnormal ESG Trading* better captures the incentives to deviate from the normal strategy in order to pursue a higher portfolio sustainability score, this evidence indicates that not following a fund's information and skills is indeed costly for its performance.

While the results in Panels A and B are obtained including fund fixed effects, in Panel C, we explore the robustness of our findings to different measures of fund performance, which capture the funds' different exposures to systematic risk factors, include time fixed effects, and test for statistical differences in the effect of ESG trading in the first and the second part of the sample after the introduction of the globe ratings.

In column 1, we continue to use a fund's excess returns as a measure of performance. In column 2, we compute the fund's monthly abnormal return as the weighted average of the monthly abnormal returns of the fund's stockholdings at the beginning of the month. To control for the risk of different stocks, we use the risk-adjustment method proposed by Daniel et al. (1997), denoted as "DGTW". Specifically, we subtract the return of the characteristic-based benchmarks obtained by sorting stocks according to size quintiles, book-to-market quintiles, and prior return quintiles from the return of each individual stock. In column 3, we measure the fund's abnormal performance by its alpha, estimated from a Fama and French (1993) three-factor model, augmented by Carhart's (1997) momentum factor. In all specifications, we find that funds that engage in more ESG trading underperform other funds in the first part of

[12] Econometrically, the variable Border Funds is too weak as an instrument for ESG trading.

[13] As shown in Fig. 5, funds with high ESG trading shrink, indicating that these funds do not underperform because of negative scale effects (Berk and Green, 2004).

[14] We also do not find that ESG funds with high ESG trading perform better (Table IA.8). This is consistent with growing evidence that US mutual funds, even those that declare an ESG objective, engage in greenwashing (see, e.g., Kim and Yoon, 2023), and on average, are unable to successfully incorporate complex ESG information in their investment process to generate alpha (Cremers, Riley, and Zambrana, 2023).

Table 5
Funds' ESG trading and performance
This table explores the relation between an active fund's *Abnormal ESG Trading* and its performance. In Panel A, the dependent variable is the fund's monthly return in excess of the risk-free rate at $t+1$. In Panel B, we interact the fund's *Abnormal ESG Trading* with an indicator variable for *Border Funds*, which equals one for funds with portfolio sustainability scores within +/−2.5 % of the globe rating cutoffs for globe 1 and globe 5, and zero otherwise. In Panels A and B, column 1 reports estimates for the full sample period (from March 2016 to September 2017), column 2 studies the first half of the sample period (from March to December 2016), and column 3 focuses on the second half (from January to September 2017). In Panel C, the dependent variable is the fund's monthly return in excess of the risk-free rate at $t+1$ in column 1, the fund's DGTW risk-adjusted portfolio return (Daniel et al., 1997) at $t+1$ in column 2, and the fund's monthly alpha from a Fama-French-Carhart four-factor model estimated on a rolling window between month t-60 to t-1 in column 3. The indicator variable *First half* equals one if the sample period is between March and December 2016. All specifications include *Abnormal ESG Turnover* and lagged fund-level controls. Panels A and B include fund fixed effects, whereas Panel C includes fund and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

Panel A. Funds' ESG trading and excess returns

	(1)	(2)	(3)
	Fund Excess Return		
	Full Sample 2016.3–2017.9	*First half* 2016.3–2016.12	*Second half* 2017.1–2017.9
Abnormal ESG Trading	−0.186	−0.479*	0.151
	(−1.398)	(−1.936)	(0.946)
Abnormal ESG Turnover	−0.649***	−2.744***	2.740***
	(−2.739)	(−5.379)	(8.175)
Globe One	0.084	0.104	0.168
	(1.066)	(0.637)	(1.611)
Globe Five	−0.050	−0.116	−0.077
	(−0.603)	(−0.674)	(−0.782)
Fund Turnover (% TNA)	0.252**	−0.084	0.473***
	(2.105)	(−0.224)	(3.234)
Flow	−0.287	5.078***	−2.261***
	(−0.412)	(3.504)	(−3.119)
Ln TNA	−2.543***	−7.926***	−2.759***
	(−10.504)	(−8.848)	(−7.304)
Age	−4.446***	−15.973***	0.674
	(−11.672)	(−12.378)	(0.791)
Constant	64.414***	204.700***	54.149***
	(13.893)	(12.138)	(7.565)
Observations	26,273	12,628	13,625
R-squared	0.043	0.103	0.162
Fixed effects	Fund	Fund	Fund

Panel B. ESG trading of *Border Funds*

	(1)	(2)	(3)
	Fund Excess Return		
	Full Sample 2016.3–2017.9	*First half* 2016.3–2016.12	*Second half* 2017.1–2017.9
Abnormal ESG Trading	−0.099	−0.351	0.194
	(−0.702)	(−1.343)	(1.149)
Border funds	0.145*	0.413***	−0.014
	(1.779)	(2.600)	(−0.170)
Abnormal ESG Trading × Border Funds	−0.761**	−1.184*	−0.356
	(−2.042)	(−1.756)	(−0.869)
Abnormal ESG Turnover	−0.689***	−2.711***	2.645***
	(−2.752)	(−5.090)	(7.748)
Abnormal ESG Turnover × Border Funds	0.370	−0.333	0.948
	(0.564)	(−0.303)	(1.101)
Globe One	0.083	0.106	0.172*
	(1.063)	(0.652)	(1.653)
Globe Five	−0.051	−0.129	−0.080
	(−0.613)	(−0.756)	(−0.810)
Fund Turnover (% TNA)	0.253**	−0.071	0.478***
	(2.119)	(−0.189)	(3.268)
Flow	−0.279	5.092***	−2.266***
	(−0.400)	(3.516)	(−3.123)
Ln TNA	−2.545***	−7.931***	−2.757***
	(−10.526)	(−8.840)	(−7.288)
Age	−4.443***	−16.012***	0.684
	(−11.682)	(−12.376)	(0.803)
Constant	64.423***	204.852***	54.088***
	(13.913)	(12.128)	(7.542)
Observations	26,273	12,628	13,625
R-squared	0.043	0.103	0.162
Fixed effects	Fund	Fund	Fund

Panel C. Alternative performance measures

	(1)	(2)	(3)
	Fund Excess Return	DGTW-Adj Return	FF4-Alpha
Abnormal ESG Trading	1.415***	0.170**	0.268***

(*continued on next page*)

Table 5 (*continued*)

Panel C. Alternative performance measures	(1) Fund Excess Return	(2) DGTW-Adj Return	(3) FF4-Alpha
	(10.290)	(2.013)	(3.084)
Abnormal ESG Trading × First half	−2.015***	−0.373***	−0.331**
	(−10.060)	(−3.145)	(−2.498)
Abnormal ESG Turnover	−0.640**	0.065	−0.168
	(−2.404)	(0.386)	(−0.930)
Abnormal ESG Turnover × First half	−0.999***	0.211	0.559**
	(−2.757)	(0.863)	(2.301)
Globe One	0.048	0.051	0.039
	(0.955)	(1.514)	(0.997)
Globe Five	−0.008	−0.039	−0.065
	(−0.145)	(−1.110)	(−1.582)
Fund Turnover (% TNA)	0.157**	0.003	0.010
	(1.970)	(0.075)	(0.182)
Flow	−1.409***	−0.255	−0.812**
	(−3.298)	(−0.918)	(−2.553)
Ln TNA	−1.184***	−0.572***	−0.512***
	(−10.283)	(−8.702)	(−7.737)
Age	0.708*	0.550**	0.424
	(1.816)	(2.169)	(1.344)
Expense Ratio	−0.528*	−0.009	0.095
	(−1.879)	(−0.057)	(0.587)
Constant	23.688***	9.944***	8.766***
	(9.827)	(6.899)	(5.905)
Observations	26,273	24,924	26,216
R-squared	0.650	0.136	0.112
Fixed effects	Fund, YM	Fund, YM	Fund, YM

the sample. Specifically, an interquartile change in ESG trading is associated with a 1.26 % (=(−2.015 + 1.415)*0.175×12)) lower annualized excess return, a 0.43 % lower DGTW-adjusted return, and a 0.13 % lower Fama-French four-factor-adjusted return.[15]

These findings assuage concerns that the negative association between ESG trading and performance is due to the fact that the stocks with high (low) ESG ratings differ along other characteristics driving their performance. The results are also consistent with evidence in Hartzmark and Sussman (2019) that if anything, globe 5 funds underperformed globe 1 funds. However, we show that the differences in performance are associated with the funds' ESG trading, even if we control for their bottom and top globe ratings. Thus, our results provide an explanation for why the underperformance of top-rated funds may have emerged. Funds that strived to be upgraded (or not to be downgraded) experienced poor performance in trading stocks with high and low ESG scores. Put differently, the association between ESG trading and poor performance during the period in which fund managers appear to have attempted to achieve a better globe rating suggests that managers may not have performed much analysis for their ESG-driven trades or lacked expertise and information to select which high (low) ESG stocks to trade; instead, they may have just focused on the objective of obtaining a better globe rating.

It is possible, however, that managers with higher ESG trading have lower skills and underperform in all trades. Being unable to achieve superior performance, these managers could instead focus on sustainability. To identify the drivers of the performance of funds with high ESG trading, we investigate which subsets of stocks in a fund's portfolio drive the poor performance we observe in the first half of the sample after the introduction of the globe ratings. Specifically, if the underperformance is driven by the manager's trades aiming to improve the fund's portfolio sustainability score, we would expect the underperformance to arise primarily from trades of stocks with high and low ESG scores, rather than from stocks without ESG scores.

We thus partition each manager's portfolio into several sub-portfolios of stocks; that is, high ESG stocks purchased, low ESG stocks sold, other stocks purchased, other stocks sold, high ESG stocks with unchanged positions, and low ESG stocks with unchanged positions. We decompose a fund's performance by considering the average abnormal performance of the stocks in each of these sub-portfolios. To estimate a stock's abnormal performance and control for its risk exposure, we continue to use the risk-adjustment method proposed by Daniel et al. (1997).

Table 6 indicates that funds that do more ESG trading underperform because of the stocks with high ESG scores they buy and the stocks with low ESG scores they sell. In column 1, the dependent variable is the average abnormal return at $t + 1$ of the high ESG stocks that fund f purchased in month t. The negative and statistically significant coefficient on the interaction between *Abnormal ESG Trading* and *First Half* clearly shows that these high ESG stocks experience lower returns relative to their benchmarks. Specifically, an interquartile increase in ESG trading is associated with a 1.00 % lower annualized return from the high ESG stocks that funds purchase.

Similarly, in column 2, the dependent variable is the average performance of the stocks with low ESG scores that a fund sells. These low ESG stocks appear to subsequently outperform their benchmarks, as seen from the positive and statistically significant coefficient on the interaction between *Abnormal ESG Trading* and *First Half*. The effect is not only statistically significant, but also economically large: An average level of ESG trading is associated with an annualized loss from the sales of low ESG stocks of 0.19 %.

Thus, the performance of funds that intentionally attempt to improve their globe ratings suffers because they sell low ESG stocks that end up performing well and purchase high ESG stocks that subsequently perform poorly. As seen in columns 3 and 4, we do not observe similar patterns for the stocks with average sustainability ratings or without sustainability ratings that these funds trade. These trades are more likely to have been driven by the funds' information and usual trading strategies because these stocks have limited or no impact on the funds' portfolio sustainability scores, and consequently, on changes in the globe ratings. These findings suggest that the funds' underperformance

[15] We do not find any clear patterns for Abnormal ESG Turnover, which is negative and significant on average as well as during the first half of the sample in column 1, insignificant in column 2, and positive and significant in column 3. Table IA.9 shows that our results are invariant if we do not include Abnormal ESG Turnover as a control.

Table 6
Funds' ESG trading and performance in different sub-portfolios of stocks
This table explores the relation between an active fund's *Abnormal ESG Trading* and the performance of sub-portfolios of stocks based on the stocks' ESG ratings. The dependent variables are the fund's average DGTW risk-adjusted returns of different sub-portfolios of stocks. The indicator variable *First half* equals one if the sample period is between March and December 2016. *High ESG (Low ESG)* stocks are those with Sustainalytics ESG scores in the top (bottom) tercile; *Other* stocks are those with no ESG scores or stocks with ESG scores not in the top or bottom tercile. All specifications include *Abnormal ESG Turnover* and lagged fund-level controls as well as fund and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

	(1) Buy High ESG	(2) Sell Low ESG	(3) Buy Other	(4) Sell Other	(5) No-Trade High ESG	(6) No-Trade Low ESG
Abnormal ESG Trading	0.557**	−0.663***	0.051	0.297*	0.294**	−0.359**
	(2.356)	(−2.750)	(0.233)	(1.663)	(2.201)	(−2.129)
Abnormal ESG Trading × First half	−1.032***	0.753**	−0.035	−0.681***	−0.277	0.713***
	(−3.215)	(2.476)	(−0.132)	(−2.748)	(−1.378)	(3.289)
Abnormal ESG Turnover	−0.522	0.448	0.115	0.308	0.019	−0.214
	(−1.348)	(1.005)	(0.308)	(0.949)	(0.086)	(−0.656)
Abnormal ESG Turnover × First half	1.871***	0.374	−0.053	0.198	0.221	0.342
	(3.738)	(0.703)	(−0.109)	(0.509)	(0.793)	(0.904)
Fund Turnover (% TNA)	0.207*	0.222	−0.167*	−0.187*	0.137	0.042
	(1.674)	(1.585)	(−1.793)	(−1.846)	(1.604)	(0.455)
Flow	−0.210	0.054	−1.115*	−0.429	−0.658	−0.491
	(−0.259)	(0.065)	(−1.906)	(−0.739)	(−1.234)	(−0.877)
Ln TNA	−0.587***	−0.397***	−0.553***	−0.428***	−0.073	−0.297***
	(−3.737)	(−2.640)	(−4.809)	(−4.266)	(−0.747)	(−2.674)
Age	0.065	1.954**	0.710	0.102	−0.526	0.947*
	(0.089)	(2.340)	(1.199)	(0.165)	(−1.076)	(1.713)
Expense Ratio	1.283**	0.963*	−0.028	0.332	−0.382*	0.047
	(2.006)	(1.941)	(−0.083)	(1.122)	(−1.916)	(0.200)
Constant	9.936***	1.393	9.262***	8.021***	3.196	3.307
	(2.783)	(0.366)	(3.578)	(3.094)	(1.402)	(1.277)
Observations	24,924	24,924	24,924	24,924	24,924	24,924
R-squared	0.096	0.077	0.093	0.092	0.188	0.193
Fixed effects	Fund, YM	Fund, YM	Fund, YM	Fund, YM	Fund, YM	Fund, YM

is directly related to their ESG trades rather than driven by poor managerial skills.

Another possible concern is that stocks with high and low ESG scores are affected by unexpected shocks, such as the unanticipated outcome of the US presidential election in the last quarter of 2016, which could have driven the poor performance of the funds trading in these stocks. If this were the case, we would expect to observe that these funds underperformed also in the portfolio of stocks with high ESG scores that they held and for which they did not vary their positions during the month. In column 5, we do not find any evidence that a fund's ESG trading is associated with underperformance in the high ESG stocks in which the fund did not change positions, suggesting that underperformance in the portfolio of stocks with high ESG scores is due to bad trades.

In column 6, however, we find outperformance in the sub-portfolio of low ESG stocks that funds with high ESG trading hold, suggesting that the funds' performance suffers from excluding low ESG stocks, which are potentially subject to positive shocks during the sample period. Importantly, the sub-portfolio of other stocks that high-ESG-trading funds sell underperforms (column 4), indicating that fund managers exhibit skills in selecting which stocks to sell when they are not encumbered by ESG considerations.

Overall, this evidence suggests that funds' underperformance is driven by their ESG-related trading, that is, by purchasing stocks with high ESG scores at prices that are too high and selling stocks with low ESG scores that end up performing well during the period in which we observe particularly strong incentives for funds to improve their globe ratings. This may be the case because stocks with high (low) ESG scores become over- (under)-valued in the first few months after the introduction of the globe ratings due to the demand pressure created by the mutual funds pursuing better globe ratings. It is equally possible, however, that the managers of funds striving for better globe ratings did not use their information and usual investment strategies in their ESG-driven trades and consequently underperformed. In either case, the ESG-driven trades were bad trades.

In the next section, we show how the funds' underperformance and the relative importance of performance and sustainability ratings in attracting flows can explain why fund managers appear to have stopped trading to improve their globe ratings.

5. Consequences for fund flows

5.1. Dynamic effects of the globe ratings on flows

In this section, we consider fund flows and study how the apparent tradeoff between sustainability and performance we describe in Section 4 affected investors' and fund managers' incentives. Managers' compensation depends on the fees they earn, which in turn are driven by the funds' net assets under management (Chevalier and Ellison, 1997; Ibert et al., 2018; Ma et al., 2019). Thus, funds aim to maximize net flows. If some investors value sustainability over performance in their fund selection, there might exist an equilibrium in which some funds pursue better sustainability ratings, while other funds strive for better performance, even if the funds that achieve the top globe rating underperform. However, if investors do not trade off sustainability and performance but consider sustainability as a signal of superior future performance, sustainability signals that become associated with poor performance should stop affecting flows.

Table 7 explores to what extent this is the case focusing on active funds as in the earlier tests.[16] As is evident from columns 2 and 5 of Panels A and B in Table 7, during the first nine months after the globe ratings' introduction, funds with the top globe rating experienced higher inflows, while those with the bottom globe rating suffered outflows. Such a finding is revealed in Panel A, where we estimate specifications similar to those in Hartzmark and Sussman (2019), controlling for the funds' prior-month categorical star ratings. We confirm these results in

[16] In Table IA.10, we show that our results would be invariant if we considered all funds (including passive funds), as fund investors may not necessarily apply different selection criteria when they choose among passive funds.

Table 7
Effects of the globe ratings on fund flows
Panel A reports the effects of the globe ratings on monthly active funds' flows. Panel B performs a horse race between the star and globe ratings at *t-1* to evaluate their effects on fund flows. Columns 1 and 4 show results for the full sample period (from March 2016 to September 2017), columns 2 and 5 report results for the first half of the sample (March to December 2016), and columns 3 and 6 report results for the second half (January to September 2017). Columns 1–3 use globe 3 as the baseline, whereas columns 4–6 use the three middle globe ratings as the baseline. All specifications include lagged controls for the fund's returns, size, age, and expense ratio as well as interactions of the fund's Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

Panel A. Globe ratings and fund flows

	(1)	(2)	(3)	(4)	(5)	(6)
	Flow (%TNA)					
	Full Sample	*First half*	*Second half*	*Full Sample*	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
One Globe	−0.000	−0.003**	0.002	−0.001	−0.003***	0.001
	(−0.443)	(−2.119)	(1.167)	(−0.748)	(−2.674)	(1.206)
Two Globes	−0.000	−0.000	0.000			
	(−0.135)	(−0.425)	(0.146)			
Four Globes	0.001	0.002*	0.000			
	(1.286)	(1.900)	(0.116)			
Five Globes	0.003**	0.004***	0.002	0.003**	0.004**	0.001
	(2.430)	(2.649)	(1.119)	(2.429)	(2.559)	(1.170)
Star Rating	0.008***	0.008***	0.008***	0.008***	0.008***	0.008***
	(16.843)	(14.618)	(14.514)	(16.878)	(14.692)	(14.513)
Fund return	0.004***	0.005***	0.003***	0.003***	0.005***	0.003***
	(7.790)	(8.609)	(4.375)	(7.752)	(8.555)	(4.376)
Ln TNA	−0.002***	−0.002***	−0.001***	−0.002***	−0.002***	−0.001***
	(−6.256)	(−5.504)	(−4.950)	(−6.273)	(−5.597)	(−4.953)
Age	−0.003***	−0.002**	−0.003***	−0.003***	−0.002**	−0.003***
	(−4.278)	(−2.152)	(−4.764)	(−4.222)	(−2.019)	(−4.779)
Expense Ratio	−0.001	0.001	−0.001	−0.001	0.001	−0.001
	(−0.503)	(0.595)	(−1.065)	(−0.475)	(0.653)	(−1.068)
Constant	0.001	−0.004	0.004	0.001	−0.004	0.004
	(0.253)	(−0.681)	(0.732)	(0.285)	(−0.627)	(0.748)
Observations	25,108	11,212	13,896	25,108	11,212	13,896
Adjusted R-squared	0.096	0.111	0.085	0.096	0.111	0.085
Fixed effects	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM

Panel B. Star and globe ratings and fund flows

	(1)	(2)	(3)	(4)	(5)	(6)
	Flow (%TNA)					
	Full Sample	*First half*	*Second half*	*Full Sample*	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
One Globe	−0.001	−0.003**	0.001	−0.001	−0.004***	0.001
	(−0.694)	(−2.364)	(1.024)	(−0.985)	(−2.903)	(1.068)
Two Globes	−0.000	−0.000	0.000			
	(−0.269)	(−0.421)	(0.003)			
Four Globes	0.001	0.002*	0.000			
	(1.227)	(1.711)	(0.172)			
Five Globes	0.002**	0.004**	0.001	0.002*	0.003**	0.001
	(1.963)	(2.239)	(0.808)	(1.950)	(2.148)	(0.843)
One Star	−0.010***	−0.009***	−0.011***	−0.010***	−0.009***	−0.011***
	(−6.104)	(−4.943)	(−5.012)	(−6.126)	(−4.926)	(−5.029)
Two Stars	−0.006***	−0.007***	−0.006***	−0.006***	−0.007***	−0.006***
	(−6.929)	(−6.069)	(−5.392)	(−6.938)	(−6.052)	(−5.394)
Four Stars	0.008***	0.009***	0.008***	0.008***	0.009***	0.008***
	(11.246)	(8.694)	(9.431)	(11.262)	(8.728)	(9.428)
Five Stars	0.022***	0.024***	0.020***	0.022***	0.024***	0.020***
	(12.186)	(10.722)	(10.484)	(12.219)	(10.848)	(10.490)
Fund return	0.004***	0.005***	0.003***	0.004***	0.005***	0.003***
	(8.039)	(8.580)	(4.612)	(7.999)	(8.532)	(4.609)
Ln TNA	−0.002***	−0.002***	−0.001***	−0.002***	−0.002***	−0.001***
	(−6.563)	(−5.815)	(−5.174)	(−6.579)	(−5.906)	(−5.176)
Age	−0.002***	−0.001*	−0.003***	−0.002***	−0.001	−0.003***
	(−3.748)	(−1.676)	(−4.406)	(−3.690)	(−1.552)	(−4.409)
Expense Ratio	−0.001	−0.000	−0.002	−0.001	−0.000	−0.002
	(−1.169)	(−0.092)	(−1.544)	(−1.141)	(−0.045)	(−1.544)
Constant	0.025***	0.021***	0.027***	0.026***	0.022***	0.027***
	(4.732)	(3.290)	(4.444)	(4.787)	(3.400)	(4.462)
Observations	25,108	11,212	13,896	25,108	11,212	13,896
Adjusted R-squared	0.101	0.118	0.087	0.101	0.118	0.088
Fixed effects	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM

Panel B, where we instead include dichotomous variables for each of the lagged star ratings, using the middle globe/star ratings as the omitted variables. The estimates are economically significant: For instance, in column 2 of Panel B, achieving a globe rating of 5 in the first half of the sample period is associated with a 0.36 % increase in fund flows, which is equivalent to about 22.4 % of the interquartile variation in flows. Figure IA.1 further shows that in the aftermath of the ratings' introduction, the dynamics of flows to globe 1 and globe 5 funds are fully

consistent with Hartzmark and Sussman (2019).

However, performance is also important for fund flows, and as seen in Table 5, ESG trading is associated with worse performance. For example, lower performance attributable to an average amount of ESG trading is associated with a decrease of about 0.09 % in flows, which offsets around 25 % of the inflows from achieving a globe 5 status. Given that many funds attempt to improve their portfolio sustainability scores, high ESG trading only slightly increases the probability of achieving or maintaining a top globe rating. Thus, even a small drop in performance may translate into a net loss. In addition, *Border Funds* do more ESG trading than the average fund. Based on Panel A of Table 3, about 10 % of the funds that are closest to the cutoffs for the top and bottom ratings have strong incentives to trade to achieve a better globe rating or avoid being downgraded. According to our estimates in column 2 of Panel B, if a fund's ESG trading is in the top decile, the associated poor performance could lead to a 0.22 % decrease in flows, offsetting more than 59.7 % of the inflows from achieving a top globe status, which is an uncertain and very low-probability outcome.

Poor performance can lead to lower flows also through a fund's star rating. In this respect, a comparison of the coefficients on the globe and star ratings is also informative: the star ratings have larger effects on flows than the corresponding globe ratings even in the first half of the sample. Thus, poor performance, increasing the likelihood of a star rating downgrade in the future, can again lead to lower assets under management because collectively investors appear to care more about performance. Overall, it appears that even during the first half of the sample, when the globe ratings affected flows, the tradeoff between sustainability and performance may have been such that managers that care about assets under management had incentives to disregard the sustainability of their portfolios and the globe ratings.

Managerial incentives should have further weakened in the second half of the sample period because the globe ratings stopped affecting flows, as seen from the statistically insignificant coefficients on the globe rating dummies in columns 3 and 6 of Panels A and B. It is unlikely that the globe ratings lose power in attracting flows just because all investors that wanted to hold sustainable mutual funds quickly reallocated their portfolios in the immediate aftermath of the globe ratings' introduction. This could be the case if fund investors did not need to switch funds because the globe ratings are rarely changed once they are assigned. However, Table 1 shows that the globe ratings continued to exhibit high turnover throughout our sample period.

The insignificant effect of the globe ratings on flows suggests that investors put performance ahead of sustainability, and the globe ratings may initially have affected flows because they were interpreted as a sign of future performance. This conclusion is confirmed when we consider different specifications. Table 8 estimates the reaction of flows to globe rating upgrades and downgrades, controlling for the initial rating. We find no evidence that investors respond to upgrades and downgrades from/to the bottom/top globe rating. Only a fund's performance and its star rating changes appear to matter.

These findings indicate that flows stop responding to the globe ratings after their initial disclosure, arguably because investors gradually become aware of the tradeoff with performance, as Panel B of Fig. 1 suggests. In Fig. 5, we relate the globe rating changes and ESG trading to fund flows. Upgraded funds with low ESG trading, which were less likely to have underperformed, did not attract flows. Upgraded funds with high ESG trading, that is, the funds that were more likely to experience worse performance as a result of their trading of ESG stocks, experienced outflows. Overall, investors may have started associating globe rating upgrades with poor future performance (on average) and stopped considering them as a predictor of superior performance.

Table 9 provides more direct evidence that the initially coveted upgrade from the bottom rating or to the top rating failed to increase flows because of the poor performance of the managers that achieved an upgrade. To test this conjecture, we rank funds' returns each month into deciles, and define *Poor Performance* as an indicator variable that equals

Table 8
Effects of globe rating upgrades and downgrades on fund flows
This table reports the effects of star and globe rating upgrades and downgrades on monthly active funds' flows. Panel A considers upgrades/downgrades to/from all globes, whereas Panel B includes only upgrades/downgrades to/from the top/bottom globe ratings. Column 1 presents results for the full sample period (March 2016 to September 2017), column 2 reports results for the first half of the sample (March to December 2016), and column 3 reports results for the second half (January to September 2017). All specifications include lagged controls for the fund's returns, size, age, and expense ratio as well as interactions of the fund's Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

Panel A. Upgrades/downgrades to/from all globes

	(1)	(2)	(3)
	Flow (%TNA)		
	Full Sample	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
Globe Downgrade	−0.001	−0.000	−0.001
	(−0.947)	(−0.311)	(−1.052)
Globe Upgrade	−0.001	−0.001	−0.001
	(−1.259)	(−0.760)	(−1.032)
Star Downgrade	−0.004***	−0.004***	−0.004***
	(−4.622)	(−3.252)	(−3.540)
Star Upgrade	0.004***	0.006***	0.003**
	(5.189)	(5.040)	(2.536)
One Globe	−0.001	−0.003**	0.001
	(−0.698)	(−2.324)	(0.981)
Two Globes	−0.000	−0.000	0.000
	(−0.233)	(−0.400)	(0.038)
Four Globes	0.001	0.002	0.000
	(1.153)	(1.574)	(0.162)
Five Globes	0.002*	0.004**	0.001
	(1.886)	(2.209)	(0.737)
One Star	−0.010***	−0.010***	−0.011***
	(−6.450)	(−5.341)	(−5.202)
Two Stars	−0.006***	−0.007***	−0.006***
	(−7.245)	(−6.413)	(−5.572)
Four Stars	0.008***	0.009***	0.008***
	(11.597)	(9.095)	(9.635)
Five Stars	0.023***	0.025***	0.021***
	(12.365)	(10.952)	(10.618)
Fund return	0.004***	0.005***	0.003***
	(7.971)	(8.579)	(4.561)
Ln TNA	−0.002***	−0.002***	−0.002***
	(−6.789)	(−6.053)	(−5.327)
Age	−0.002***	−0.001	−0.003***
	(−3.680)	(−1.607)	(−4.372)
Expense Ratio	−0.001	0.000	−0.002
	(−1.056)	(0.043)	(−1.470)
Constant	0.027***	0.023***	0.028***
	(4.927)	(3.454)	(4.604)
Observations	25,108	11,212	13,896
Adjusted R-squared	0.103	0.121	0.089
Fixed effects	Cat*YM	Cat*YM	Cat*YM

Panel B. Upgrades/downgrades to/from the top/bottom ratings

	(1)	(2)	(3)
	Flow (%TNA)		
	Full Sample	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
Globe Downgrade to Globe 1/4	−0.001	−0.000	−0.002
	(−1.175)	(−0.221)	(−1.283)
Globe Upgrade to Globe 2/5	−0.001	−0.001	−0.001
	(−0.539)	(−0.395)	(−0.438)
Star Downgrade	−0.004***	−0.004***	−0.004***
	(−4.605)	(−3.254)	(−3.507)
Star Upgrade	0.004***	0.006***	0.003**
	(5.185)	(5.034)	(2.541)
One Globe	−0.001	−0.003**	0.001
	(−0.574)	(−2.195)	(1.061)
Two Globes	−0.000	−0.000	0.000
	(−0.135)	(−0.400)	(0.187)

(*continued on next page*)

Table 8 (*continued*)

Panel B. Upgrades/downgrades to/from the top/bottom ratings	(1)	(2)	(3)
	Flow (%TNA)		
	Full Sample	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
Four Globes	0.001	0.002*	0.000
	(1.241)	(1.660)	(0.214)
Five Globes	0.002**	0.004**	0.001
	(2.104)	(2.276)	(1.001)
One Star	−0.010***	−0.010***	−0.011***
	(−6.443)	(−5.337)	(−5.198)
Two Stars	−0.006***	−0.007***	−0.006***
	(−7.253)	(−6.420)	(−5.568)
Four Stars	0.008***	0.009***	0.008***
	(11.603)	(9.089)	(9.655)
Five Stars	0.023***	0.025***	0.021***
	(12.365)	(10.947)	(10.627)
Fund return	0.004***	0.005***	0.003***
	(7.986)	(8.597)	(4.565)
Ln TNA	−0.002***	−0.002***	−0.002***
	(−6.789)	(−6.038)	(−5.328)
Age	−0.002***	−0.001	−0.003***
	(−3.671)	(−1.607)	(−4.355)
Expense Ratio	−0.001	0.000	−0.002
	(−1.055)	(0.038)	(−1.468)
Constant	0.026***	0.022***	0.028***
	(4.900)	(3.431)	(4.572)
Observations	25,108	11,212	13,896
Adjusted R-squared	0.103	0.121	0.089
Fixed effects	Cat*YM	Cat*YM	Cat*YM

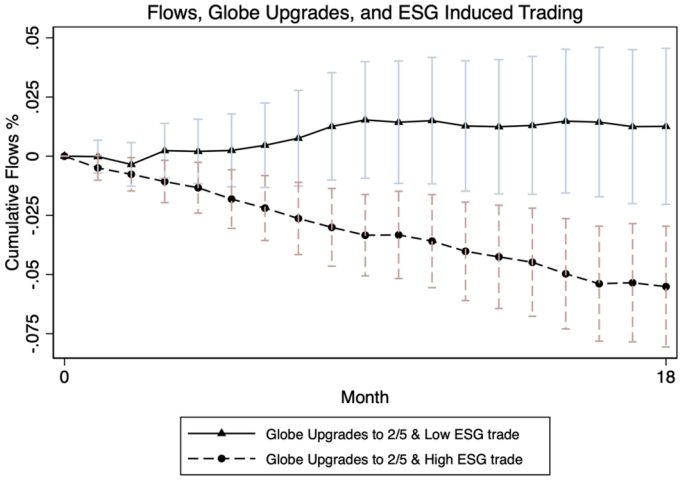


Fig. 5. Fund flows, ESG trading, and globe upgrades
This figure compares the cumulative flows for active mutual funds that were upgraded to globe 2 or globe 5 at $t = 0$. Specifically, we separate active funds into two groups based on the extent to which they have engaged in *Abnormal ESG Trading* (as defined in the Appendix) between $t=-1$ and $t = 0$. We classify a fund as *High (Low) ESG trade* if the fund's *Abnormal ESG Trading* belongs to the top (bottom) quartile, compared to other funds within the same Morningstar category during the same month. Fund flows are adjusted by the average fund flows within each Morningstar category during each month. The 90 % confidence intervals are also reported.

one if a fund's monthly return belongs to the bottom decile. We also introduce an interaction between the dummy for a fund's poor performance and its upgrade from globe 1 or to globe 5, respectively. The estimates show that even funds that managed to be upgraded, which as shown in Table 4 was an uncertain event due to competition with other funds, did not attract flows. While the direct effect of an upgrade is positive but not statistically significant in the first part of the sample, upgraded funds lost assets under management when their performance was remarkably poor in the nine months after the introduction of the globe ratings (column 2). Interestingly, an upgrade from globe 1 to globe 2 does not magnify the negative effect of poor performance (columns 7–9), but we find that this negative effect on flows is larger for funds that are upgraded to the top rating, as the coefficient on the interaction term between *Poor Performance* and *Upgrade to Globe 5* is negative and statistically significant (column 5). The effect is also economically significant – an upgraded fund with a record of poor performance experiences an additional 1.7 % outflows (equivalent to 38.9 % of the standard deviation of fund flows). This suggests that some investors in the upgraded funds redeemed, possibly fearing that a change in strategy towards sustainability would have resulted in persistently poor performance.

This evidence provides an explanation for why fund managers stopped trading to improve their portfolio ESG scores. Realizing that globe rating upgrades and downgrades did not matter for flows, and that high ESG trading 'backfired' because of the negative effects on performance, fund managers stopped tilting their portfolios towards stocks with higher sustainability scores.

Overall, the findings we have presented so far suggest that in the long term, the globe ratings are unlikely to lead to an increase in financial flows to sustainable investments. Nevertheless, it could be that a top globe rating insulates funds from redemptions following weak performance (Bollen, 2007). In turn, this could give underperforming asset managers incentives to invest in sustainable stocks. In Table IA.11 in the Internet Appendix, we show that a top globe rating does not mitigate the negative effects of weak performance. The interactions between bottom and top globe ratings and fund performance are not statistically significant.[17]

In sum, the globe ratings appear to leave flows unaffected in the second half of the sample period. These findings are confirmed in Table IA.13, where we distinguish between funds' institutional and retail share classes. While immediately after the introduction of the globe ratings, institutional investors allocate capital to funds with the top globe rating (column 2) and retail investors also redeem capital from funds with the bottom globe rating (column 5), the sustainability ratings lose power in explaining the flows for both categories of investors in the second half of the sample.

The evidence that mutual fund investors pay close attention to performance and the star rating upgrades and downgrades further suggests that the poor performance of the funds that achieved the highest sustainability rating may have led investors to subsequently ignore the globe ratings. This effect is likely to have been stronger for institutional share classes as more sophisticated institutional investors realized that a top globe rating was not a costless marketing tool, but instead came at the expense of performance.

It is also possible that some or all investors that value sustainability over performance are inattentive and do not track changes in the globe ratings. However, even if mutual fund investors were inattentive to the globe rating changes, we would still conclude that increased transparency about the sustainability of funds' portfolios does not provide long-term incentives for fund managers to tilt their portfolios towards sustainable investments. Furthermore, the insignificant interaction term between the globe ratings and fund performance in Table IA.11 does not support such an interpretation.

5.2. Do globe ratings still matter for ESG-focused funds?

We also consider funds that we identity as having an explicit sustainability focus based on their prospectuses as those that are more likely to have investors that value sustainability. We then test whether the top and bottom globe ratings continued to be relevant for these ESG-focused funds, which are upgraded and downgraded as frequently as

[17] In addition, Table IA.12 shows that globe 5 funds do not attract flows even if they have a top star rating.

Table 9

Globe rating upgrades, fund performance, and flows

This table studies the effects of the interaction between negative performance and globe rating upgrades on active funds' flows. Each month, we rank funds' returns into deciles and define *Poor Performance* as an indicator variable that equals one if a fund's monthly return belongs to the bottom decile. The dependent variable is a fund's monthly flow. Columns 1, 4, and 7 show results for the full sample period (March 2016 to September 2017), columns 2, 5, and 8 report results for the first half (March to December 2016), and columns 3, 6, and 9 report results for the second half (January to September 2017). All specifications include lagged controls for the fund's categorical star rating, returns, size, age, and expense ratio as well as interactions of the fund's Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1% level is denoted by *, **, and ***, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Flow (%TNA)								
	Full Sample	*First half*	*Second half*	*Full Sample*	*First half*	*Second half*	*Full Sample*	*First half*	*Second half*
	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9	2016.3–2017.9	2016.3–2016.12	2017.1–2017.9
Poor Performance	−0.004***	−0.005***	−0.003***	−0.004***	−0.005***	−0.003***	−0.004***	−0.006***	−0.003***
	(−5.218)	(−5.106)	(−2.684)	(−5.307)	(−5.208)	(−2.712)	(−5.463)	(−5.456)	(−2.787)
Upgrade to Globe 2/5	−0.000	−0.001	0.001						
	(−0.088)	(−0.737)	(0.330)						
Poor Performance ×	−0.007**	−0.011**	−0.003						
Upgrade to Globe 2/5	(−2.099)	(−2.210)	(−0.714)						
Upgrade to Globe 5				−0.001	−0.003	−0.000			
				(−0.584)	(−1.053)	(−0.087)			
Poor Performance ×				−0.014**	−0.017*	−0.010			
Upgrade to Globe 5				(−2.241)	(−1.938)	(−1.144)			
Upgrade to Globe 2							0.001	0.000	0.001
							(0.419)	(0.047)	(0.518)
Poor Performance ×							−0.002	−0.005	0.001
Upgrade to Globe 2							(−0.503)	(−1.150)	(0.169)
Observations	25,108	11,212	13,896	25,108	11,212	13,896	25,108	11,212	13,896
Adjusted R-squared	0.096	0.111	0.084	0.096	0.111	0.085	0.096	0.110	0.084
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed effects	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM

Table 10
Effects of the globe ratings on ESG funds' flows
This table reports the effects of the globe ratings on monthly active ESG funds' flows. *ESG Funds* are defined by searching words associated with social and environmental objectives in the funds' prospectuses. Column 1 shows results for the full sample period (March 2016 to September 2017), column 2 reports results for the first half of the sample (March to December 2016), and column 3 reports results for the second half (January to September 2017). All columns use the three middle globe ratings as the baseline. All specifications include lagged controls for the fund's returns, size, age, and expense ratio as well as interactions of the fund's Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

	(1) *Full Sample* 2016.3–2017.9	(2) *First half* 2016.3–2016.12	(3) *Second half* 2017.1–2017.9
One Globe	−0.001	−0.004***	0.001
	(−1.018)	(−2.909)	(1.022)
Five Globes	0.002*	0.003*	0.001
	(1.742)	(1.743)	(0.917)
ESG Funds	0.001	0.001	0.001
	(0.440)	(0.409)	(0.408)
One Globe × ESG Funds	0.002	0.003	0.000
	(0.619)	(0.852)	(0.020)
Five Globes × ESG Funds	0.001	0.004	−0.002
	(0.180)	(0.644)	(−0.430)
One Star	−0.010***	−0.009***	−0.011***
	(−6.139)	(−4.910)	(−5.042)
Two Stars	−0.006***	−0.007***	−0.006***
	(−6.935)	(−6.056)	(−5.396)
Four Stars	0.008***	0.009***	0.008***
	(11.238)	(8.714)	(9.373)
Five Stars	0.022***	0.024***	0.020***
	(12.188)	(10.870)	(10.463)
Fund return	0.004***	0.005***	0.003***
	(7.990)	(8.525)	(4.622)
Ln TNA	−0.002***	−0.002***	−0.001***
	(−6.548)	(−5.869)	(−5.168)
Age	−0.002***	−0.001	−0.003***
	(−3.651)	(−1.485)	(−4.403)
Expense Ratio	−0.001	−0.000	−0.002
	(−1.178)	(−0.079)	(−1.570)
Constant	0.025***	0.022***	0.027***
	(4.746)	(3.352)	(4.449)
Observations	25,108	11,212	13,896
Adjusted R-squared	0.101	0.118	0.087
Fixed effects	Cat*YM	Cat*YM	Cat*YM

other funds during the sample period.

Table 10 estimates the specifications in columns 4 to 6 in Panel A of Table 7 adding an interaction between the top (bottom) globe rating and an indicator variable for ESG funds. The interaction terms are never statistically significant, while the top (bottom) globe rating appears to be associated with positive (negative) flows only during the first part of the sample.

Interpreting these results with caution due to the fact that our sample includes only 118 funds with an explicit sustainability focus, we conclude that investors in ESG funds are similar to investors in other funds and value performance over sustainability. More specifically, they may have invested in ESG funds in expectation of superior performance, that is, for the same reason as investors who initially invested in funds with the top globe ratings and spurned funds with the bottom globe ratings.

5.3. Robustness

5.3.1. Sustainability scores vs. globe ratings

The globe ratings may no longer affect flows because investors rely on other portfolio sustainability metrics. For instance, investors could consider the funds' portfolio sustainability scores as opposed to their globe ratings. The sustainability score is displayed by Morningstar together with other information about the fund, albeit less prominently than the fund's globe rating. It has the advantage to give an absolute ranking of the sustainability of the fund's portfolio rather than ranking the fund relative to the other funds in the same category and may therefore be preferred by investors with pro-social preferences. In this case, the sustainability of a fund's portfolio could attract flows, even if the globe ratings stop being relevant.

To evaluate this possibility, in Table IA.14, we substitute the fund's globe rating with its sustainability score. Consistent with our earlier findings, the sustainability score appears to be positively related to flows only in the first half of the sample period, confirming that only the fund's performance matters for flows in the long term.

5.3.2. New globe rating methodologies and other sustainability metrics

Since the globe ratings' initial introduction, Morningstar has made several changes to the methodology to compute them. These modifications occurred after the sample period on which we have focused so far. Specifically, in October 2018, Morningstar announced some changes to the criteria used to assign the globe ratings, which became effective in November 2018. First, Morningstar started assigning the globe ratings based on a fund's historical sustainability score, which also considers the sustainability of the fund's portfolio in the past, even though more recent scores are assigned higher weights. Second, instead of ranking funds within the Morningstar category, Morningstar started considering the Morningstar Global category, a coarser classification. In this way, funds have a larger number of peers.

The methodology was once again changed in November 2019, when Morningstar started also considering the absolute *Historical Portfolio Sustainability Score* of a fund. Funds in categories like energy could score well within their categories even if their portfolios have poor sustainability. The new methodology does not allow these funds to have a globe rating above 3. Morningstar also introduced a 1 % buffer around the rating cutoffs so that a fund must move by at least 1 % above (below) the threshold to be upgraded (downgraded).

These changes in the methodology for the globe rating computation may indicate that Morningstar wanted to address some of the problems arising from funds' attempts to improve their globe status. Making a fund's globe rating less sensitive to the current portfolio holdings, increasing the number of peers and allowing for a buffer should have decreased funds' incentives to manipulate their globe ratings.

However, in columns 1 to 3 of Table 11, we find no evidence that the arguably improved methodology may have increased the relevance of the sustainability ratings for fund flows. We also consider whether a higher historical sustainability score attracts flows. In column 4, we find that a fund's *Historical Portfolio Sustainability Score* is not statistically significant. These findings mirror our results for the latter part of our main sample period and confirm that the globe ratings and portfolio sustainability scores do not contribute much to the allocation of capital across different funds because investors seem to focus mostly on performance, as captured by the funds' past returns and star ratings.

Finally, we consider an alternative measure to evaluate whether our results can be generalized to other sustainability metrics. This is particularly important because several recent papers have raised concerns about the informativeness of ESG ratings (see, e.g., Freiberg et al., 2020; Cohen et al., 2023). Thus, investors with pro-social preferences may have started using other measures of sustainability. Specifically, we exploit that in April 2018, Morningstar introduced the *Low Carbon Designation*, identifying mutual funds that have portfolios aligned with the transition to a low carbon economy. In column 5, we find no evidence that this new measure affects fund flows, supporting our interpretation that when evaluating the tradeoff between sustainability and performance, mutual fund managers and their investors overwhelmingly choose performance.

Table 11
Morningstar's modified methodologies and fund flows
This table reports the effects of an active fund's globe rating and *Historical Portfolio Sustainability Score* on monthly fund flows after November 2018 (columns 1, 3 and 4) and November 2019 (column 2), when Morningstar implemented two modifications of its globe rating methodology. In columns 1–3, we use globe 3 as the baseline. In column 4, we replace a fund's globe rating with its *Historical Portfolio Sustainability Score*. In column 5, we consider instead the effect of Morningstar's *Low Carbon Designation* after its introduction in April 2018. All specifications include lagged controls for the fund's returns, size, age, and expense ratio as well as interactions of the fund's Morningstar category and year-month fixed effects. Standard errors are clustered at the fund level. Statistical significance at the 10 %, 5 %, and 1 % level is denoted by *, **, and ***, respectively.

	(1) Flow (%TNA) 2018.11–2019.9	(2) 2019.11–2020.9	(3) 2018.11–2020.9	(4) 2018.11–2020.9	(5) 2018.4–2020.9
One Globe	−0.002	0.001	−0.000		
	(−1.097)	(0.870)	(−0.332)		
Two Globes	0.001	−0.000	0.000		
	(0.730)	(−0.080)	(0.516)		
Four Globes	−0.001	−0.000	−0.000		
	(−0.749)	(−0.133)	(−0.520)		
Five Globes	−0.000	−0.001	−0.000		
	(−0.144)	(−0.597)	(−0.217)		
Historical Portfolio Sustainability Score				−0.000 (−0.011)	
Low Carbon Designation					−0.001 (−1.317)
One Star	−0.003**	−0.004**	−0.004***	−0.004***	−0.005***
	(−2.322)	(−2.324)	(−2.848)	(−2.901)	(−4.170)
Two Star	−0.004***	−0.001	−0.003***	−0.003***	−0.003***
	(−3.737)	(−1.269)	(−3.365)	(−3.362)	(−4.011)
Four Star	0.010***	0.009***	0.010***	0.010***	0.009***
	(10.374)	(9.643)	(13.014)	(12.828)	(13.778)
Five Star	0.021***	0.024***	0.023***	0.023***	0.022***
	(11.456)	(12.497)	(14.561)	(14.380)	(15.743)
Fund return	0.004***	0.003***	0.003***	0.003***	0.003***
	(6.683)	(6.892)	(8.526)	(8.739)	(9.968)
Ln TNA	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(−3.650)	(−2.794)	(−3.792)	(−3.499)	(−4.313)
Age	0.002	0.001	0.002*	0.002*	0.001
	(1.626)	(1.022)	(1.687)	(1.718)	(1.301)
Expense Ratio	−0.000**	0.000	−0.000	−0.000	−0.000
	(−2.363)	(1.024)	(−0.720)	(−0.759)	(−1.441)
Constant	−0.002	−0.008	−0.006	−0.007	−0.001
	(−0.306)	(−1.057)	(−0.991)	(−0.877)	(−0.211)
Observations	12,742	12,316	26,207	26,371	33,939
Adjusted R-squared	0.104	0.090	0.096	0.096	0.101
Fixed effects	Cat*YM	Cat*YM	Cat*YM	Cat*YM	Cat*YM

6. Conclusion

Rating financial intermediaries on the basis of the sustainability of their portfolios may appear to be an effective mechanism that allows investors to allocate funds in accordance with their environmental and social preferences. We show that if most investors care to a larger extent about performance, a tradeoff between portfolio sustainability and performance arises, which reduces the subsequent effectiveness of the sustainability ratings.

The behavior of mutual funds and their investors is consistent with evidence showing that a majority of ESG proposals is not supported by shareholders, and in particular by mutual funds He et al., (2023), suggesting that ultimately these investors care predominantly about performance. Our findings indicate that increased transparency may be insufficient, and regulation may be necessary to direct capital to more sustainable investments.

Finally, our results can inform on the drivers of socially responsible investing (SRI) growth. The returns of sustainable stocks have been benefitting from flows into sustainable investments (Pastor et al., 2022). Hence, flows into SRI funds may not necessarily have been driven by investor preferences for sustainable investments because investors may have interpreted sustainability as a signal of superior future performance. Our findings suggest that a stop in flows may translate to a large setback for sustainable funds because sustainable stocks would stop outperforming in the absence of inflows.

CRediT authorship contribution statement

Nickolay Gantchev: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mariassunta Giannetti:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Rachel Li:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

Nickolay Gantchev declares that he has no relevant conflicts or material financial interests that relate to the research described in this paper.

Mariassunta Giannetti declares that she has no relevant conflicts or material financial interests that relate to the research described in this

paper.

Rachel Li declares that she has no relevant conflicts or material financial interests that relate to the research described in this paper.

Appendix: Variable Definition

Variable Name	Definition
Panel A: Fund Trading	
Effective ESG Score	The normalized company-level ESG score minus a Sustainalytics controversy deduction. The company-level ESG score is normalized using a *z*-score transformation within each company's peer group. Morningstar's Portfolio Sustainability Score is based on the weighted average of the stocks' effective scores, with the funds' portfolio shares as weights.
Abnormal ESG Trading	The abnormal ESG trading of fund f in month t is defined as: $Abnormal\ ESG\ Trading(f,t) = ESG\ Trading(f,t) - \frac{1}{24} \times \sum_{\tau=March16-36}^{March16-12} ESG\ Trading(f,\tau)$ where $ESG\ Trading(f,t) = \frac{\sum_{j=1}^{g} abs(NumShares(f,j,t) - NumShares(f,j,t-1)) \times Price(j,t-1)}{\sum_{i=1}^{n} abs(NumShares(f,i,t) - NumShares(f,i,t-1)) \times Price(i,t-1)}$, i is any stock held by fund f and j is $j \in \{High\ ESG\ stocks | NumShares(f,j,t) - NumShares(f,j,t-1) > 0\}\ U\ \{Low\ ESG\ stocks | NumShares(f,j,t) - NumShares(f,j,t-1) < 0\}$
Position Change	The position change in stock i of fund f in quarter t, defined as: $Position\ Change(f,i,t) = \frac{Price(i,t-1)*[(NumShares(f,i,t) - NumShares(f,i,t-1)]}{TNA(f,t-1)}$
Fund turnover (% TNA)	Fund f's quarterly portfolio turnover, computed as the aggregate absolute value of the position change between quarters *t-1* and *t* across all stock holdings, computed using the stock price at time *t-1*, divided by the fund's TNA at the end of quarter *t-1*, multiplied by two.
Abnormal ESG turnover	Fund f's quarterly ESG turnover, computed as the absolute value of the aggregate fund position change between quarters *t-1* and *t* across the fund's holdings of *High ESG stocks* and *Low ESG stocks*, valued using the stocks' prices at time *t-1*, divided by the fund's TNA at the end of quarter *t-1*. $Abnormal\ ESG\ Turnover\ (f,t) = ESG\ Turnover(f,t) - \frac{1}{24} \times \sum_{\tau=March2016-36}^{March2016-12} ESG\ Turnover(f,\tau)$
Panel B: Fund Performance	
Fund excess return	Fund f's monthly net return in excess of the risk-free rate.
DGTW-Adj return	Fund f's monthly portfolio return, risk-adjusted following the methodology of Daniel et al. (1997). Portfolio weights are based on the value of the fund's portfolio holdings at *t-1*.
FF4-Alpha	Fund f's monthly alpha, estimated using Fama-French-Carhart four-factor model on a rolling-window between month *t-60* to *t-1*.
Buy High ESG	The average abnormal return of the high ESG stocks (defined as those with Sustainalytics ESG scores in the top tercile) that fund f has purchased in month t, risk-adjusted following the methodology of Daniel et al. (1997).
Sell Low ESG	The average abnormal return of the low ESG stocks (defined as those with Sustainalytics ESG scores in the bottom tercile) that fund f has sold in month t, risk-adjusted following the methodology of Daniel et al. (1997).
Buy Other	The average abnormal return of other stocks (i.e., stocks with no Sustainalytics ESG scores or stocks with Sustainalytics ESG scores not in the top tercile) that fund f has purchased in month t, risk-adjusted following the methodology of Daniel et al. (1997).
Sell Other	The average abnormal return of other stocks (i.e., stocks with no Sustainalytics ESG scores or stocks with Sustainalytics ESG scores not in the bottom tercile) that fund f has sold in month t, risk-adjusted following the methodology of Daniel et al. (1997).
No-Trade High ESG	The average abnormal return of the high-ESG stocks (defined as those with Sustainalytics ESG scores in the top tercile) that fund f held in month t and did not trade in month t, risk-adjusted following the methodology of Daniel et al. (1997).
No-Trade Low ESG	The average abnormal return of the low-ESG stocks (defined as those with Sustainalytics ESG scores in the bottom tercile) that fund f held in month t and did not trade in month t, risk-adjusted following the methodology of Daniel et al. (1997).
Panel C: Fund Characteristics	
Flow (% TNA)	A fund's quarterly flows, defined as $Flows_{j,q} = \frac{TNA_{j,q} - TNA_{j,q-1} \times (1 + R_{j,q})}{TNA_{j,q-1}}$.
Expense Ratio	Ratio of total fees (as a percentage) that shareholders pay for a fund's operating expenses, including 12b-1 fees.
Ln TNA	Natural logarithm of the fund's month-end total net assets.
Fund Age	Natural logarithm of the fund's age, calculated as the number of years since the oldest share class was made available to investors.
Fund Return	Monthly net return of a fund.
Star Rating	Rating based on a fund's Morningstar Risk-Adjusted Return% Rank for all funds in a given category. Morningstar calculates ratings based on the fund's historical performance in the previous three-, five-, and ten-year periods. The fund must have at least 36 continuous months of historical performance in order to receive a rating. More stars mean better performance. A fund's peer group for the three-, five-, and ten-year ratings is based on the fund's current category without adjusting for category changes. The overall star rating is based on a weighted average (rounded to the nearest integer) of the number of stars received for the past three-, five-, and 10-year performance.
Globe Rating	A fund's sustainability rating, based on its portfolio sustainability scores. Funds are ranked within their Morningstar categories. A fund rating is based on its percentile rank within the fund's Morningstar category. To receive a globe rating, the fund's Morningstar category must have at least 10 funds with portfolio sustainability scores.
Low Carbon Designation	A fund is assigned a Low Carbon Designation by Morningstar if its portfolio holdings have low carbon risk scores and low levels of fossil fuel exposure. The designation is an indicator that the companies held in a portfolio are in general alignment with the transition to a low carbon economy.
Panel D: Stock Characteristics	
Monthly Abnormal Return	A firm's monthly abnormal return calculated using the Fama-French four-factor model, with betas estimated over the previous 36-months, computed using the quarter-end stock price.
Ln Market Cap	Natural logarithm of a firm's market capitalization.
Book to Market	Book-to-market ratio, calculated as book value of equity scaled by market value of equity, computed using the quarter-end stock price.
Leverage	Calculated as the sum of long-term debt and debt in current liabilities scaled by total assets.
ROA	Return on assets, calculated as operating income, divided by lagged total assets.
Sales Growth	Net sales at *t* minus net sales at *t-1*, divided by net sales at *t-1*.
Stock Ret	Quarterly stock return.

References

Albuquerque, R., Koskinen, Y., Zhang, C., 2019. Corporate social responsibility and firm risk: theory and empirical evidence. Manage. Sci. 65, 4451–4469.

Albuquerque, R., Koskinen, Y., Santioni, R., 2023. Mutual Fund Trading and ESG Clientele. Boston College. Unpublished working paper.

Amel-Zadeh, A., Serafeim, G., 2017. Why and how investors use ESG information: evidence from a global survey. Financ. Anal. J. 74, 87–103.

Ammann, M., Bauer, C., Fischer, S., Müller, P., 2018. The impact of the Morningstar sustainability rating on mutual fund flows. Eur. Financ. Manage. 25, 520–553.

Barber, B.M., Morse, A., Yasuda, A., 2021. Impact investing. J. Financ. Econ. 139, 162–185.

Bauer, R., Ruof, T., Smeets, P., 2021. Get real! Individuals prefer more sustainable investments. Rev. Financ. Stud. 34, 3976–4043.

Ben-David, I., Li, J., Rossi, A., Song, Y., 2022. What do mutual fund Investors really care about? Rev. Financ. Stud. 35, 1723–1774.

Berk, J.B., Green, R.C., 2004. Mutual fund flows and performance in rational markets. J. Pol. Econ. 112, 1269–1295.

Bollen, N.P.B., 2007. Mutual fund attributes and investor behavior. J. Financ. Quant. Anal. 42, 683–708.

Bordalo, P., Gennaioli, N., Shleifer, A., 2013. Salience and consumer choice. J. Pol. Econ. 121, 803–843.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. J. Pol. Econ. 105, 1167–1200.

Carhart, M., 1997. On persistence in mutual fund performance. J. Finance 52, 57–82.

Ceccarelli, M., Ramelli, S., Wagner, A.F., 2024. Low carbon mutual funds. Rev. Financ. 28, 45–74.

Cohen, L., Gurun, U.G., Nguyen, Q.H., 2023. The ESG-innovation disconnect: Evidence from Green Patenting. NBER. Unpublished working paperNo. 27990.

Cremers, M.K.J., Riley, T.B., Zambrana, R., 2023. The Complex Materiality of ESG ratings: Evidence from Actively Managed ESG Funds. University of Notre Dame. Unpublished working paper.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. J. Finance 52, 1035–1058.

Del Guercio, D., Tkac, P.A., 2009. Star power: the effect of Morningstar ratings on mutual fund flow. J. Financ. Quant. Anal. 43, 1–30.

Edmans, A., 2011. Does the stock market fully value intangibles? Employee satisfaction and equity prices. J. Financ. Econ. 101, 621–640.

Evans, R.B., Sun, Y., 2021. Models or stars: the role of asset pricing models and heuristics in investor risk adjustment. Rev. Financ. Stud. 34, 67–107.

Fama, E.F., French, K.R., 1993. Common risk factors in the returns on stocks and bonds. J. Financ. Econ. 33, 3–56.

Freiberg, D., Park, D.G., Serafeim, G., Zochowski, R., 2020. Corporate Environmental impact: measurement, Data and Information. Harvard Business School. Unpublished working paper.

Geczy, C., Jeffers, J.S., Musto, D.K., Tucker, A.M., 2021. Contracts with (social) benefits: the implementation of impact investing. J. Financ Econ 142, 697–718.

Han, X., Roussanov, N.L., Hongxun, R., 2022. Mutual Fund Risk Shifting and Risk anomalies. Unpublished working Paper. Jacobs Levy Equity Management Center for Quantitative Financial Research.

Hartzmark, S.M., 2015. The worst, the best, ignoring all the rest: the rank effect and trading behavior. Rev. Financ. Stud. 28, 1024–1059.

Hartzmark, S.M., Sussman, A.B., 2019. Do investors value sustainability? A natural experiment examining ranking and fund flows. J. Finance 74, 2789–2837.

He, Y.E., Kahraman, B., Lowry, M., 2023. ES risks and shareholder voice. Rev. Financ. Stud. 36, 4824–4863.

Holmstrom, B., 1982. Moral hazard in teams. Bell. J. Econ. 13, 324–340.

Kim, S.H., 2022. Do mutual funds manipulate star ratings? Evidence from Portfolio pumping. Unpublished Working Paper. Wilfrid Laurier University.

Heath, D., Macciocchi, D., Michaely, R., Ringgenberg, M.C., 2023. Does socially responsible investing change firm behavior? Rev. Financ. 27, 2057–2083.

Ibert, M., Kaniel, R., Nieuwerburgh, S.V., Vestman, R., 2018. Are mutual fund managers paid for investment skill? Rev. Financ. Stud. 31, 715–772.

Kisgen, D.J., 2006. Credit ratings and capital structure. J. Finance 61, 1035–1072.

Li, T., Naaraayanan, L.S., Sachdeva, K., 2023. Conflicting Objectives of ESG funds: Evidence from Proxy Voting. Unpublished Working Paper. University of Florida.

Lins, K.V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. Journal of Finance 72, 1785–1824.

Ma, L., Tang, Y., Gomez, J.-P., 2019. Portfolio manager compensation in the US mutual fund industry. J. Finance 74, 587–638.

Pastor, L., Vorsatz, M.B., 2020. Mutual fund performance and flows during the COVID-19 crisis. Rev. Asset. Pricing Stud. 10, 791–833.

Pastor, L., Stambaugh, R.F., Taylor, L.A., 2022. Dissecting green returns. J. Financ. Econ. 146, 403–424.

Reuter, J., Zitzewitz, E., 2021. How much does size erode mutual fund performance? A regression discontinuity approach. Rev. Financ. 25, 1395–1432.

Riedl, A., Smeets, P., 2017. Why do investors hold socially responsible mutual funds? J. Finance 72, 2505–2550.

Starks, L.T., 2023. Presidential address: sustainable finance and ESG issues – value versus values. Journal of Finance 78, 1837–1872.