RESEARCH   **Open Access**

ESG scores, scandal probability, and event returns

Wenya Sun[1], Yichen Luo[2], Siu-Ming Yiu[1], Luping Yu[3] and Wenzhi Ding[4*]

*Correspondence:
wenzhi.ding@polyu.edu.hk

[1] The University of Hong Kong, Pok Fu Lam, Hong Kong, China
[2] University College London, London, UK
[3] Xiamen University, Xiamen, Fujian, China
[4] The Hong Kong Polytechnic University, Hung Hom, Hong Kong, China

Abstract

The informativeness of environmental, social, and governance (ESG) scores and their actual impact on firms remains understudied. To address this gap in the literature, we make theoretical predictions and conduct empirical research revealing that a high ESG score is associated with a lower probability of ESG scandals and lower stock returns during a scandal event. Our results suggest that ESG scores are heterogeneous but informative, and that a strong ESG reputation may have both positive and negative consequences for firms. Drawing on our findings, we develop a model and showcase that firms face an optimization problem when determining optimal ESG investment levels. Two equilibria may exist based on the trade-off between ESG scandal losses and ESG adjustment costs. Our model explains why certain firms make heterogeneous ESG decisions

Keywords: ESG score performance, ESG scandal, ESG adjustment cost

JEL Classification: G12, G24, L82, Q01

Introduction

The rise of environmental, social, and governance (ESG)-oriented investment in recent years has called for a deeper understanding of the effects of ESG ratings on stock markets. People tend to believe in human goodwill and to, in turn, perceive that socially responsible firms can be rewarded in the capital market. However, the current empirical evidence does not fully support this view. For example, Hartzmark and Sussman (2019) found that while investors predicted that high-sustainability firms would have better future performance, they did not find empirical evidence to support this view in the fund market. Meanwhile, Serafeim and Yoon (2023) recognized that ESG ratings could predict future ESG news, but they did not observe any connection between ESG ratings and market reactions to ESG news. Larcker and Watts (2020) similarly documented that investors perceive green and non-green securities issued by the same issuer as nearly identical substitutes, with no premium for being green. Hsu et al. (2023) found that highly-toxic emission firms generate higher stock returns than their counterparts, which can be attributed to the risk premium associated with the environmental policy risk they face.

While the literature generally suggests that ESG has a muted or negative effect on stock market performance, some evidence supports a positive correlation. Ferrell et al. (2016) found a positive relationship between corporate social responsibility and firm value.[1] Avramov et al. (2021) theoretically inferred that a green-minus-brown portfolio can generate significantly positive payoffs over a reasonably long investment horizon. In two field surveys conducted in the United States of America, Bauer et al. (2021) found that participants were willing to expand their fund engagement based on ESG, although they expected this engagement to potentially harm financial performance; this inflow of funds into higher ESG-rated firms would naturally increase their stock prices. Ding et al. (2021) also observed that high ESG-rated firms experienced smaller negative stock returns during the stock market turbulence caused by the COVID-19 pandemic.

From a corporate finance perspective, there is evidence to support the link between firm valuation and ESG performance. Regarding the impact of ESG performance on cash flows (product market sales), Dai et al. (2021) found that corporate customers can influence a supplier's corporate social responsibility performance. Ding et al. (2022) discovered that firms enhance their ESG performance to differentiate their products in response to intense product market competition, and that, currently, firms generally use high ESG performance to attract or retain both corporate and retail customers. Regarding the impact of ESG performance on discount rates, Cheng et al. (2014) documented that higher corporate social responsibility performance enables better access to financing, which in turn lowers the cost of capital and, accordingly, the discount rate. Some past studies therefore suggest that higher ESG performance enhances firm valuation. However, considering all the aforementioned studies, it is still unclear how ESG affects long-term stock performance.

Several factors may contribute to the inconclusiveness of this topic. One main argument is the low quality of ESG ratings. There is a growing number of ESG rating agencies worldwide, including Thomson Reuters ASSET4, S&P ESG (formerly RobecoSAM), S&P Trucost, MSCI ESG (formerly MSCI KLD and MSCI IVA), Sustainalytics, and Bloomberg ESG; the issue here is that not only do the rating methodologies and data sources vary across these raters, but they also change over time, and this is even within a specific vendor (Gibson et al. 2021; Christensen et al. 2021; Avramov et al. 2022; Serafeim and Yoon 2023). These characteristics imply that, despite substantial investment in the accumulation of ESG data and the construction of rating methodologies, it is unlikely that every evaluation accurately mirrors a company's true ESG profile. Therefore, both the industry and academia are seeking a better understanding of the validity of the various ESG scores. In this study, we combine a micro-founded model with empirical evidence to study the quality of ESG scores and how they can be utilized to generate better predictive performance.

We first constructed a model to illustrate how a rational investor combines different ESG scores and adjusts a firm's valuation; this model assumes a constant true ESG level for a firm that can influence both ESG ratings and the likelihood of an ESG scandal. Initially, the investor receives inaccurate ESG scores and trades the stock accordingly.

---

[1] Corporate social responsibility and ESG are highly-related concepts with some differences; for more on the topic, see
https://blog.worldfavor.com/esg-vs-csr-what-is-the-difference.

Sun *et al. Financial Innovation* (2024) 10:121

Page 3 of 21

Subsequently, the firm may or may not face an ESG scandal and the investor may adjust its valuation based on the scandal's occurrence. In the final stage, firm cash flow is realized, with a negative adjustment if an ESG scandal occurs in the previous stage. This reflects the market penalty imposed on firms with low ESG performance, as Ding et al. (2022) suggested.

This model generated two testable predictions. First, it suggests that if an ESG scandal occurs, a firm will experience a negative stock market response. Second, although it is commonly believed that higher ESG-rated firms are more resilient during crises, our model implies that higher ESG-rated firms experience more negative stock returns in the event of an ESG scandal. These two perspectives do not contradict each other, as a high ESG-rated firm may be robust against overall market crises and less so when facing its own scandal. This is because ESG scandal occurrence not only implies future penalties in the product market but also a higher likelihood of an overestimation of company ESG score. Both factors contribute to reduced valuation.

To validate our model's key assumptions, we conducted empirical analyses using large-scale ESG ratings and news data. Our findings demonstrate that a higher ESG score generally indicates a lower probability of an ESG scandal. From the stock market perspective, we observed a negative cumulative abnormal return surrounding ESG scandals, and higher ESG-rated firms exhibited lower returns during ESG scandals. We also found that a greater disagreement in ESG scores is associated with lower scandal-related returns.

Based on our empirical evidence, which indicated that higher ESG ratings reduce the likelihood of a scandal but increase the associated losses, we developed a model to illustrate that firms face a trade-off between ESG investment cost and the impact of a scandal. Our findings showcase that this trade-off determines the optimal level of ESG investment for a firm. There are two types of equilibrium in this context, as described herein: a firm with a suboptimal ESG evaluation may incur prohibitive costs when attempting to substantially enhance its ESG rating; then, if the advantages of improved ESG metrics (e.g., minimized repercussions from scandals) fail to surpass the expenditures associated with ameliorating ESG perception, the optimal strategy for the entity would likely be to maintain a suboptimal ESG rating. Our model further suggests that social preferences regarding ESG (i.e., determine the payoff in the product market), firm size, and initial ESG rating levels are key factors in determining a firm's optimal ESG investment decisions.

Our study contributes to several literature fields, mostly studies on the signaling effect of ESG scores. Yoon et al. (2006) found that the sincerity of motives determines the effectiveness of corporate social responsibility activities. Dunbar et al. (2021) observed that the motivation for a firm's ESG engagement is more questionable when the firm is of a high reputation and involved in misconduct. Our study contributes to this field by deriving separate equilibria for firms' optimal ESG investments. These equilibria have implications for practitioners when considering their ESG investment decisions based on the social preferences for ESG, firm size, and initial ESG level. To the best of our knowledge, this study is the first to deliver these pieces of evidence.

Our study also contributes to the literature on ESG rating properties. As aforementioned, there is an ongoing debate on whether ESG scores are informative (Bartov

et al. 2020; Gibson et al. 2021; Christensen et al. 2021; Berg et al. 2022; Avramov et al. 2022). Meanwhile, this study shows that despite the heterogeneity of ESG score, related scores remain capable of providing predictive power, at least to some extent, for ESG scandal probability and scandal-related returns.

Finally, we contribute to discussions on the real impact of ESG. Several studies show that corporate customers' and investors' ESG preferences shape firm ESG practices and financial decisions (Ferrell et al. 2016; Capelle-Blancard and Petit 2019; Chen et al. 2020; Dai et al. 2021). We add to this field by developing a model that characterizes the optimization problem faced by firms with ESG concerns. In particular, we introduce four ESG investment-related dimensions that can be empirically tested in future studies: firm size, market ESG preference, and initial reputation level. Furthermore, investors' ESG preferences can be split into value alignment and impact seeking (Bonnefon et al. 2022) and individually tested.

The remainder of this paper is organized as follows. "Data" Section describes the data used in our empirical analyses. "Analysis" Section presents the development of the model used to generate testable predictions, illustrates the adopted empirical methodology, and discusses the empirical results. "Firm's optimal ESG investment" Section introduces the firm's optimal ESG investment problem based on our empirical findings. "Conclusion" Section concludes the manuscript.

Data

ESG scores were originally developed in the 1980s to supplement traditional financial data and provide investors with additional information about companies. These scores quantify company performance in terms of ESG issues. This study collected ESG data from four notable ESG rating providers, as follows: KLD (now MSCI), ASSET4 (now Refinitiv ESG), Sustainalytics (now Morningstar), and S&P Global. These vendors cover a significant proportion of the companies that have ESG performance ratings. Table 1 illustrates the similarities between the ESG scores of the four rating entities under scrutiny. The average ESG score of a company was computed as the mean of the four scores from the rating providers.

This study also used another score, the principal component analysis (PCA) ESG score, which represents the collective perspective of rating agencies toward a company at a given time. We conducted PCA on the ratings obtained from the four vendors to derive the first principal component, subsequently defining it as the PCA ESG

Table 1 Input similarity matrix

Correlation	ASSET4	S&P	Sustainalytics	KLD
ASSET4	1.00			
S&P	0.49	1.00		
Sustainalytics	0.53	0.63	1.00	
KLD	0.39	0.45	0.53	1.00

The presented table displays correlation coefficients between ASSET4, S&P, Sustainalytics, and KLD, providing insights into the interrelationships among these ESG rating entities, with correlation values indicating the degree of association between their respective ESG rating

Table 2 Principle components

Component	Eigenvalue	Difference	Proportion	Cumulative
Comp1	2.52	1.90	0.63	0.63
Comp2	0.62	0.10	0.15	0.78
Comp3	0.51	0.16	0.13	0.91
Comp4	0.35		0.09	1.00

This table presents the principal components obtained from the analysis, including their eigenvalues, differences between consecutive eigenvalues, proportions of total variance explained, and cumulative proportions. These metrics provide insights into the significance and contribution of ESG scores from 4 different data agencies

score. The PCA ESG score captures 63% of the variance in the overall ESG scores rated by the four ESG rating agencies, as presented in Table 2.

Disagreement among ESG providers regarding a company's performance can also influence stock reactions following ESG scandals. Therefore, we calculated ESG disagreement as the standard deviation of the ratings provided by the four agencies (see Appendix B for a more extensive discussion on the development of ESG scores, their distinctions, and recent studies on the subject).

To identify ESG scandals, we used event data compiled by RepRisk. RepRisk continuously monitors over 100,000 public sources and stakeholders for ESG incidents. The methodology employed was event-driven rather than company-driven. The RepRisk news dataset catalogues ESG-related incidents between 2007 and 2020. We classified incidents recorded in this dataset as ESG scandals by creating a binary variable called ESG Scandal, which is one if a firm has at least one RepRisk ESG scandal in the next 12 months, and zero otherwise.

For event returns analysis, we first compiled a list of ESG scandals and their corresponding event dates from RepRisk. We then obtained stock returns data for the involved firms around the scandal event date from the CRSP. Following Brown and Warner (1985) approach, we employ capital asset pricing model to estimate the predicted returns. The model parameters were estimated between 250 and 20 trading days prior to the scandal events. Expected and cumulative abnormal returns were subsequently calculated for the period between 10 trading days before and after the scandal events.

We also retrieved financial performance data from the Compustat Fundamentals database to address the control variables in our study. Annual financial information was systematically gathered, encompassing a range of factors such as total assets, cash and short-term investments, current assets, current liabilities, intangible assets, earnings before interest, long-term debt due in one year, total debt including current, sales, operating activities net cash flow, capital expenditures, total dividends, and net income. Drawing on the established literature on market returns correlations (Haque and Sarwar 2013; Sarwar et al. 2013; Brown and Huang 2020), 11 key financial characteristics were calculated as control variables. These variables were size, cash ratio, current ratio, intangibility, return on assets, maturing debt, leverage, growth, cash flow volatility, capital expenditure, and dividend payout ratio. The calculation for each control variable is detailed in Appendix A, along with a succinct textual explanation of the variable for clarity. Our two samples are described in "Scandal probability", and "Event returns" Section, respectively.

Sun *et al. Financial Innovation*   (2024) 10:121

Page 6 of 21

Analysis

Theoretical predictions

In this section, we build a simple model to predict the relationship between ESG scores and event returns when a scandal occurs. The model considers one firm over three periods. The firm has a real ESG level $v \in [0, 1]$, which is unknown to investors. At $t = 0$, ESG rating agencies disclose ESG score $s \sim N(v, \tau)$ for the firm. τ is the rating difficulty of the firm, s follows normal distribution. In $t = 1$, the firm is exposed to an ESG scandal at a probability $p(v)$, where $p(\cdot) \in [0, 1]$ and $p' < 0$. This leads to the following hypothesis:

Hypothesis 1   Firms with high ESG rating have a lower probability of experiencing an ESG scandal.

At $t = 2$, the firm realizes its product sales $D(S, \lambda) = 1 - S\lambda$, where $\lambda \in [0, 1]$ is the aggregated customer's ESG preference and $S \in \{0, 1\}$ is a dummy indicating whether the scandal happened or not. Importantly, the formulation assumes a constant level of price elasticity throughout the analysis, implying that the relationship between product sales and aggregated customer's ESG preference, as captured by the function $D(S, \lambda)$, is based on a consistent price elasticity factor. Although this assumption simplifies the analysis, we acknowledge that price elasticity variations can affect revenue sustainability; accordingly, sensitivity analyses that explore different elasticity scenarios may provide additional insights into our model's robustness.

We then derive the model to generate predictions. At $t = 0$, the value (product sale) of the firm can be calculated by computing the mathematical expectation of ESG scandal occurrence, as follows[2]:

$$
\begin{aligned}
E[D|s] &= p(s)D(1, \lambda) + [1 - p(s)]D(0, \lambda) \\
&= 1 - p(s)\lambda.
\end{aligned}
\tag{1}
$$

Here $D(1, \lambda) = 1 - \lambda$ is the product sales when the scandal happens, and $D(0, \lambda) = 1$ is the product sales when the scandal does not happen. After $t = 1$, it is revealed whether the ESG scandal happens. The value of the firm should be $1 - S\lambda$, making the return for period $t = 1$ be

$$
R = \frac{1 - S\lambda}{1 - p(s)\lambda} - 1
\tag{2}
$$

If there is no scandal revealed in $t = 1$, the return $R_{S=0}$ is

$$
R_{S=0} = \frac{1}{1 - p(s)\lambda} - 1 > 0.
\tag{3}
$$

If a scandal happened, the return $R_{S=1}$ is

$$
R_{S=1} = \frac{1 - \lambda}{1 - p(s)\lambda} - 1 < 0.
\tag{4}
$$

---

[2] For technical brevity, we assume that the market has no prior belief on v, so the probability of ESG scandal is simply $p\left(s\right)$, where s is the ESG score aggregated by various ESG ratings ${s}_{i}$ on the market (i indexes different rating agencies). The conclusions are qualitatively unchanged if there is a prior of v.

Sun *et al. Financial Innovation*   (2024) 10:121

Page 7 of 21

This generates a testable prediction, as follows,

Hypothesis 2   The investor penalizes the firm that experiences an ESG scandal on the stock market.

Since $p'(s) < 0$, the partial derivative with regard to s is

$$\frac{\partial R_{S=1}}{\partial s} = \frac{(1-\lambda)\lambda p'(s)}{(1-p(s)\lambda)^2} < 0. \tag{5}$$

This generates another testable prediction, as described herein,

Hypothesis 3   If the ESG score is higher, the return on scandal event is lower.

Scandal probability

Following the methodology outlined by Anderson et al. (2018) and Agrawal et al. (2022), we employed a panel logistic Poisson regression with fixed effects to investigate the influence of ESG scores on the probability of scandals within companies during a corresponding month. Our primary objective was to discern whether ESG scores impact the likelihood of a company experiencing a scandal, and if so, to what extent. To achieve this, we estimate scandal probability using the following model:

$$Scandal_{c,t} = \beta_1 Score_{c,t} + \beta_2 Disagree_{c,t} + \Gamma F_{c,t} + \delta_c + \gamma_t + \varepsilon_{c,t}, \tag{6}$$

where $Scandal_{c,t}$ is a dummy indicating whether firm c will experience an ESG scandal in the next 12 months after time t. $Score_{c,t}$ is the ESG score of firm c at time t. We calculate $Disagree_{c,t}$ as the standard deviation of the ESG scores from the four rating agencies of S&P, ASSET4, KLD, and Sustainalytics. $F_{c,t}$ is a vector of company fundamentals comprising 11 factors, namely size, cash ratio, current ratio, intangibility, return on assets, maturing debt, leverage, growth, cash flow volatility, capital expenditure, and dividend payout ratio. δ_c is the company fixed effect, which serves to condition out the company-invariant effect on scandal probability. Similarly, the time fixed effect γ_t is to condition out the time-invariant impact on scandal probability.

In our experiments, we conducted six fixed-effect linear regressions using different sets of independent variables. To provide more details, we employed six different ESG scores (i.e., S&P, ASSET4, KLD, Sustainalytics, average scores of the four, and the first PCA ESG score derived from the four) in combination with the same set of company characteristics to evaluate the probability of a company scandal.

We present the summary statistics for the sample in Panel A of Table 3, which consists of 127,478 firm-month observations. An average ESG scandal score of 0.26 suggests a 26% likelihood of a firm experiencing an ESG scandal within any given 12-month interval. This figure may appear elevated, reflecting RepRisk's expansive criteria for identifying ESG scandals; that is, some scandal events reported by RepRisk may not be significant or well-known. To ensure consistency, we standardize all ESG scores to have a mean of zero and a standard deviation of one. Standardized ESG scores were then used to calculate the average and the PCA ESG scores.

Sun *et al. Financial Innovation*   (2024) 10:121

Page 8 of 21

Table 3 Summary descriptives

	N	Mean	Std. Dev	10th Percentile	Median	90th Percentile
Panel A. Event probability sample						
ESG Scandal	127,478	0.26	0.44	0.00	0.00	1.00
ESG disagree	127,478	0.00	1.00	− 1.20	− 0.12	1.36
ASSET4	127,478	0.00	1.00	− 1.19	− 0.12	1.31
S&P	127,478	0.00	1.00	− 1.24	− 0.21	1.50
Sustainalytics	127,478	0.00	1.00	− 1.27	− 0.05	1.30
KLD	127,478	0.00	1.00	− 1.03	− 0.20	1.47
Average	127,478	0.00	0.62	− 0.81	0	0.80
PCA	127,478	0.00	1.31	− 1.61	− 0.13	1.81
Size	127,478	0.28	0.59	− 0.12	0.17	0.85
Cash ratio	127,478	0.00	0.06	− 0.06	0.00	0.06
Current ratio	127,478	− 0.04	0.72	− 0.50	− 0.02	0.39
Intangibility	127,478	0.01	0.08	− 0.05	0.00	0.08
ROA	127,478	0.00	0.05	− 0.05	0.00	0.04
CAPEX	127,478	− 0.01	0.02	− 0.03	0.00	0.01
Maturing debt	127,478	0.01	0.12	− 0.08	0.00	0.11
Leverage	127,478	0.22	0.41	0.01	0.17	0.49
Dividend payout	127,478	0.03	0.92	− 0.52	− 0.01	0.42
Sales growth	127,478	− 0.02	0.10	− 0.12	− 0.02	0.06
Cash flow volatility	127,478	0.00	0.02	− 0.02	0.00	0.01
Panel B. Event return sample						
ESG scandal returns	40,045	− 0.29	6.20	− 7.58	− 0.34	6.97
ESG disagree	40,045	0.00	1.00	− 1.13	− 0.14	1.33
ASSET4	40,045	0.00	1.00	− 1.46	0.45	0.74
S&P	40,045	0.00	1.00	− 1.21	− 0.21	1.51
Sustainalytics	40,045	0.00	1.00	− 1.39	0.06	1.32
KLD	40,045	0.00	1.00	− 1.27	− 0.11	1.33
Average	40,045	0.00	0.79	− 1.16	0.10	1.00
PCA	40,045	0.00	1.58	− 2.30	0.20	2.03
Size	40,045	0.21	0.37	− 0.11	0.13	0.71
Cash ratio	40,045	− 0.01	0.05	− 0.06	− 0.01	0.05
Current ratio	40,045	− 0.13	0.73	− 0.50	− 0.06	0.24
Intangibility	40,045	0.01	0.09	− 0.06	0.00	0.10
ROA	40,045	− 0.01	0.05	− 0.06	− 0.01	0.04
CAPEX	40,045	0.00	0.02	− 0.02	0.00	0.01
Maturing debt	40,045	0.01	0.08	− 0.06	0.00	0.08
Leverage	40,045	0.30	0.22	0.08	0.28	0.53
Dividend Payout	40,045	0.06	0.97	− 0.52	0.00	0.30
Sales growth	40,045	− 0.02	0.07	− 0.10	− 0.02	0.05
Cash flow volatility	40,045	0.00	0.02	− 0.01	0.00	0.02

This table documents the statistics of the ESG scandal, ESG ratings (ASSET4, S&P, Sustainalytics, and KLD), and firm fundamental characteristics. Panel A reports the summary statistics for the event probability sample, corresponding to the analysis in "Scandal probability" Section. *N* represents the number of firm-month observations. Panel B reports the summary statistics for the event returns sample, corresponding to the analysis in "Event returns" Section. *N* represents the number of event observations. All continuous variables are winsorized at a 1% level

The empirical effect of ESG scores on scandal probability was tested, and the regression results are presented in Table 4. As shown in Table 4, the S&P, KLD, average, and PCA ESG scores had significant and negative effects on scandal occurrence. The

Sustainalytics score also exhibited negative associations with scandal probability, albeit not to a statistically significant extent. In terms of economic magnitude, a one standard deviation increase in ESG scores corresponds to a 0.1–1.1% decrease in scandal probability. These findings generally support Hypothesis 1. We provide additional empirical results based on the alternative specifications in Panel B of Table 4.

The coefficient estimates of the control variables supported the validity of the measures. Larger firms tended to have a higher probability of ESG scandals, possibly due to the greater scrutiny they receive. Firms with higher intangibility and profitability had a lower probability of ESG scandals, indicating that they can prioritize good governance and social responsibility because of less earnings pressure. Conversely, high leverage was associated with a higher probability of ESG scandals, highlighting the correlation between financial and ESG risks. A robust sales growth not only provides companies with financial stability and resources to invest in ethical practices (Bint-Tariq and Nobanee 2020), but also incentivizes a commitment to ESG principles, fostering stakeholder trust, regulatory compliance, and a long-term perspective, thereby reducing ESG scandal likelihood.

Event returns

To study the relationship between ESG scores and event returns, we first used the following ordinary least squares regression:

$$CAR_{c,i,t} = \beta_1 Score_{c,t} + \beta_2 Disagree_{c,t} + \Gamma F_{c,t} + \varepsilon_{c,i,t} \qquad (7)$$

where c, i, and t are the index company, scandal event, and time, respectively. *Score* is the particular ESG score being studied, and *Disagree* is the disagreement in ESG scores among the four data agencies. *F* is the vector of the firm's fundamental characteristics. We estimate an ordinary least squares regression with robust standard errors clustered at the company level. Similar to the findings for scandal probability, we conducted six regression analyses (i.e., ASSET4, S&P, Sustainalytics, KLD, the average score of the four, and the PCA ESG score).

Panel B of Table 3 presents the summary statistics of the sample.[3] The total number of scandals was 40,045,[4] and the average ESG scandal return was −0.29, indicating that the cumulative abnormal return for an ESG scandal event was −0.29%. This descriptive statistic supports Hypothesis 2. However, it is important to note that this statistic is not economically significant as RepRisk identifies ESG scandals in a broad manner, including events that may not be significant or well-known.

As Table 5 shows, firms with higher ESG scores tended to experience more severe ESG scandals, as all coefficients related to ESG scores were statistically significant and negative. In terms of economic significance, a one-standard-deviation increase in ESG score was responsible for an average decrease of 7–20 basis points in event returns. This

---

[3] To partially address multicollinearity problems, we conducted correlation analysis between our key explanatory variables and control variables. The results are presented in Appendix C, and the correlation coefficients were small in absolute terms.

[4] The ratio between these two sample sizes was 31% (40,045/127,478) and larger than 26% (average ESG scandal for the first sample) because some firms may have multiple scandals within 12 months. Such cases contribute multiple observations to the scandal return sample, but are underrepresented in the scandal probability sample because the dependent variable for this sample is a dummy.

Sun *et al. Financial Innovation*   (2024) 10:121

Page 10 of 21

Table 4 ESG score and scandal probability

Score	ESG Scandal					
	ASSET4 (1)	S&P (2)	Sustainalytics (3)	KLD (4)	Average (5)	PCA (6)
Panel A. Baseline						
Score	0.001 (0.03)	− 0.109** (− 2.31)	− 0.098 (− 0.85)	− 0.065** (− 2.03)	− 0.022** (− 2.31)	− 0.012*** (− 2.62)
ESG Disagree	0.038 (0.41)	0.061 (0.65)	0.036 (0.39)	0.135 (1.27)	0.107 (1.10)	0.113 (1.15)
Size	0.029*** (3.97)	0.029*** (3.97)	0.029*** (3.98)	0.029*** (3.97)	0.030*** (4.03)	0.030*** (3.99)
Cash Ratio	− 0.014 (− 0.12)	− 0.010 (− 0.09)	− 0.013 (− 0.11)	− 0.018 (− 0.15)	− 0.012 (− 0.10)	− 0.012 (− 0.10)
Current Ratio	0.001 (0.12)	0.001 (0.12)	0.001 (0.10)	0.002 (0.16)	0.001 (0.08)	0.001 (0.10)
Intangibility	− 0.147** (− 2.14)	− 0.145** (− 2.10)	− 0.146** (− 2.12)	− 0.146** (− 2.12)	− 0.145** (− 2.10)	− 0.144** (− 2.08)
ROA	− 0.437*** (− 4.37)	− 0.441*** (− 4.41)	− 0.433*** (− 4.32)	− 0.443*** (− 4.43)	− 0.437*** (− 4.37)	− 0.436*** (− 4.36)
CAPEX	0.326 (1.09)	0.343 (1.15)	0.326 (1.09)	0.336 (1.12)	0.340 (1.14)	0.341 (1.14)
Maturing Debt	− 0.066 (− 1.58)	− 0.068 (− 1.63)	− 0.066 (− 1.59)	− 0.064 (− 1.55)	− 0.067 (− 1.60)	− 0.066 (− 1.59)
Leverage	0.029** (2.03)	0.029** (2.02)	0.029** (2.02)	0.029** (2.02)	0.029** (2.00)	0.029** (1.99)
Dividend Payout	− 0.003 (− 0.80)	− 0.003 (− 0.81)	− 0.003 (− 0.79)	− 0.003 (− 0.75)	− 0.003 (− 0.75)	− 0.003 (− 0.76)
Sales Growth	− 0.102* (− 1.71)	− 0.101* (− 1.71)	− 0.102* (− 1.72)	− 0.099* (− 1.67)	− 0.103* (− 1.74)	− 0.101* (− 1.71)
Cash Flow Vol	0.219 (1.30)	0.217 (1.29)	0.216 (1.28)	0.228 (1.35)	0.215 (1.28)	0.219 (1.30)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	127,478	127,478	127,478	127,478	127,478	127,478
# of firms	2238	2238	2238	2238	2238	2238

Score	ESG Scandal					
	Average (1)	PCA (2)	Average (3)	PCA (4)	Average (5)	PCA (6)
Panel B. Alternative specifications						
Score	− 0.007* (− 1.93)	− 0.004** (− 2.02)	− 0.006** (− 2.29)	− 0.004** (− 2.54)	− 0.017* (− 1.81)	− 0.010** (− 2.17)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes				
Year-month FE			Yes	Yes	Yes	Yes
Sample period	Sub-sample 2016–2020	Sub-sample 2016–2020	Full sample 2013–2020	Full sample 2013–2020	Full sample 2013–2020	Full sample 2013–2020
Estimation	OLS	OLS	OLS	OLS	Logit	Logit
N	95,202	95,202	127,478	127,478	116,388	116,388
Adjusted R2	0.406	0.406	0.405	0.405		
# of firms	2236	2236	2238	2238		

This table presents the relationship between the ESG score and the probability of a scandal. The dependent variable is the occurrence of an ESG scandal, which is represented by a value of one if the firm experiences at least one scandal within the next twelve months, and zero otherwise. The explanatory variables consist of six ESG scores. In Panel A, we utilize a full sample from 2013 to 2020 and estimate the results using a robust statistical approach—panel Poisson regression, incorporating fixed effects for firms and years. In Panel B, we provide results for (1) a sub-sample from 2016 to 2020 (5-year period); (2) firm and year-month fixed effects; and (3) estimation using panel Poisson regression. Detailed definitions of all variables can be found in Table 4. We have clustered the standard errors at the firm level. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively

finding supports our Hypothesis 3. We also provide the empirical results for alternative specifications in Panel B of Table 5.

Furthermore, we observed that disagreements among ESG rating providers significantly affected event returns, wherein the greater the disagreement, the more negative the impact of the scandal on the company's stock returns. Specifically, a one-standard-deviation increase in ESG disagreement was associated with a decrease of 22–26 basis points.

Additionally, we compared the predictability of various ESG scores for scandal probability and event returns (Table 6), so as to demonstrate our main results' robustness. Importantly, this study did not specifically focus on comparing predictability across different ESG scores.

Firm's optimal ESG investment

In this section, we construct a model to show that, based on our empirical findings (see "Analysis" Section), each firm faces an optimization problem in determining ESG investment. We define firm cash flow as

$$X_t = (\Pi_t - \Phi_t) - (\Gamma_t + \Theta_t) \tag{8}$$

where X_t is the residual cash for the investor of a firm; Π_t is firm gross profit; Φ_t is the adjustment costs of capital (Bonnefon et al. 2022)[5]; Γ_t is the adjustment costs of reputation capital (the cost of changing ESG level); Θ_t is the expected loss from an ESG scandal. Furthermore, and similar to the approach used by Liu et al. (2009), the adjustment costs of capital are defined as

$$\Phi_t = \phi\left(\frac{I_t^K}{K_t}\right) K_{t-1}, \tag{9}$$

$$\Gamma_t = \gamma\left(\frac{I_t^G}{G_t}\right) G_{t-1}, \tag{10}$$

where $\phi(\cdot)$ increases with the capital investment level $\frac{I_t^K}{K_t}$, and $\gamma(\cdot)$ increases with the reputation investment level $\frac{I_t^G}{G_t}$. These are the unit costs of capital adjustment. Therefore, the total cost should be scaled by the original size of capital, K_{t-1}, and reputation, G_{t-1}.

Relating to our empirical findings, we calculate the expected loss from ESG scandal (Θ_t) similar to the calculation for expected loan loss in a commercial bank[6]:

$$\Theta_t = P\left(\frac{G_t}{K_t}\right) \theta\left(\frac{G_t}{K_t}; \lambda\right) K_t \tag{11}$$

where $P(\cdot)$ is scandal probability and $P \in [0, 1]$; $\theta(\cdot)$ is the loss given a scandal, while $\theta' > 0$; $\frac{G_t}{K_t}$ is the ESG level relative to firm size; λ is the aggregated ESG preference of a market.

Sun *et al. Financial Innovation*   (2024) 10:121

Page 12 of 21

Table 5 ESG score and scandal returns

Score	ESG scandal returns (%)					
	ASSET4	**S&P**	**Sustainalytics**	**KLD**	**Average**	**PCA**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel A. Baseline						
Score	−0.101** (−2.23)	−0.145*** (−3.38)	−0.061 (−1.41)	−0.134*** (−3.24)	−0.188*** (−3.30)	−0.093*** (−3.25)
ESG disagree	−0.264*** (−6.48)	−0.246*** (−6.27)	−0.245*** (−6.16)	−0.261*** (−6.40)	−0.267*** (−6.61)	−0.266*** (−6.59)
Size	−0.495 (−0.48)	−0.497 (−0.48)	−0.503 (−0.48)	−0.465 (−0.45)	−0.367 (−0.35)	−0.369 (−0.35)
Cash ratio	0.370*** (5.47)	0.379*** (5.63)	0.358*** (5.27)	0.366*** (5.42)	0.366*** (5.43)	0.366*** (5.42)
Current ratio	−0.378 (−0.94)	−0.286 (−0.71)	−0.339 (−0.83)	−0.310 (−0.77)	−0.269 (−0.67)	−0.269 (−0.66)
Intangibility	11.909*** (9.88)	11.900*** (9.86)	12.060*** (10.04)	11.708*** (9.72)	11.881*** (9.84)	11.895*** (9.86)
ROA	3.487 (1.18)	3.350 (1.13)	3.031 (1.02)	2.899 (0.98)	3.139 (1.06)	3.131 (1.06)
CAPEX	0.479 (0.90)	0.448 (0.84)	0.450 (0.84)	0.429 (0.81)	0.477 (0.90)	0.477 (0.90)
Maturing debt	−0.074 (−0.43)	−0.074 (−0.43)	−0.074 (−0.43)	−0.114 (−0.66)	−0.092 (−0.54)	−0.091 (−0.53)
Leverage	−0.074* (−1.73)	−0.079* (−1.84)	−0.077* (−1.80)	−0.078* (−1.82)	−0.076* (−1.79)	−0.076* (−1.79)
Dividend payout	−2.643*** (−3.47)	−2.459*** (−3.22)	−2.557*** (−3.35)	−2.390*** (−3.14)	−2.525*** (−3.32)	−2.527*** (−3.32)
Sales growth	−9.071*** (−2.98)	−9.170*** (−3.02)	−8.992*** (−2.96)	−8.363*** (−2.76)	−9.087*** (−2.99)	−9.103*** (−2.99)
Cash flow Vol	−0.101** (−2.23)	−0.145*** (−3.38)	−0.061 (−1.41)	−0.134*** (−3.24)	−0.188*** (−3.30)	−0.093*** (−3.25)
N	40,045	40,045	40,045	40,045	40,045	40,045
Adjusted R2	0.014	0.014	0.014	0.014	0.014	0.014
# of firms	767	767	767	767	767	767

Score	ESG scandal returns (%)					
	Average	**PCA**	**Average**	**PCA**	**Average**	**PCA**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
Panel B alternative specifications						
Score	−0.054* (−1.69)	−0.026 (−1.64)	−0.106*** (−2.60)	−0.051** (−2.54)	−0.117** (−2.06)	−0.057** (−2.03)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Sample period	Full sample 2013–2020	Full sample 2013–2020	Full sample 2013–2020	Full sample 2013–2020	Sub−sample 2016–2020	Sub−sample 2016–2020
Event window	[−3, 3] days	[−3, 3] days	[−5, 5] days	[−5, 5] days	[−10, 10] days	[−10, 10] days
N	40,045	40,045	40,045	40,045	29,139	29,139
Adjusted R2	0.005	0.005	0.009	0.009	0.017	0.017
# of firms	767	767	767	767	753	753

This table presents the relationship between ESG scores and scandal returns. The dependent variable is *ESG Scandal Returns*, which represents the cumulative abnormal returns from −10 to 10 days relative to the date of the scandal event. The explanatory variables consist of six ESG scores. Panel A utilizes the full sample from 2013 to 2020 and estimates the results using cross-sectional OLS. Panel B displays the results for: (1) cumulative abnormal returns from −3 to 3 days, (2) cumulative abnormal returns from -5 to 5 days, and (3) a sub-sample from 2016 to 2020 (5-year period). All variables mentioned here are defined in Table 4. We have clustered the standard errors at the firm level. *, **, and *** represent significance levels of 10%, 5%, and 1%, respectively

Sun *et al. Financial Innovation*   (2024) 10:121

Page 13 of 21

Table 6 Predictability of ESG scores

ESG Score	Scandal probability			Scandal return		
	Adjusted R^2	Increment (10^{-6})	Ranking	Adjusted R^2	Increment (10^{-6})	Ranking
–	0.40547			0.01366		
ASSET4	0.40548	1.1	5	0.01386	199.2	5
S&P	0.40550	22.2	3	0.01426	601.0	3
Sustainalytics	0.40547	− 2.2	6	0.01376	103.3	6
KLD	0.40548	3.4	4	0.01492	1258.9	1
Average	0.40550	26.0	2	0.01426	605.8	2
PCA	0.40550	28.2	1	0.01424	584.4	4

This table showcases the variation in adjusted R-squared when incorporating each ESG score into the regression model. The baseline is the regression model without any ESG score being considered as an independent variable

Therefore, a firm faces two sub-problems in order to maximize its cash flow (X_t): (1) maximize operating profit $\Pi_t - \Phi_t$ and (2) minimize ESG cost $\Gamma_t + \Theta_t$. To make this problem easier to understand, we arbitrarily assume that the function forms

$$\begin{cases} P = \left(1 - \frac{G}{K}\right)^2 \\ \theta = \lambda \ln\left(1 + \frac{G}{K}\right) \\ \Gamma = \gamma G \\ \gamma\left(\frac{I^G}{G}\right) = \left(\frac{I^G}{G}\right) + \left(\frac{I^G}{G}\right)^2 \end{cases} \tag{12}$$

Figure 1 shows how expected losses from scandals change by ESG levels. When a firm is considered 100% unethical, ESG scandals imply no loss because the market already acknowledges that the firm regularly engages in unethical practices. Similarly, if a firm is perceived as 100% ethical, it has nothing to lose from an ESG scandal because the probability of its involvement in such an event is zero. Still, as the market becomes more concerned about ESG, the expected loss increases, as illustrated in Fig. 2.
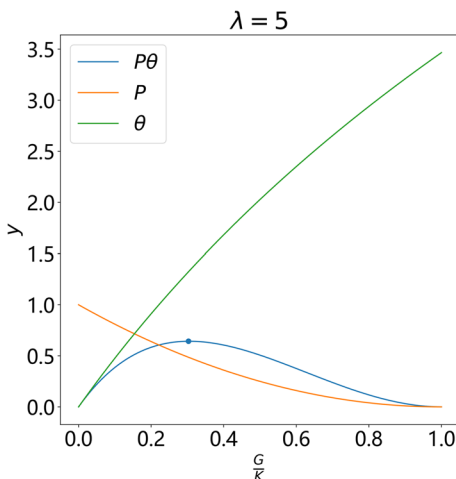

Fig. 1 Expected Loss from Scandal. This figure shows how expected loss from scandal is determined. *P* is the probability of scandal events. θ is the loss given ESG scandal happens. $P\theta$ is the expected loss from scandal. Both *P* and θ are functions of $\frac{G}{K}$, which is the standardized reputation level. *G* is reputation level. *K* is firm size
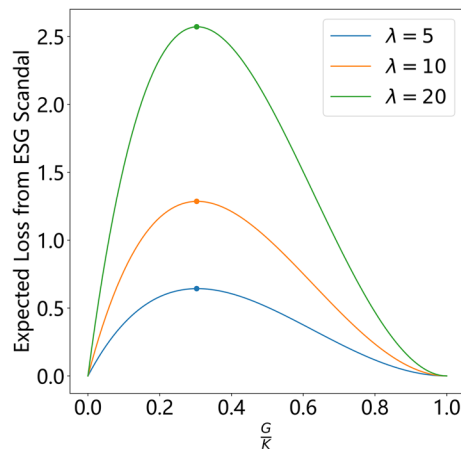
Fig. 2 Expected Loss from Scandal under Different ESG Preference. This figure shows how expected loss from scandal is affected by different levels of ESG preference λ. $\frac{G}{K}$ is the standardized reputation level. G is reputation level. K is firm size

It also remains that firms cannot merely strive to be 100% ethical or unethical to minimize expected losses, as there is the existence of an adjustment cost associated with reputation, which increases proportionally to the magnitude of the adjustment. Consequently, firms must determine an optimal ESG investment level to minimize overall ESG-related costs, as Fig. 3 demonstrates. As shown in Panel A of Fig. 4, in a market where ESG preference is absent, the optimal choice for firms is not to invest in ESG, whereas an increase in ESG preference leads to an increase in optimal ESG investment. Meanwhile, Panel B of Fig. 4 indicates that for firms with poor ESG performance, increasing ESG investment may not yield sufficient benefits. Consequently, these firms may choose to remain at their original levels. For firms with high ESG performance, the required investment is relatively low. Panel C of Fig. 4 highlights the significance of matching firm size with the optimal ESG investment level.
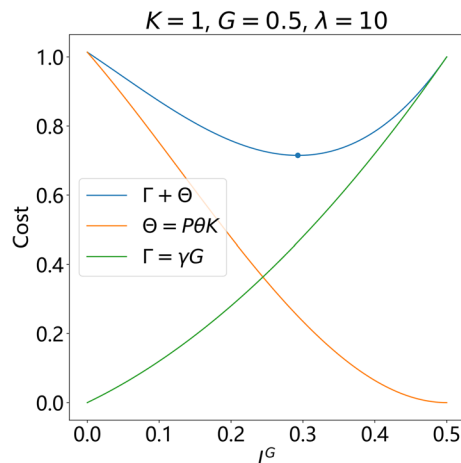

Fig. 3 Optimal ESG Investment. This figure shows how firm's optimal ESG investment is determined. Θ is the expected loss from ESG scandal. Γ is the costs of ESG investment. $\Gamma + \Theta$ is the total costs of ESG. I^G is the ESG investment level
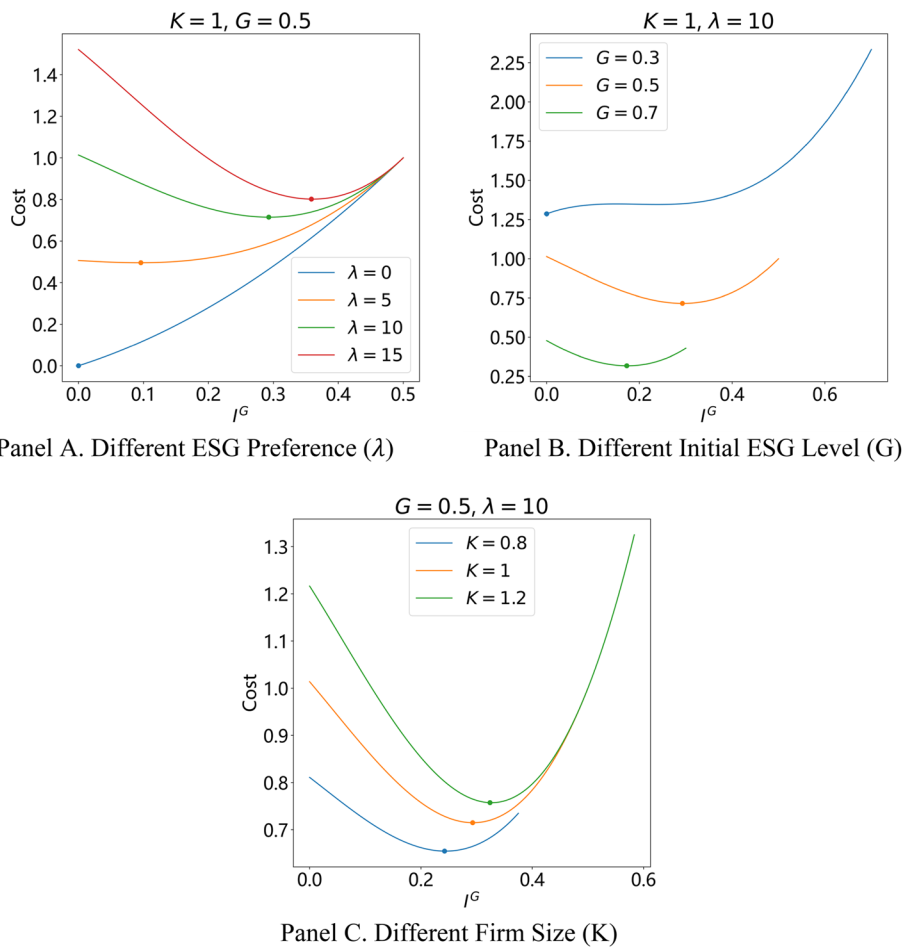
Fig. 4 Optimal ESG investment under different conditions. These figures show how optimal ESG investment level is changed by the changes in different parameters. K is firm size. G is reputation level. I^G is the ESG investment level. λ is ESG preference

Although this study does not provide an analytical solution to the model, it presents several additional topics for future exploration. For emple, by calibrating the ESG prefer-ence parameter (λ), one can determine how the market views ESG. In the case of using a cross-country sample, the simplest method to estimate ESG preference is by calculating the country-fixed effects for ESG scandal event returns (the cumulative abnormal return, labeled as CAR in Table 5). Another less direct approach is to link the data on ESG scan-dal event returns to the World Values Survey data; this Survey dataset comprises hun-dreds of questions on people's values and beliefs, and researchers can apply machine learning methods (e.g., least absolute shrinkage and selection operator, also known as LASSO) to identify the dimensions most relevant to scandal event returns and compile them into an ESG preference index for each country. There is also the more complex approach, namely, applying structural estimation by matching the reputation assets/investment (G), tangible assets (K), scandal probability (P), and firm market value times scandal return (Θ) data. By applying Eq. (9), we can then backout the function of

$\theta\left(\frac{G_t}{K_t}; \lambda\right)$. With additional function form assumption, researchers can estimate ESG preference λ for a country. It is also the case that this model was designed as a dynamic model, and it may thus be worth trying to replace λ to λ_t to see how ESG preference changes over time.

However, estimating a country's preference for ESG requires a consistent measure of its reputational assets and investments, which in turn requires the development of accounting standards that enable the accurate measuring of firm ESG investments and capital. By doing so, investors can value a firm in a fair manner.

Conclusion

In this study, we firstly theoretically and empirically demonstrate that a higher ESG score is associated with (1) lower ESG scandal probability and (2) higher losses given a scandal. These findings convey two key messages, as follows: (1) ESG scores are informative, as they can predict negative ESG events; (2) high ESG-rated firms should maintain elevated standards to avoid scandals, as the losses incurred otherwise will be more significant compared to low ESG-rated firms.

Further, we outline a model that provides insights into how firms optimize their ESG investments. Our model reveals that firms have distinct equilibria for ESG investment, primarily because of the trade-off between ESG practice cost adjustments and losses incurred from ESG scandals. For low ESG-rated firms, the costs of improving and maintaining high ESG standards is higher than the losses from ESG scandals, leading their optimal strategy to be the maintenance of a low ESG. Our findings also emphasize the significance of ESG preferences in the market and suggest potential research directions for studies on ESG preferences and optimal ESG investments.

Our evidence demonstrates that if the costs of adjusting ESG practices are too high, or if the market does not consider ESG factors important, firms may choose not to invest in ESG as the optimal solution. This makes it vital for governments aiming to promote a socially responsible commercial society to reduce ESG investment costs. Policies could include tax incentives for ESG investment, such as green investment (E), donations (S), and recruitment of sustainability auditors (G), and the promotion of high ESG firms. Moreover, citizen education on ESG issues holds potential to foster more socially-responsible behavior among firms.

In conclusion, this study highlights the critical connections between ESG scores, scandal probabilities, and scandal returns. Although our findings are descriptive, they underscore the potential for causal exploration using rigorous identification strategies. Future research could extend our model to incorporate additional dimensions, including the impact of capital structure on reputation investment decisions and the nuanced influence of ESG preferences. This study also delivers a roadmap for future investigations (see "Firm's optimal ESG investment" Section) and contributes to the ongoing discourse of the complex interplay between ESG considerations and financial outcomes.

Appendix A: Variable definition

This table describes the definition and data source of each variable we used in our analyses.

Variable	Definition	Source
Dependent Variables		
ESG Scandal	Dummy equals one if there is an ESG scandal event in the next twelve months	RepRisk
ESG Scandal Returns	Cumulative abnormal returns for ESG scandal events. The estimation window is $[-250, -20]$ trading days and the event window is $[-10, 10]$ trading days	CRSP, RepRisk
ESG Scores		
ASSET4	ASSET4 ESG score	ASSET4
S&P	S&P global ESG score	S&P Global
Sustainalytics	Sustainalytics ESG score	Sustainalytics
KLD	KLD ESG score	KLD
PCA	Principal component of ASSET4, S&P, Sustainalytics, and KLD	
ESG Disagree	The standard deviation of ASSET4, S&P, Sustainalytics, and KLD	
Average	Mean of ASSET4, S&P, Sustainalytics, and KLD	
Company Fundamentals		
Size	Natural logarithm of total assets plus one $size = \log(ACT) + 1$	Compustat
Cash Ratio[a]	Cash and short-term investments (CHE) divided by total assets (AT) $Cash \; ratio = \frac{CHE}{AT}$	Compustat
Current Ratio	Current assets (ACT) divided by current liabilities (LCT) $current \; ratio = \frac{ACT}{LCT}$	Compustat
Intangibility	Intangible assets (INTAN) divided by total assets (AT) $Intangibility = \frac{INTAN}{AT}$	Compustat
ROA	Earnings before interest (EBITDA) divided by total asset (AT) $ROA(Return \; on \; Assets) = \frac{EBITDA}{AT}$	Compustat
CAPEX	Capital expenditures (CAPX) divided by total asset (AT) $CAPEX(Capital \; Expenditure) = \frac{CAPX}{AT}$	Compustat
Maturing Debt	Long-term debt due in one year (DD1) divided by current asset (ACT) $Maturing \; Debt = \frac{DD1}{ACT}$	Compustat
Leverage	total debt including current (DT) divided by total assets (AT) $Leverage = \frac{DT}{AT}$	Compustat
Dividend Payout	Total dividends (DVT) divided by net income (NI) $Dividend \; Payout = \frac{DVT}{NI}$	Compustat
Sales Growth	The compound growth rate of last 5 years' sales $Sales \; Growth = \left(\frac{SALES}{SALES_{5 \; years \; ago}} \right)^{\frac{1}{5}} - 1$	Compustat
Cash Flow Vol	The standard deviation of last 5 years' Operating Activities Net Cash Flow (OANCF) $Cash \; Flow \; Volatility = \sigma(OANCF \; of \; last \; 5 \; years)$	Compustat

[a] Please note that CHE/LCT also makes sense since it captures the ability of a firm to use its cash to repay its current debt, which is a common practice in academia. CHE/AT provides useful information on cash holding since not all cash is prepared for repaying short-term debt. Cash can also be used for more general purposes like capital investment, financial policy flexibility, etc. This is also common in academia, e.g., Palazzo (2012), Cao et al. (2021), and Ding et al. (2021). We try using CHE/LCT to measure cash holding, and the result also holds

Appendix B. ESG scores

B.1: ESG scores history

ASSET4 was acquired by Thomson Reuters in 2009, but the ESG data was made available under the old name of ASSET4. After the acquisition, the name was changed to Thomson Reuters ESG Scores. However, since the name ASSET4 is widely known, we continue to use it for simplicity. It is important to note that as of 2018, the ESG ratings data of Thomson Reuters is now part of Refinitiv and is also known as Refinitiv ESG.

In 2017, Morningstar acquired approximately 40 percent stake in Sustainalytics, and later in 2020, they purchased the remaining approximate 60 percent of Sustainalytics equity.

The data from KLD originates from Kinder, Lydenberg, and Domini (KLD) Inc., which was acquired by Riskmetrics in 2009. In 2010, MSCI acquired Riskmetrics. For the purposes of this paper, we refer to this data as either KLD or MSCI KLD. Eccles et al. (2020) provide detailed information on the history of KLD.

B.2: Research on ESG score qualities

With the increasing availability of ESG ratings from different raters, the disagreement among ESG rating providers has become a concern for investors. As evidenced by Chatterji et al. (2016) and Berg et al. (2022), the disagreement among ESG raters is substantial, with the correlation between ESG ratings ranging from 0.38 to 0.71 as shown by Berg et al. (2022). Chatterji et al. (2016) first investigated the main drivers for the divergence, they provide two reasons for such divergence: "theorization" and "commensurability". They evidenced that both differences in theorization and commensurability will have an impact on ESG rating divergence. Berg et al. (2022) extended the study to analyze the extent of the impact of each source that drives the divergence. They distinct the ESG rating divergence into three sources: "Scope divergence" comes from the situation where ratings are calculated based on different sets of attributes. "Measurement divergence" emerges when different rating providers measure the same attribute by distinct indicators. "Weight divergence" appears when rating agencies give different importance to attributes. Berg et al. (2022) documented the possibility of estimating the aggregation rule used by ESG rating agencies and showed an accuracy of 79–99%. Berg et al. (2022) also ranked the contributions to ESG divergence of the above-mentioned three sources: the measurement divergence contributes the most (56%), scope divergence contributes the second (38%), and weight divergence is the last meaningful, contributing merely 6%. Berg et al. (2022) further show that the measurement divergence is in part driven by the "rater effect", also known as the "halo effect", where a high score in one category would induce high scores in other categories from the same rater, it showed that 15% variation of category scores can be explained by rater effect when controlling firm and category.

Capelle-Blancard and Petit (2019) found negative market reactions to negative ESG news. Serafeim and Yoon (2023) analyzed how ESG rater disagreement affects the predictive ability of ESG ratings. They documented that the predictive value of ESG consensus will be weakened by the presence of significant disagreement. By examining the three components of disagreement proposed by Berg et al. (2022), they find that the predictive

ability of consensus rating diminishes for firms with large measurement divergence, and such findings are not evident for scope divergence and weight divergence.

The ESG scores are still under development. The above-mentioned literature studies the quality of previous version of ESG scores. Whether these conclusions can be applied to the next generation of ESG scores is still not clear and calls for more research.

Appendix C: Correlation

This table presents the correlation between the explanatory variables and other control variables.

	Average ESG score	
	Scandal return analysis	Scandal probability analysis
ESG disagree	− 0.21***	0.28***
Size	− 0.15*	− 0.08
Cash ratio	0.12	0.01
Current ratio	0.09	0.01
Intangibility	0.04	− 0.01
ROA	− 0.06	− 0.01
CAPEX	− 0.04	0.01
Maturing debt	− 0.02	− 0.02
Leverage	− 0.03	− 0.01
Dividend Payout	0.01	0.02
Sales growth	− 0.08	− 0.07
Cash flow Vol	− 0.03	− 0.02
N	40,045	127,496

Abbreviations

ESG	Environmental, social, and governance
E	Environmental
S	Social
G	Governance
CSR	Corporate social responsibility
PCA	Principal component analysis
AT	Total assets
CHE	Cash and short-term investments
ACT	Current assets
LCT	Current liabilities
INTAN	Intangible assets
EBITDA	Earnings before interest
DD1	Long-term debt due in one year
DT	Total debt including current
SALES	Sales
OANCF	Operating activities net cash flow
CAPX	Capital expenditures
DVT	Total dividends
NI	Net income
ROA	Return on assets
CAR	Cumulative abnormal return
WVS	World Values Survey
LASSO	Least absolute shrinkage and selection operator

Sun *et al. Financial Innovation*   (2024) 10:121

Page 20 of 21

Author contributions
WS: Data preprocessing, data analysis, writing. YL: Data analysis, writing. SMY: Conceptual framework, writing. LY: Data download, data preprocessing. WD: Conceptual framework, theoretical modeling, data analysis, writing.

Funding
The study was funded by Hong Kong Polytechnic University under grand number P0047740.

Availability of data and materials
All the data we use is proprietary, so we cannot disclose it to the public. The codes will be provided upon request.

Declarations

Competing interests
Authors declares that they have no competing interest.

References
Agrawal A, González-Uribe J, Martínez-Correa J (2022) Measuring the ex-ante incentive effects of creditor control rights during bankruptcy reorganization. J Financ Econ 143:381–408. https://doi.org/10.1016/j.jfineco.2021.09.020

Anderson RW, Bustamante MC, Guibaud S, Zervos M (2018) Agency, firm growth, and managerial turnover. J Finance 73:419–464. https://doi.org/10.1111/jofi.12583

Avramov D, Cheng S, Lioui A, Tarelli A (2022) Sustainable investing with ESG rating uncertainty. J Financ Econ 145:642–664. https://doi.org/10.1016/j.jfineco.2021.09.009

Avramov D, Lioui A, Liu Y, Tarelli A (2021) Dynamic ESG equilibrium. Social Science Research Network, Rochester, NY

Bartov E, Marra A, Momenté F (2020) Corporate social responsibility and the market reaction to negative events: evidence from inadvertent and fraudulent restatement announcements. Account Rev 96:81–106. https://doi.org/10.2308/tar-2018-0281

Bauer R, Ruof T, Smeets P (2021) Get real! Individuals prefer more sustainable investments. Rev Financ Stud 34:3976–4043. https://doi.org/10.1093/rfs/hhab037

Berg F, Kölbel JF, Rigobon R (2022) Aggregate confusion: the divergence of ESG ratings. Rev Financ. https://doi.org/10.1093/rof/rfac033

Bint-Tariq M-N, Nobanee H (2020) Business ethics and finance: a mini-review

Bonnefon J-F, Landier A, Sastry PR, Thesmar D (2022) The Moral Preferences of Investors: Experimental evidence

Brown JR, Huang J (2020) All the president's friends: Political access and firm value. J Financ Econ. https://doi.org/10.1016/j.jfineco.2020.05.004

Brown SJ, Warner JB (1985) Using daily stock returns: the case of event studies. J Financ Econ 14:3–31. https://doi.org/10.1016/0304-405X(85)90042-X

Cao J, Han B, Zhan X, Tong Q (2021) Option return predictability. Rev Financ Stud. https://doi.org/10.1093/rfs/hhab067

Capelle-Blancard G, Petit A (2019) Every little helps? ESG news and stock market reaction. J Bus Ethics 157:543–565. https://doi.org/10.1007/s10551-017-3667-3

Chatterji AK, Durand R, Levine DI, Touboul S (2016) Do ratings of firms converge? Implications for managers, investors and strategy researchers. Strateg Manag J 37:1597–1614. https://doi.org/10.1002/smj.2407

Chen T, Dong H, Lin C (2020) Institutional shareholders and corporate social responsibility. J Financ Econ 135:483–504. https://doi.org/10.1016/j.jfineco.2019.06.007

Cheng B, Ioannou I, Serafeim G (2014) Corporate social responsibility and access to finance. Strateg Manag J 35:1–23. https://doi.org/10.1002/smj.2131

Christensen DM, Serafeim G, Sikochi A (2021) Why is corporate virtue in the eye of the beholder? The case of ESG ratings. Account Rev 97:147–175. https://doi.org/10.2308/TAR-2019-0506

Dai R, Liang H, Ng L (2021) Socially responsible corporate customers. J Financ Econ 142:598–626. https://doi.org/10.1016/j.jfineco.2020.01.003

Ding W, Levine R, Lin C, Xie W (2021) Corporate immunity to the COVID-19 pandemic. J Financ Econ 141:802–830. https://doi.org/10.1016/j.jfineco.2021.03.005

Ding W, Levine R, Lin C, Xie W (2022) Competition laws, ownership, and corporate social responsibility. J Int Bus Stud 53:1576–1602. https://doi.org/10.1057/s41267-022-00536-4

Dunbar CG, Li ZF, Shi Y (2021) Corporate social (ir)responsibility and firm risk: the role of corporate governance

Eccles RG, Lee L-E, Stroehle JC (2020) The social origins of ESG: an analysis of innovest and KLD. Organ Environ 33:575–596. https://doi.org/10.1177/1086026619888994

Ferrell A, Liang H, Renneboog L (2016) Socially responsible firms. J Financ Econ 122:585–606. https://doi.org/10.1016/j.jfineco.2015.12.003

Gibson R, Krueger P, Schmidt PS (2021) ESG rating disagreement and stock returns. Financ Anal J. https://doi.org/10.2139/ssrn.3433728

Haque A, Sarwar S (2013) Effect of fundamental and stock market variables on equity return in Pakistan. Sci Int 25:981–987

Hartzmark SM, Sussman AB (2019) Do investors value sustainability? A natural experiment examining ranking and fund flows. J Financ 74:2789–2837. https://doi.org/10.1111/jofi.12841

Hsu P-H, Li K, Tsou C-Y (2023) The pollution premium. J Financ. https://doi.org/10.1111/jofi.13217

Sun *et al. Financial Innovation*   (2024) 10:121

Page 21 of 21

Larcker DF, Watts EM (2020) Where's the greenium? J Account Econ 69:101312. https://doi.org/10.1016/j.jacceco.2020.101312

Liu LX, Whited TM, Zhang L (2009) Investment-based expected stock returns. J Political Econ 117:1105–1139. https://doi.org/10.1086/649760

Palazzo B (2012) Cash holdings, risk, and expected returns. J Financ Econ 104:162–185. https://doi.org/10.1016/j.jfineco.2011.12.009

Sarwar S, Hussan W, Malhi SN (2013) Empirical relation among fundamentals, uncertainty and investor sentiments evidence of karachi stock exchange. Int Rev Manag Bus Res 2:674–681

Serafeim G, Yoon A (2023) Stock price reactions to ESG news: the role of ESG ratings and disagreement. Rev Account Stud 28:1500–1530. https://doi.org/10.1007/s11142-022-09675-3

Yoon Y, Gürhan-Canli Z, Schwarz N (2006) The effect of corporate social responsibility (CSR) activities on companies with bad reputations. J Consum Psychol 16:377–390. https://doi.org/10.1207/s15327663jcp1604_9

Publisher's Note