Does investor communication improve corporate social responsibility? A machine learning-based textual analysis

Fan, Siyu; Kong, Dongmin; Lu, Jie; Yu, Honghai

Published in:
China Journal of Accounting Research

Published: 01/09/2024

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

Published version (DOI):
10.1016/j.cjar.2024.100370

Publication details:
Fan, S., Kong, D., Lu, J., & Yu, H. (2024). Does investor communication improve corporate social responsibility? A machine learning-based textual analysis. *China Journal of Accounting Research*, *17*(3), Article 100370. https://doi.org/10.1016/j.cjar.2024.100370

China Journal of Accounting Research 17 (2024) 100370

HOSTED BY

Contents lists available at ScienceDirect

China Journal of Accounting Research

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

Does investor communication improve corporate social responsibility? A machine learning-based textual analysis

Siyu Fan [a,b], Dongmin Kong [c], Jie Lu [a], Honghai Yu [a,*]

[a] School of Management and Engineering, Nanjing University, Nanjing, China
[b] Department of Accountancy, City University of Hong Kong, Hong Kong, China
[c] School of Economics, Huazhong University of Science and Technology, Wuhan, China

ARTICLE INFO

Article history:
Received 22 May 2023
Accepted 22 May 2024
Available online 13 June 2024

Keywords:
Corporate Social Responsibility
Institutional Investors
Site Visits
Textual Analysis

ABSTRACT

In this study, we take a machine learning-based approach to measure institutional investor attention to corporate social responsibility (CSR) issues when communicating with firms during site visits. We find that institutional investors can effectively enhance CSR performance through CSR-related communication. This effect remains robust to various checks and is more pronounced for non-state-owned enterprises and firms with lower levels of institutional ownership and in periods following the issuance of Green Investment Guidelines. We also identify information asymmetry and financing constraints as the two mechanisms underlying this effect. Overall, our findings highlight the importance of private interactions between management and institutional investors in promoting CSR.

1. Introduction

Socially responsible investing (SRI) has grown rapidly in recent years, with institutional investors increasingly incorporating environmental, social and governance (ESG) factors into their investment decisions. These investors now manage trillions of dollars in assets under sustainable investment strategies to better manage risks and generate long-term returns (Renneboog et al., 2008; Nofsinger et al., 2019; Hoepner et al., 2023).

* Corresponding author at: #22 Hankou Road, Nanjing, China, 210046.
E-mail address: hhyu@nju.edu.cn (H. Yu).

https://doi.org/10.1016/j.cjar.2024.100370
1755-3091/

As major providers of capital in financial markets, they have both the incentive and power to influence corporate social responsibility (CSR) activities[1] (Dyck et al., 2019; Kim et al., 2019; Chen et al., 2020; Liu et al., 2023). While many studies examine the mechanisms through which investors can engage with management and influence firms' CSR performance, such as shareholder proposals and voting (Dimson et al., 2015; Dyck et al., 2019; Kim et al., 2019; Dikolli et al., 2022; Hwang et al., 2022; Hoepner et al., 2023), few investigate the role of private meetings between managers and institutional investors. Survey evidence suggests that institutional investors actively engage in private communications with firms to express their concerns about the environment, labor relations, human rights and other topics related to social performance (Hockerts and Moir, 2004; Ziek, 2009; Du et al., 2010). In April 2022, the China Securities Regulatory Commission revised the Guidelines for Investor Relations Management of Listed Companies for the first time, requiring sustainable information to be included in communications between listed companies and investors. Anecdotal evidence also reveals that Chinese investors increasingly consider corporate sustainability and any relevant information, as this can be essential when they interact with firms.[2] However, private communication, as a channel through which institutional investors can influence CSR performance, remains understudied. We use a unique Chinese corporate site visit dataset to examine whether investors can effectively enhance CSR performance through direct interactions.

Previous studies show that institutional investors play an important role in shaping CSR. Although numerous studies explore the relationship between institutional ownership and CSR performance based on total institutional shareholder holdings, the empirical findings remain mixed (Borghesi et al., 2014; Fernando et al., 2017; Dyck et al., 2019; Chen et al., 2020; Hwang et al., 2022; Kim and Yoon, 2022), partly because institutional investors do not monitor all of their holdings equally. This variation may thus not accurately reflect institutional investors' monitoring behavior regarding CSR issues, as their preferences, objectives and strategies will differ (Gloßner, 2019; Nofsinger et al., 2019). Other studies mainly focus on shareholder activism or voting on shareholder proposals and present direct evidence of institutional investors' influence (Dimson et al., 2015; Barko et al., 2022; Hoepner et al., 2023). However, shareholder activism is an infrequent form of costly intervention and often only involves (and therefore can only influence) a few firms (Chapman et al., 2022). The generalizability of these findings may then be limited, warranting cautious interpretation.

Our focus in this study is on private interactions between institutional investors and firms, thus filling an important gap in the literature. Institutional investors can enhance the transparency of CSR policies, risks and impacts by engaging with firms on CSR issues, thus enabling better monitoring and advocacy for improvement (Hockerts and Moir, 2004; Jiang et al., 2022). Communication also helps to reduce the uncertainty around CSR practices, thus reducing investors' perceptions of risk and firms' capital costs. This provides firms with more resources to invest in CSR (Lins et al., 2017; Chapman et al., 2022). However, communication can serve investors' interests in terms of influencing firms, or may lack constraining power compared with activist tactics such as shareholder proposals (Chapman et al., 2022; Heath et al., 2023). Thus, whether communications lead to CSR improvements remains debatable.

Using the context of site visits to examine institutional investors' CSR-related communications offers several advantages. First, site visit transcripts capture real-time interactions between company managers and investors and avoid post-editing or embellishment, thus reflecting institutional investors' authentic concerns over CSR issues. Second, compared with regular investor communication (e.g., earnings conference calls), site visits can be initiated whenever investors deem them necessary.[3] The transcripts can therefore promptly capture investors' concerns regarding CSR issues. Third, unlike interactive platforms aimed at retail investors

[1] According to the Global Sustainable Investment Review (2022), $30.3 trillion is invested globally in sustainable investing assets. The market size for major types of responsible investments in China was approximately CNY 31.59 trillion.

[2] According to the manager of the Investor Service Department at the Shanghai Stock Exchange, "CSR has emerged as one of the critical areas of focus for investors when evaluating companies." (News source: https://www.amac.org.cn/hdjl/esgtzlt/2020lt/2020ltzjgd/202101/t20210125_24199.html).

[3] According to the "Investor Relations Management Guidelines" published by the Shenzhen Stock Exchange, listed companies typically do not decline site visit requests from investors and only advise rescheduling if the timing coincides with sensitive periods (e.g., before profit announcements or other significant corporate event disclosures).

(Lee and Zhong, 2022), site visits primarily involve expert analysts and institutional investors, whose expertise, research resources and market influence better position them to analyze CSR performance and drive changes.

We introduce a machine learning-based approach to capture CSR-related communication during site visits. First, we use *word2vec*, a neural network-based word embedding model, to identify words that are semantically related to our selected CSR seed words, and construct a CSR lexicon tailored to Chinese-listed firms. Next, we perform a content analysis on the discussions between managers and institutional investors, utilizing our CSR lexicon to identify CSR-related communication in the site visit transcripts, and construct two key independent variables. The first variable computes the cosine similarity between two vectors: one representing the transcript text and the other the CSR lexicon. This provides a quantifiable measure of the semantic alignment between the transcripts and CSR vocabulary. The second variable evaluates the relative frequency of CSR keywords in the transcripts by normalizing the total keyword count by the total word count. Higher values for these two variables indicate institutional investors' increased emphasis on CSR issues during site visits.

Using Chinese A-share firms listed in the Shenzhen Stock Exchange (SZSE) from 2013 to 2021, we first explore whether institutional investors can effectively enhance CSR performance through CSR-related communication. We find that this communication significantly enhances firms' subsequent CSR performance. A one-standard-deviation increase in CSR-related communication results in an approximate 2 % improvement in CSR performance in the following year. We also find positive and significant correlations between five CSR communication sub-dimensions and future CSR performance. Our results indicate that site visits represent an interactive communication channel through which institutional investors influence CSR. The results are robust to several checks, including using alternative measures of CSR performance and CSR-related communication, controlling for firm fixed effects, using alternative model specifications and accounting for possible endogeneity issues.

We then investigate cross-sectional variation in how the effect of CSR communication on CSR performance varies with state ownership, institutional shareholdings, and regulatory environment. We find that the effect is more pronounced for non-state-owned enterprises and firms with lower institutional shareholdings in periods following the issuance of the Green Investment Guidelines. Finally, we identify two channels through which institutional investors' CSR communication can improve firms' CSR performance: enhancement of the information environment and alleviation of financial constraints.

This study makes several contributions to the literature. First, it provides insights into the effect institutional investors can have on CSR performance. Whether institutional investors can truly improve corporate sustainability remains a matter of debate. Some argue that institutional investors can improve CSR through engagement and monitoring (Dimson et al., 2015; Gloßner, 2019), but others suggest that they may lack the motivation or capability to affect CSR change across large portfolios or consider short-term financial priorities (Gloßner, 2019). "Greenwashing," a situation in which there is no actual ESG impact, may also be an issue (Gibson Brandon et al., 2022; Heath et al., 2023). Our study provides evidence of the positive influence of institutional investors on CSR performance. In addition, most studies focus on ownership, shareholder proposals and voting, and only a few report that site visits can affect CSR performance, despite the prevalence and importance of private interactions between investors and managers. They also examine the impact of site visits or their frequency on CSR ratings without considering the communication content (Hu et al., 2020; Jiang et al., 2022). As investors' horizons and preferences differ (Gloßner, 2019; Hwang et al., 2022), the frequency of their site visits may not accurately capture the levels of CSR monitoring they engage in. We address this limitation by conducting a content analysis of CSR-related communication based on site visit transcripts, thus providing richer insights into the channel through which institutional investors influence CSR performance.

Second, our study enriches the literature on the economic consequences of investor–manager communication. Investors can communicate with public firms through various channels such as earnings calls, capital market conferences, non-deal roadshows and site visits. Prior studies indicate that these interactions provide information advantages (Cheng et al., 2016; Liu et al., 2017; Bushee et al., 2018; Chapman et al., 2022; Rennekamp et al., 2022), reduce information asymmetry (Brown et al., 2004; Jiang and Yuan, 2018) and facilitate the scrutiny and monitoring of firms (Reiter, 2021), but they do not explore the impact of such communications on CSR. Our study is the first to provide empirical evidence that investor–manager communications significantly influence firms' CSR performance specifically through site visits.

Finally, our study makes important methodological contributions to analyses of investor engagement in CSR. We innovatively apply a machine learning-based textual analysis approach, including seeded word embedding and bag-of-words-based content analysis techniques, to quantify CSR-related communication during site visits. We develop a specific Chinese CSR lexicon, which incorporates both common CSR terms and China-specific CSR vocabulary and concepts. This enables us to more effectively analyze investor engagement in CSR issues in China's unique institutional context (Shen et al., 2023).

2. Literature review and hypothesis development

2.1. Institutional investors' corporate site visits

Site visits are important channels through which both shareholding and non-shareholding institutional investors can privately interact with firms (Soltes, 2014; Cheng et al., 2016; Bowen et al., 2018). However, these private meetings are generally unobservable, so information regarding such meetings in U.S. public firms is limited. By contrast, the mandatory disclosure regulations introduced in China by the SZSE in 2009 mean that investor relationship management reports are available and private meetings observable. We can therefore examine these previously unobservable activities. Through site visits, institutional investors can observe firms' operating and production activities first-hand and have the opportunity to engage in face-to-face discussions with managers (Cheng et al., 2016). These private interactions facilitate institutional investors' information acquisition and provide them an informational advantage, that enables them to make better investment decisions and more accurate forecasts (Cheng et al., 2016; Liu et al., 2017; Han et al., 2018; Hong et al., 2019). The information gained from site visits is eventually conveyed to the market and incorporated into stock prices (Cheng et al., 2018). Site visits also enable institutional investors to more effectively monitor managers, thus helping to mitigate their myopic decisions (Jiang et al., 2022). Although some studies suggest that site visits can affect firms' activities related to social responsibility (Hu et al., 2020), they only examine whether or not institutional investors conduct site visits or the frequency of their visits, without considering the communication content. We analyze the CSR-related communications in site visit transcripts to address this limitation.

2.2. Institutional investors and CSR performance

Institutional investors are often known as "universal owners" due to their large, diversified and long-term equity holdings. They play an important external governance role as they can influence management via "voice" and "the threat of exit" (Hirschman, 1970; Gillan and Starks, 2000; Chen et al., 2007; McCahery et al., 2016). Due to their importance in financial markets, a substantial body of literature has examined whether and how institutional investors affect corporate social performance. Some studies indicate that institutional shareholders can help to improve the social impact of their portfolio firms (Dyck et al., 2019; Chen et al., 2020). As major equity owners, they can directly engage with firms to address environmental protection, employee rights and other social issues (Dimson et al., 2015; Barko et al., 2022; Hwang et al., 2022). Through their monitoring role, they can also ensure that managers pursue CSR strategies in the shareholders' interests (Gloßner, 2019). Numerous studies indicate that active CSR engagement allows investors to reduce the risk of costly incidents (Nofsinger and Varma, 2014; Nofsinger et al., 2019; Hoepner et al., 2023), generate social benefits (Kim et al., 2019) and ultimately enhance firm value (Gloßner, 2019).

However, institutional investors may not attempt to change firms' CSR practices and may even have negative effects. They may lack the ability or resources to effectively monitor CSR practices across all of their portfolio firms, and the costs of extensive monitoring may outweigh the expected benefits. Moreover, some institutional investors may also have pecuniary motives for prioritizing financial performance over CSR. For example, Gloßner (2019) finds that short-term investors induce managerial short-termism, in which CSR spending is reduced to increase short-term profits. Furthermore, socially responsible investment (SRI) funds themselves may engage in "greenwashing" or "impact washing," limiting their influence on firms'

CSR practices. Empirical evidence suggests that SRI funds attract capital but do not exhibit better ESG performance than traditional funds (Gibson Brandon et al., 2022; Kim and Yoon, 2022), or they simply select firms that already have high environmental and social performance without significantly changing firm behavior (Heath et al., 2023).

Thus, it remains an open question whether institutional investors successfully improve the CSR performance of their portfolio firms. Investors' preferences and investment horizons regarding CSR will also differ. Some may aim to maximize the financial performance of their portfolios, while others may have broader objectives that encompass social responsibility. This heterogeneity among institutional investors can determine the extent to which they actively pursue and achieve CSR improvements across the firms they have invested in.

2.3. Hypothesis development

We propose that institutional investors have the potential to influence firms' CSR performance through CSR-related communication. First, through effective communication, these investors can make CSR information more transparent, which then improves the monitoring of management and leads to CSR performance improvements. Institutional investors can obtain detailed, timely and accurate information about firms' CSR policies, impacts, risks and opportunities by engaging with them, and thus can confidently advocate for improvements (Hockerts and Moir, 2004). For example, conversations with executives can inform assessments of how to implement CSR strategies, while conversations with operations managers can provide insights into resource usage and waste. Discussions with staff can elicit their perspectives on working conditions and employee safety. Such increased transparency enables institutional investors to better fulfill their monitoring role and thus pushes managers to improve their sustainability performance (Jiang et al., 2022). Issues highlighted by institutional investors also gain attention from the public and regulators, increasing the scrutiny firms face regarding their social responsibility activities. Institutional investors' engagement in CSR also provides a signal that such issues are material to long-term value creation, which incentivizes firms to improve their CSR practices (Chen et al., 2020; Barko et al., 2022). In summary, CSR-related communication promotes transparency, which helps to improve corporate sustainability.

Second, communication can reduce investors' uncertainty about CSR practices, thus reducing the perceived investment risks. Firms will then face fewer financing constraints and can direct more resources toward CSR initiatives. Chapman et al. (2022) find that direct and ongoing interactions can help increase investors' understanding of a firm's strategy and build mutual trust, resulting in greater alignment with and support from management. Firms regarded as more trustworthy by investors may also receive valuation premiums from them (Guiso et al., 2008; Lins et al., 2017). Through proactive CSR engagement, investors can gain more clarity about firms' policies, impacts and exposure to risk. If lower risk premiums are factored into capital costs, firms will face fewer restrictions when accessing affordable capital from investors and lenders (Dhaliwal et al., 2011; Lins et al., 2017). This reduction in financing constraints will provide firms with more resources and the flexibility to further invest in CSR initiatives.

Nevertheless, whether institutional investors can truly improve corporate sustainability through CSR-related communication remains debatable. First, they may engage in such communication primarily because their investments and performance metrics involve social responsibility, rather than to influence corporate practices (Heath et al., 2023). Second, compared with litigation and shareholder proposals, communication represents a relatively mild form of intervention that may lack management constraining power (Chapman et al., 2022). Executives may simply react with superficial responses or temporary actions to satisfy investors' transient CSR concerns.

In summary, the net effect of institutional investors' CSR-related communication on CSR performance is an empirical question. Therefore, we propose the following null hypothesis:

H1: Institutional investors' CSR-related communication with firms has no effect on the CSR performance of these firms.

3. Data and methodology

3.1. Sample and data sources

We first assess firms listed on the SZSE between 2013 and 2021 and select our sample according to the following criteria.[4] We exclude (1) financial firms and B-share (foreign share) firms; (2) site visits involving non-institutional investors[5]; (3) transcripts in which any questions or answers do not consist of at least two tokens; and (4) firm-year observations missing data for the variables used in our analysis. Our final sample consists of 7,781 firm-year observations. To alleviate the potential influence of extreme observations, all continuous variables are winsorized at the 1 % level in each tail.

The transcripts of site visits for Chinese listed firms and firm characteristics are collected from the China Stock Market and Accounting Research (CSMAR) database. Corporate social responsibility data are retrieved from Huazheng ESG ratings, accessed through the WIND database. This leading Chinese ESG rating agency has several advantages over other mainstream sustainability rating frameworks. First, it provides comprehensive ratings dating back to 2009 for all A-share listed firms. Second, it ensures the accuracy of these ratings by issuing quarterly updates and timely adjustments in response to major ESG incidents. Consequently, the Huazheng rating score is applied extensively in studies of ESG performance (see, e.g., Lin et al., 2021; Jiang et al., 2022). We manually collected the signatories of the Chinese Principles of Responsible Investment (PRI) from the online United Nations PRI signatory directory, and matched them with institutional investor names in the site visit transcripts using fuzzy matching.[6]

3.2. Measuring CSR-related communication

We introduce a machine learning-based approach using *word2vec*, a neural network-based word embedding model, to identify words semantically related to CSR seed words and construct a CSR lexicon.[7] We then use this lexicon to quantify institutional investors' considerations of CSR issues when communicating with firms.

First, following the literature (Li et al., 2021; Wu, 2023), we initiated the construction of a dictionary by identifying relevant seed words. The Company Law of 2006 requires Chinese companies to undertake social responsibilities in their business activities, and in 2008 the SZSE mandated companies in the Shenzhen 100 Index to release CSR reports. These should include at least the following dimensions: (1) protection of the interests of shareholders and creditors; (2) protection of workers' rights; (3) protection of suppliers, customers and consumers; (4) environmental protection and sustainable development; and (5) public relations and social welfare services. We selected seed words for our initial CSR dictionary construction based on these five dimensions, as our sample consisted of SZSE-listed companies. We reviewed 100 randomly selected CSR reports from Shenzhen 100 Index companies and took an independent extraction and cross-validation approach to ensure seed word quality. From this, we compiled a seed word repository containing 563 words, with 108 related to "shareholders and creditors," 119 to "employees", 112 to "suppliers, customers and consumers", 118 to "environmental protection and sustainability" and 106 related to "public relations and social welfare." Table IA1 in the presents the list of selected seed words in Chinese across all CSR dimensions.

[4] The sample period begins in 2013 because the SZSE updated the requirements on the timeliness of the information disclosure related to site visits of listed companies in 2012. Since 2013, all listed companies have standardized the disclosure of investors' site visit activities following the requirements of the SZSE. As our model uses lagged CSR communication data, the CSR performance data span 2014–2021, while the CSR communication and control variables span 2013–2020.

[5] Site visits include various visitors such as individual investors, institutional investors and other market participants. Among them, institutional investors are the primary participants. The literature shows that the participation of non-institutional investors, such as media, in site visits could also have an impact on corporate social responsibility. To avoid confounding our results, we exclude site visits involving non-institutional investors.

[6] Data source: https://www.unpri.org/signatories/signatory-resources/signatory-directory.

[7] For lack of space, we are unable to report all details of our textual analysis in this section. Technical details of the natural language processing pipeline and machine learning methods are provided in the online.

Next, we scraped CSR reports released on cninfo.com[8] from 2006 to 2020 and preprocessed the raw CSR report text to train the embedding model. As these reports focus specifically on CSR issues, they are likely to contain more domain-specific words around the relevant CSR concepts than are more general texts, thus providing more finely tuned and precise word vector representations than models trained on generic corpora. Preprocessing involved text tokenization, named entity recognition, stopword removal and phrase extraction to prepare the data for effective model training. We applied the same preprocessing procedures used for the CSR reports to the site visit transcripts we used for constructing our text-based measures to ensure consistency in the text processing.

We then used the textual corpus of these CSR reports to train a word2vec model, a widely used embedding algorithm that converts words into word vectors (Mikolov et al., 2013), to capture semantic similarities between words based on their co-occurrence patterns. word2vec is based on a neural network model that builds a vocabulary from the training corpus and learns vector representations of words. It trains word vectors using either the continuous bag-of-words or skip-gram architecture. Both methods work to capture semantic similarity through vector cosine distance, and words with similar meanings will have closer vectors. We divided the preprocessed CSR reports into three training rounds to feed into the word2vec model to generate word embeddings. The model maps the tokens to 300-dimensional vectors, constructing a vector space where each unique token is assigned a specific vector representation. Words with similar semantic meanings have a smaller cosine distance between them in this vector space. Using our selected seed words, we first constructed a CSR dictionary by searching for semantically similar words. We used the average word vector of all seed words within the same dimension to represent that dimension. We then identified semantically similar words by calculating the cosine similarity between this averaged dimension vector and the vectors of other words. We then expanded the CSR lexicon by incrementally adding synonyms in three rounds of training. In each iteration, based on cosine similarity, words most similar to the seed word vectors were selected to expand the dictionary for each dimension. Finally, after manually checking the output dictionaries, we constructed a CSR lexicon containing 3,879 words. Table IA3 in the illustrates the expansion of the dictionary in each round and the number of words in each dimension.

Using the constructed CSR lexicon, we performed a content analysis of dialogues between managers and institutional investors to quantify the relevant CSR-related communication in the site visit transcripts, and then aggregated these measures at the firm-year level. This involved creating two independent variables to measure institutional investor attention to CSR issues when communicating with firms during site visits.

First, CSR_cosine_k is the cosine similarity between the TF-IDF weighted[9] word vectors that represent the site visit transcripts and the CSR dictionary, formulated as in Eq. (1).

$$CSR\_cosine_k = \cos\left(V_k, V_{CSR}\right) \tag{1}$$

where V_k is the TF-IDF weighted word vector of the k^{th} transcript and V_{CSR} is the CSR dictionary vector. For expositional purposes, we multiplied all variables based on cosine similarity by 100.

Second, CSR_freq_k measures the relative frequency of the TF-IDF weighted CSR dictionary words appearing in transcript k. This is calculated as the weighted count of CSR words divided by the total number of words in the transcript, formulated as in Eq. (2).

$$CSR\_freq_k = \frac{\sum_{j=1}^{N} I\left(w_j \in W_{\mathrm{CSR}}\right) \cdot \mathrm{tfidf}\left(w_j, k\right)}{\sum_{i=1}^{M} \mathrm{tfidf}\left(w_i, k\right)} \tag{2}$$

where w_j is a word appearing in the transcript,W_{CSR} is the CSR dictionary, $I(\cdot)$ is an indicator function that equals 1 if w_j is in W_{CSR} and 0 otherwise, $tfidf\left(w_j, k\right)$ is the TF-IDF weight of w_j in the k^{th} transcript and M is the total number of words in the k^{th} transcript.

[8] The official website designated by the China Securities Regulatory Commission for public companies disclosure.

[9] The TF-IDF weighting technique aims to highlight words that are frequent within a specific document but relatively rare across the entire corpus. By multiplying the term's frequency (TF) by its inverse document frequency (IDF), TF-IDF assigns higher weights to terms that are both prominent in the current document and distinctive across the broader dataset.

We aggregated these two measures to the firm-year level using Eqs. (3) and (4), where $N_{i,t}$ is the number of meeting transcripts for company i in year t. By construction, the $CSR\_cosine_{i,t}$ and $CSR\_freq_{i,t}$ measures quantify the extent to which institutional investors consider CSR issues. Increased values for these two independent variables indicate a higher level of consideration of CSR issues during site visits.

$$CSR\_cosine_{i,t} = \sum_{k=1}^{N_{i,t}} CSR\_cosine_k \tag{3}$$

$$CSR\_freq_{i,t} = \sum_{k=1}^{N_{i,t}} CSR\_freq_k \tag{4}$$

In addition, as CSR is a multi-dimensional concept encompassing many issues, we created variables to measure institutional investor attention toward each specific CSR dimension in communications. We calculated both the cosine similarity $CSR\_cosine_{d,i,t}$ and the relative frequency $CSR\_freq_{d,i,t}$ for each CSR dimension d. These are formulated similarly to Eqs. (1)–(4) but use the words of the sub-dimensions instead of the full CSR lexicon. We labeled these dimension-specific CSR communication variables with suffixes indicating each dimension: "share" for shareholder and creditor interests; "emp" for employees; "ssc" for suppliers, customers and consumers; "env" for environmental protection and sustainability; and "social" for public relations and social welfare services. Higher values of these variable indicate greater attention to issues within the specific dimension.

To assess the robustness of our main results, we constructed several alternative measures. First, we created equally weighted versions without TF-IDF weighting to evaluate whether term importance affects the measures. Second, we restricted our identification to only the question portion of the transcripts rather than the full questions and answers, as this provided us with institutional investors' initiations. Third, we calculated the average of the TF-IDF weighted CSR communication across all transcripts of firms i in year t. Together, these variants enabled us to evaluate whether our main CSR communication measures are sensitive to the weighting method, text section or aggregation level. All of the variables are defined in Appendix A.

3.3. Measuring CSR performance

Huazheng ESG ratings are updated and released quarterly (at the end of January, April, July and October each year). The evaluated firms are assigned one of nine rating grades from "AAA" to "C." We quantified the rating levels by assigning 100 points to an AAA grade with a decrement of 10 points for each lower grade, down to 20 points for a C grade. To calculate the CSR score for each firm-year, we took the average of the four quarterly rating scores:

$$Score_{i,t} = \frac{1}{4}\sum_{q=1}^{4} RawScore_{i,q,t} \tag{5}$$

where $RawScore_{i,q,t}$ is company i's rating score for the q quarter of year t.

3.4. Control variables

Following the literature (Dyck et al., 2019; Gloßner, 2019; Chen et al., 2020), we included numerous firm-level variables to control for factors that may affect CSR performance: firm size (*Size*), firm age (*Age*), asset tangibility (*Tangible*), leverage ratio (*Lev*), return on equity (*ROE*), Tobins' Q(*Tobinq*), institutional ownership (*Inshold*), state ownership (*SOE*), board independence (*Indp*) and CEO duality (*Dual*). We provide detailed definitions in Appendix A.

3.5. Empirical model

We estimated the following baseline regression model to examine whether investors can effectively enhance CSR performance through CSR-related communication:

$$Score_{i,t+1} = \beta_0 + \beta_1 * CSR\_Communication_{i,t} + Controls_{i,t} + IndustryFE + YearFE + \varepsilon_{i,t} \tag{6}$$

Table 1
Summary Statistics.

Panel A CSR-related communication variables

	Obs	P25	Mean	Median	P75	SD
CSR_cosine	7781	1.5304	5.5445	3.3293	7.0936	6.2028
cosine_share	7781	0.5638	2.5413	1.4277	3.2306	3.1284
cosine_emp	7781	0.0903	0.8986	0.3726	1.0776	1.3663
cosine_scc	7781	1.2898	5.6460	3.0327	6.8812	7.2361
cosine_env	7781	0.1633	2.7042	0.7027	2.7342	5.0835
cosine_social	7781	0.0000	0.5875	0.1877	0.6831	1.0323
CSR_freq	7781	1.9653	7.0417	4.2642	9.0043	7.7974
freq_share	7781	0.3068	1.2392	0.7182	1.5728	1.4599
freq_emp	7781	0.0823	0.5045	0.2438	0.6158	0.7032
freq_scc	7781	0.8761	3.6156	1.9836	4.4624	4.5119
freq_env	7781	0.1187	1.3732	0.4207	1.4184	2.4384
freq_social	7781	0.0000	0.2103	0.0781	0.2479	0.3567

Panel B Other Variables

	Obs	P25	Mean	Median	P75	SD
Score	7781	45.0000	50.9857	50.0000	60.0000	9.9820
Size	7781	21.3084	22.0905	21.9659	22.7219	1.0991
Age	7781	1.6094	1.9895	1.9459	2.3979	0.6745
Tangible	7781	0.8849	0.9072	0.9459	0.9723	0.1022
Lev	7781	0.2321	0.3862	0.3775	0.5243	0.1881
ROE	7781	0.0367	0.0700	0.0731	0.1170	0.1030
Tobinq	7781	1.3972	2.2526	1.8502	2.6495	1.3208
Inshold(%)	7781	14.5509	37.7714	38.0032	58.5055	24.6086
SOE	7781	0.0000	0.2109	0.0000	0.0000	0.4080
Dual	7781	0.0000	0.3380	0.0000	1.0000	0.4731
Indp(%)	7781	33.3300	37.6471	33.3300	42.8600	5.3695

This table presents summary statistics for the sample of 7,781 firm-year observations over 2013–2021. Panel A reports descriptive statistics for the CSR communication measures. Panel B shows descriptive statistics for the dependent variable and control variables. All continuous variables are winsorized at the 1st and 99th percentiles. Variable definitions are provided in Appendix A.

where the dependent variable $Score_{i,t+1}$ represents firm i's annual average CSR rating score in the following year. $CSR\_Communication_{i,t}$ refers to a series of text-based measures of CSR-related communication for firm i in year t. *Controls* are a set of firm-specific characteristics. We included industry- and year-fixed effects to absorb time-invariant industry differences and common time trends, respectively. The coefficient estimate of β_1 captures the impact of CSR-related communication on firms' subsequent CSR performance, when controlling for other factors. A positive and statistically significant β_1 would suggest that increased CSR-related communication is associated with improvements in future CSR performance. The model is estimated using OLS with standard errors clustered at the firm level.

4. Empirical results

4.1. Descriptive statistics

Table 1 provides the descriptive statistics for the CSR-related communication measures and fundamental firm characteristics included in our regression models. Panel A demonstrates that the median values of our CSR communication metrics are generally lower than the mean, suggesting a right-skewed distribution. This indicates that a majority of firms exhibit limited CSR-related communication. The mean values for the independent variables *CSR_cosine* and *CSR_freq* are 5.5445 and 7.0417, respectively. By comparing the statistics

Table 2
Institutional Invests' CSR Communication and CSR Performance: Baseline Results.

	Depend Variable = $Score_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
CSR_cosine	0.232***	0.239***	0.155***			
	(0.025)	(0.026)	(0.026)			
CSR_freq				0.183***	0.188***	0.118***
				(0.020)	(0.020)	(0.020)
Size			2.057***			2.063***
			(0.271)			(0.271)
Age			−0.993***			−0.992***
			(0.334)			(0.334)
Tangible			8.506***			8.481***
			(1.631)			(1.632)
Lev			−11.060***			−11.071***
			(1.165)			(1.165)
ROE			21.437***			21.418***
			(1.563)			(1.566)
Tobinq			−0.371***			−0.374***
			(0.136)			(0.137)
Inshold			−0.028***			−0.028***
			(0.009)			(0.009)
SOE			2.530***			2.535***
			(0.566)			(0.566)
Dual			−0.521			−0.534
			(0.366)			(0.367)
Indp			0.154***			0.153***
			(0.031)			(0.031)
Constant	49.700***	46.616***	−5.506	49.694***	46.809***	−5.418
	(0.251)	(2.137)	(6.279)	(0.252)	(2.150)	(6.289)
Industry Fixed	No	Yes	Yes	No	Yes	Yes
Year Fixed	No	Yes	Yes	No	Yes	Yes
Observations	7781	7781	7781	7781	7781	7781
Adjusted R^2	0.021	0.058	0.178	0.020	0.057	0.177

This table reports the estimation results for the effect of institutional investors' CSR communication on firms' subsequent CSR performance. Columns (1)-(3) present results using the CSR communication variable based on cosine similarity. Columns (4)-(6) show results based on the relative frequency of CSR keywords. The dependent variable is the firms' CSR rating score in year $t + 1$, with all independent and control variables measured in year t. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10 %, 5 %, and 1 % levels, respectively.

across the specific dimensions, we find that "public relations and social welfare services" has the lowest mean values of 0.59 for *cosine_social* and 0.21 for *freq_social*.[10] In Panel B, the average CSR rating score is 51 out of 100, and other firm-level control variables are consistent with the samples of SZSE A-share companies used in other studies.

4.2. Baseline multivariate regression results

Table 2 reports the main results of our regressions. Columns (1) to (3) show that the independent variable is CSR-related communication calculated using the cosine similarity method. Columns (4) to (6) show CSR-related communication calculated by the relative frequency method. Columns (1) and (4) show a significant correlation between the independent and dependent variables when no control variables are added. Fixed effects are added in Columns (2) and (5), and the estimates persist with statistical significance. Control variables and controls for industry- and year-fixed effects are added in Columns (3) and (6). The coefficients of

[10] In the descriptive statistics of Table 1, Panel A, the sum of the relative frequency means of the five dimensions should theoretically equal the mean of *csr_freq*. However, as these variables are winsorized separately, a slight discrepancy arises.

CSR_cosine and *CSR_freq* are positive and significant at the 1 % level, suggesting that institutional investors can improve CSR performance through CSR-related communication during site visits. This effect has both statistical and economic significance. The coefficient estimates of 0.155 in Column (3) and 0.118 in Column (6) suggest that a one-standard-deviation increase in CSR-related communication leads to an improvement in CSR performance of 1.9 % (=0.155*6.20/50.99) in the following year and 1.8 % (=0.118*7.79/50.99) if all other variables are controlled. In summary, the baseline model results show that CSR-related communication during site visits can lead to subsequent improvements in firms' CSR performance, and thus our null hypothesis is rejected.

Table 3
Dimension-specific CSR Communication and CSR Performance.

Panel A Cosine Similarity Communication Measures

	Depend Variable =$Score_{t+1}$				
	(1)	(2)	(3)	(4)	(5)
cosine_share	0.145***				
	(0.049)				
cosine_emp		0.341***			
		(0.101)			
cosine_scc			0.099***		
			(0.022)		
cosine_env				0.187***	
				(0.034)	
cosine_social					0.621***
					(0.133)
Controls	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781	7781
Adjusted R^2	0.171	0.171	0.174	0.178	0.173

Panel B Relative Frequency Communication Measures

	Depend Variable = $Score_{t+1}$				
	(1)	(2)	(3)	(4)	(5)
freq_share	0.351***				
	(0.106)				
freq_emp		0.769***			
		(0.201)			
freq_scc			0.163***		
			(0.035)		
freq_env				0.411***	
				(0.069)	
frq_social					2.016***
					(0.388)
Controls	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781	7781
Adjusted R^2	0.172	0.172	0.174	0.178	0.174

This table presents the estimation results for the effects of dimension-specific CSR communication on firms' overall CSR performance. Panel A reports findings using the cosine similarity communication variables, while Panel B shows analyses based on relative frequency measures. All models include control variables, industry fixed effects, and year fixed effects. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

4.3. Dimension-specific CSR communication

We further examined the effect of sub-dimension CSR communication on firms' CSR performance. Panel A of Table 3 reports the results of estimating model (6) using the cosine similarity measures for CSR communication related to each dimension. We find positive and significant coefficients on all CSR dimension variables. Increased discussions of shareholder and creditors' rights (*cosine_share*), employee relations (*cosine_emp*), supplier/consumer/product issues (*cosine_scc*), environmental protection (*cosine_env*) and public welfare services (*cosine_social*) during site visits lead to improvements in firms' CSR scores. The relative frequency measures in Panel B produce consistent results. In summary, more dimension-specific CSR communication during site visits is related to better CSR performance. This illustrates that site visits can be a useful channel for various forms of CSR-related communication and suggests that institutional investors can effectively engage firms across various aspects of CSR through direct interaction and communication during site visits.

5. Robustness tests

We performed the following robustness tests to further verify the reliability of our main results.

5.1. Alternative CSR performance measure

Firms' CSR performance may be partly related to resource availability and thus to firm size. We addressed this following Hwang et al. (2022) and constructed an alternative CSR performance measure, *adjustedScore*, by removing size effects from the raw CSR scores. We subtracted the mean score of firms in the same total asset quintile. As shown in Table 4, our main conclusions remain robust when using this size-adjusted dependent variable.

5.2. Alternative CSR communication proxies

We assessed the robustness of our baseline results using three alternative proxies for institutional investors' CSR communication.

First, we constructed equal-weighted versions of our key variables *CSR_cosine* and *CSR_freq*. Compared with the baseline TF-IDF weighted measures, these equal-weighted variants exclude any effects of word importance, thus providing a sensitivity test to the weighting scheme. As Columns (1) and (2) of Table 5 show,

Table 4
Robustness Checks: An Alternative Proxy for CSR Performance.

	Depend Variable $=adjustedScore_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
CSR_cosine	0.214***	0.217***	0.160***			
	(0.026)	(0.026)	(0.026)			
CSR_freq				0.171***	0.170***	0.122***
				(0.020)	(0.020)	(0.020)
Constant	−1.185***	−5.229***	−35.566***	−1.202***	−5.055**	−35.467***
	(0.247)	(1.991)	(6.244)	(0.249)	(2.004)	(6.254)
Controls	No	No	Yes	No	No	Yes
Industry Fixed	No	Yes	Yes	No	Yes	Yes
Year Fixed	No	Yes	Yes	No	Yes	Yes
Observations	7781	7781	7781	7781	7781	7781
Adjusted R^2	0.018	0.051	0.156	0.018	0.050	0.155

This table reports the estimation results using an alternative CSR performance measure adjusted for firm size. The dependent variable is the firm's CSR rating in year $t + 1$ minus the quintile mean rating, with all independent and control variables measured in year t. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10 %, 5 %, and 1 % levels, respectively.

Table 5
Robustness Checks: Alternative Proxies for Institutional Investors' CSR Communication.

	Depend Variable = $Score_{t+1}$					
	Equal-weighted		Only Questions		Average Values	
	(1)	(2)	(3)	(4)	(5)	(6)
CSR_cosine	0.058***		0.183***		1.057***	
	(0.011)		(0.034)		(0.301)	
CSR_freq		0.885***		0.207***		1.123***
		(0.169)		(0.037)		(0.291)
Constant	−6.132	−6.201	−6.546	−6.321	−9.831	−9.791
	(6.289)	(6.284)	(6.226)	(6.235)	(6.191)	(6.196)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781	7781	7781
Adjusted R^2	0.176	0.175	0.176	0.176	0.173	0.173

This table reports the estimation results using three alternative constructions of the CSR communication variable. The dependent variable is the firms' CSR rating in year $t+1$, with all independent and control variables measured in year t. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10 %, 5 %, and 1 % levels, respectively.

the coefficients remain positive and statistically significant at the 1 % level when using these equal-weighted measures.

We then solely considered the question portions of the transcripts rather than the full question-and-answer section. In our baseline regression, the CSR communication measures are built using the Q&A sessions from the site visit transcripts. Interactions in site visits are dynamic, and questions posed by institutional investors are often uncertain and are not predetermined. Although this helps to ensure the accuracy and authenticity of managers' responses, some may engage in excessive self-marketing when answering. The CSR-related information we capture in investor–management conversations is then expected to be boilerplate language or "cheap talk" with no real impact. By assessing only investor questions we isolate institutional investors' initiations and thus mitigate concerns about biased manager responses. We substituted *CSR_cosine* and *CSR_freq* in our baseline model with the measures constructed from only the questions to address these concerns. The results reported in Columns (3) and (4) of Table 5 show a persistent and significant positive correlation between institutional investors' CSR communication and firms' CSR performance.

Finally, we redefined the CSR communication measures of a given firm by averaging the cosine similarity and relative frequency across all site visit transcripts for that firm in year t, thus quantifying the average proportion of CSR-related communication in the firm-year level. As Columns (5) and (6) of Table 5 show, the coefficient estimates on alternative measures are all significant and positive and thus consistent with our main results in Table 2.

In summary, the regression results are consistent with our main findings, indicating that they are robust to alternative measures and are unlikely to be driven by the method used to calculate CSR-related communication.

5.3. Endogeneity issues

Our findings demonstrate that institutional investors' CSR communications can improve firms' CSR performance. However, our results may suffer from endogeneity problems because the direct consideration of CSR-related issues by institutional investors during their interactions is not likely to be random. Endogeneity may thus occur due to potential factors simultaneously affecting CSR-related communication and firms' CSR performance. Reverse causality may also be a problem as the consideration of CSR-related issues by institutional investors during communications may be driven by changes in a firm's CSR performance. We therefore applied several procedures to address these potential problems.

Table 6
Robustness Checks: Controlling for Firm Fixed Effects.

	Depend Variable $=Score_{t+1}$			
	(1)	(2)	(3)	(4)
CSR_cosine	0.074***	0.035*		
	(0.021)	(0.020)		
CSR_freq			0.067***	0.035**
			(0.016)	(0.015)
Size		1.442***		1.437***
		(0.458)		(0.457)
Age		−0.955		−0.939
		(0.762)		(0.762)
Tangible		2.031		2.021
		(1.768)		(1.768)
Lev		−6.295***		−6.308***
		(1.367)		(1.366)
ROE		15.724***		15.671***
		(1.312)		(1.311)
Tobinq		−0.143		−0.145
		(0.123)		(0.123)
Inshold		0.014		0.014
		(0.014)		(0.014)
SOE		0.957		0.958
		(1.060)		(1.059)
Dual		−0.026		−0.024
		(0.377)		(0.377)
Indp		0.117***		0.117***
		(0.035)		(0.035)
Constant	51.373***	16.040	51.299***	16.104
	(0.251)	(10.206)	(0.250)	(10.196)
Firm Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781
Adjusted R^2	0.013	0.081	0.014	0.082

This table reports the estimation results after controlling for firm fixed effects. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

5.3.1. Controlling for firm-fixed effects

Although we controlled for a set of firm-level control variables in the baseline model, unobservable firm-level time-invariant variables may drive our results. To mitigate this concern, we re-estimated the main specifications with firm fixed effects. As Table 6 shows, the coefficients for the CSR communication measures remain positive and significant, although with smaller magnitudes. These results indicate that time-invariant firm-specific characteristics do not drive the positive relationship between institutional investors' CSR-related communication and CSR performance.

5.3.2. Controlling for CSR reports and communication tone

Various communication channels can convey information about a company's CSR activities or records (Ziek, 2009; Du et al., 2010), and annual CSR reports represent one of the most important channels. These CSR disclosures provide details of a firm's CSR activities and communicate CSR information to stakeholders, which may simultaneously determine the level of attention that institutional investors give to CSR issues and firms' future CSR practices. Investors can obtain a comprehensive picture of firms' CSR practices from these reports, and consequently can identify firms that could be persuaded to increase their CSR performance. Moreover, Chen et al. (2018) find that firms are often under pressure to increase their commitment to social responsibility when they are required to publish their CSR reports, which ultimately improves their corporate social responsibility practices. Therefore, the presence of CSR reports could be simultaneously associated with

Table 7
Endogeneity: Additional Controlling Variable.

	Depend Variable $=Score_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
CSR_cosine	0.126***	0.155***	0.126***			
	(0.023)	(0.026)	(0.023)			
CSR_freq				0.098***	0.118***	0.098***
				(0.018)	(0.020)	(0.018)
Report_cosine	2.592***		2.592***			
	(0.151)		(0.151)			
Report_freq				2.002***		2.002***
				(0.116)		(0.116)
Tone		0.173	0.225		0.031	−0.018
		(0.803)	(0.757)		(0.805)	(0.760)
Constant	20.765***	−5.620	20.619***	19.775***	−5.439	19.787***
	(6.002)	(6.340)	(6.055)	(5.990)	(6.355)	(6.050)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781	7781	7781
Adjusted R^2	0.252	0.178	0.252	0.253	0.177	0.253

This table presents the estimation results after controlling for additional variables. In Columns (1) and (4), we employed textual analysis on firms' CSR reports in year *t* to measure public CSR disclosure and additionally controlled for *Report_cosine* and *Report_freq*, corresponding to the two independent variables. In Columns (2) and (5), we include *Tone* to proxy for the sentiment in communications in year *t*. Columns (3) and (6) show the results after controlling for both additional variables. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

both the firm's CSR communication and its future CSR performance. To control for this potential confounding effect, we include CSR reporting as a control variable in our analysis.

In addition, our results may be driven by the tone conveyed rather than the specific content of the communication. The literature suggests that sentiment, as a form of non-verbal expressive behavior, can provide rich out-of-text information and have economic consequences (Loughran and McDonald, 2011; Price et al., 2012). Such omitted variables may affect the main results of our study.

We thus considered additional control variables in Eq. (6) to mitigate the endogeneity resulting from the omitted variables of firms' CSR report disclosures and the average tone of communications. Using the same CSR lexicon and textual analysis methodology, we constructed two measures for each annual CSR report, *Report_cosine* and *Report_freq*, to quantify the cosine similarity and relative frequency of the CSR information of the annual CSR report, respectively. We also controlled for the sentiment tone of the communications using a validated Chinese sentiment dictionary.[11] The sentiment is averaged across site visit transcripts to compute a firm-year level variable.

We then re-estimated the regressions with the expanded regression model. Table 7 shows the results. The coefficients on CSR communication remain significant and positive, consistent with the baseline model results. The consistency across these tests suggests that our findings are robust after further controlling for plausible omitted variables.

5.3.3. Change model analyses

Our findings may also be influenced by reverse causality, as institutional investors are more likely to discuss CSR issues with firms that already have strong CSR performance. Thus, the observed relationship may be correlational rather than causal. To account for this potential endogeneity, we estimated a change model as

[11] We measure the sentiment in a conversation as the share of positive tone words minus negative tone words using the sentiment dictionary developed by the National Taiwan University, one of the most commonly used sentiment dictionaries in Chinese natural language processing.

Table 8
Change Model Analyses.

	Depend Variable $=\Delta Score_{t+1}$					
	(1)	(2)	(3)	(4)	(5)	(6)
ΔCSR_cosine	0.042***	0.041***	0.034**			
	(0.014)	(0.014)	(0.014)			
ΔCSR_freq				0.037***	0.037***	0.031***
				(0.011)	(0.011)	(0.011)
Constant	−0.108	−1.524**	−4.537*	−0.104	−1.538**	−4.599*
	(0.072)	(0.676)	(2.603)	(0.072)	(0.674)	(2.603)
Controls	No	No	Yes	No	No	Yes
Industry Fixed	No	Yes	Yes	No	Yes	Yes
Year Fixed	No	Yes	Yes	No	Yes	Yes
Observations	5111	5111	5111	5111	5111	5111
Adjusted R^2	0.001	0.019	0.058	0.002	0.020	0.058

This table presents the estimation results of change model analyses. The dependent variable is the annual change in firms' CSR ratings from year t to $t+1$ ($\Delta Score_{i,t+1}$). The independent variables are annual changes in CSR communication measures (ΔCSR_cosine and ΔCSR_freq). Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10 %, 5 %, and 1 % levels, respectively.

an additional robustness test, in which we use the differences of the variables rather than their levels. The dependent variable is the firm's CSR performance change from year t to $t+1$. The independent variable is the change in CSR communication. These analyses filter out time-invariant unobservable firm heterogeneity and thus better identify the causal effect and alleviating endogeneity issues.

Table 8 shows that after considering the differences in the variables, the relationship between changes in CSR-related communication (ΔCSR_cosine and ΔCSR_freq) and changes in firms' CSR ratings ($\Delta Score$) remains significant and positive. This indicates that even after accounting for time-invariant unobservable firm heterogeneity, CSR communication is significantly associated with improvements in CSR performance.

5.3.4. Instrumental variable

Institutional investors' consideration of CSR-related issues may be driven by changes in firms' CSR performance, leading to a reverse causality problem. Potential endogeneity may then remain a concern after applying the above procedures. Thus, we used a two-stage least squares (2SLS) analysis to further address the endogeneity problem, in which we introduced an instrumental variable for the CSR-related communication of institutional investors in site visits.

We considered the number of non-top-10 shareholding institutional investors who participated in the site visits and are signatories to the United Nations' PRI as an instrumental variable for CSR communication.

PRI is a global initiative of the United Nations aimed at encouraging financial institutions to consider ESG factors in their investment decisions, with the goal of promoting sustainable development and long-term value creation. Signatories of the PRI must regularly report on their progress in integrating ESG factors and how they are advancing sustainability goals in their investment practices. The PRI also encourages institutional investors to engage in dialogue with the firms they invest in to promote better environmental and social responsibility practices (Gibson Brandon et al., 2022). Thus, institutional investors who are PRI signatories are likely to devote more attention to CSR issues when communicating with firms. Therefore, the proportion of CSR-related issues in the conversations may increase with the number of PRI signatories among visiting institutional investors.

The choice to become a PRI signatory is at the discretion of each institutional investor. Their decision is unlikely to be influenced by any specific firm, as investors often have diversified portfolios spanning multiple firms. To reinforce the exogeneity of our instrumental variable, we only considered the number of non-top-10 shareholding institutional investors among the site visit participants. Although comprehensive shareholder data are not available, by restricting our sample to these shareholders we only consider institutional investors with limited influence over firms' CSR practices. Investors with large shareholdings may affect CSR by exercising shareholder rights, either through voting or threatening divestment. By limiting the instrument to these

Table 9
Endogeneity: An Instrumental Variables Approach.

	1st Stage		2nd Stage	
	CSR_cosine	*CSR_freq*	$Score_{t+1}$	
	(1)	(2)	(3)	(4)
CSR_cosine			0.533***	
			(0.180)	
CSR_freq				0.411***
				(0.138)
PRI	0.292***	0.379***		
	(0.046)	(0.058)		
Constant	−23.627***	−31.719***	3.893	4.346
	(4.537)	(5.841)	(7.758)	(7.851)
Controls	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781
Adjusted R^2	0.085	0.085	0.127	0.129

This table reports the estimation results of the two-stage least squares (2SLS) analysis by using *PRI* as an instrumental variable. Columns (1) and (2) show the results for the first stage, in which we examine the relationship between PRI and CSR communication measures. Columns (3) and (4) report the results for the second stage regression. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 10
Institutional investors' CSR Communication, State-owned enterprises, CSR performance.

	Depend Variable = $Score_{t+1}$			
	SOE	*non-SOE*	*SOE*	*non-SOE*
	(1)	(2)	(3)	(4)
CSR_cosine	0.048	0.185***		
	(0.041)	(0.030)		
CSR_freq			0.053*	0.135***
			(0.032)	(0.024)
Constant	−24.933**	1.752	−24.359**	1.856
	(10.809)	(7.764)	(10.811)	(7.802)
Controls	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	1641	6140	1641	6140
Adjusted R^2	0.223	0.168	0.223	0.166
Comparison coefficients	Observed difference = 0.138		Observed difference = 0.082	
	p-value = 0.000		p-value = 0.004	

This table reports the effects of institutional investors' CSR communication on CSR performance for both state-owned and non-state-owned enterprises. The last row shows the significance of the differences in coefficient estimates between two groups. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

non-top-10 shareholding institutional investors, we exclude such potential channels of influence. We manually collected the names of Chinese institutions that had joined before 2020 from the online PRI signatory directory,[12] and thus obtained a list of Chinese investment managers. We used fuzzy matching to compare these signatories against the names of the non-top-10 institutional shareholders participating in site visits.

Table 9 presents the results of the 2SLS analysis. Columns (1) and (2) show the results of the first stage, in which the instrumental variable is significantly and positively correlated with the endogenous explanatory

[12] https://www.unpri.org/signatories.

variables. This suggests that PRI signatories involved in site visits are likely to place a greater emphasis on CSR issues during their interactions. The Cragg–Donald Wald F statistics (80.28 and 85.33, respectively) indicate no problem of weak instrumental variables. The second stage regression results in Columns (3) and (4) reveal that the coefficients corresponding to the fitted CSR communication are significant and positive at the 1 % level. Thus, our main findings remain robust even after considering potential endogeneity issues through the instrumental variable approach.

6. Cross-sectional implications and plausible mechanisms

The baseline regression confirms that the CSR communication of institutional investors may be a driver of future improved CSR performance. In this section, we report our investigation of changes in this effect. We considered three situations: state-owned enterprises (SOEs), periods following the issuance of Green Investment Guidelines, and firms with low institutional ownership. We then explore the potential mechanisms for the effect of CSR-related communication on firms' CSR performance.

6.1. SOEs

Previous research finds that unlike non-SOEs, whose primary goal is to maximize profits or shareholder wealth, SOEs have additional social responsibilities as required by the government, including infrastructure construction, addressing employment problems and participating in poverty alleviation programs (Piotroski and Wong, 2012). SOEs therefore recognize the strategic significance of CSR and may be more willing to undertake relevant endeavors. However, as non-SOEs do not face such institutional pressure, they may lack the incentive to engage in sustainable development. Therefore, the expectations of and attention from institutional investors may play a more effective role in non-SOEs that have scope for CSR improvement.

We define SOEs as firms in which the government is the ultimate controlling owner. We divided the sample into two subsamples based on whether a firm is an SOE, and re-ran our baseline model. Table 10 shows a significant effect of institutional investors' CSR communication on the CSR score of a firm that is a non-SOE. In addition, a comparison of the coefficients of CSR communication measures for the two subsamples suggests that the differences between them are statistically significant. Overall, non-SOEs do not face the same

Table 11
Institutional Investors' CSR Communication, Policy Implementation, CSR Performance.

	Depend Variable = $Score_{t+1}$			
	Before 2018	*After 2018*	*Before 2018*	*After 2018*
	(1)	(2)	(3)	(4)
CSR_cosine	0.141***	0.200***		
	(0.026)	(0.045)		
CSR_freq			0.106***	0.155***
			(0.021)	(0.034)
Constant	−1.324	−20.726**	−1.397	−20.564**
	(6.540)	(9.471)	(6.541)	(9.482)
Controls	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	5098	2683	5098	2683
Adjusted R^2	0.170	0.180	0.169	0.179
Comparison coefficients	Observed difference = −0.060		Observed difference = −0.050	
	p-value = 0.047		p-value = 0.033	

This table reports the effects of institutional investors' CSR communication on CSR performance across the periods before and after the issuance of the Green Investment Guidelines. The last row shows the significance of the differences in coefficient estimates between two groups. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

external pressure from the government's sustainable development strategies as SOEs, and the attention from institutional investors regarding CSR issues can make up for this lack of external drive.

6.2. Implementation of Green Investment Guidelines

Changes in the regulatory environment may lead to differences in the effectiveness of institutional investor communication in different periods. The China Securities Investment Fund Industry Association introduced the "Green Investment Guidelines (Trial)" in 2018, which represents the first comprehensive and systematic set of green investment self-regulatory standards. Institutional investors are required to prioritize investments in sustainable firms, engage in responsible investment and use their rights to urge investee firms to improve E&S performance and corporate information disclosure. This policy has strengthened the incentives for institutional investors to implement green investment, thus encouraging asset managers to focus on sustainability. Therefore, we expect the effect of institutional investors' CSR communication on CSR performance to be greater after the introduction of these guidelines.

We therefore divided the samples into two subsamples before and after 2018, as the time of the policy release. We re-estimated our baseline model using these subsamples. The results presented in Table 11 indicate that the positive correlation between CSR communication measures and CSR scores is high for the period following the implementation of the Green Investment Guidelines (Trial), and the difference between the coefficients of the two groups has statistical significance.

6.3. Institutional ownership

Chen et al. (2020) provide evidence that institutional investors can push for high firm-level E&S performance and generate real social impact by influencing the CSR policies of their portfolio firms through the rights that come with their shareholdings. Our main results show that the voice of institutional investors in private communication can affect CSR performance. We also find that this influence is not limited to shareholding institutional investors. Those who do not yet hold shares can also impact CSR improvement if they participate in private meetings. Thus, we further examine whether institutional investor communication in

Table 12
Institutional Investors' CSR Communication, Institutional Ownership, CSR Performance.

	Depend Variable = $Score_{t+1}$			
	High Institutional Shareholding	*Low Institutional Shareholding*	*High Institutional Shareholding*	*Low Institutional Shareholding*
	(1)	(2)	(3)	(4)
CSR_cosine	0.121***	0.204***		
	(0.036)	(0.036)		
CSR_freq			0.102***	0.144***
			(0.028)	(0.029)
Constant	−16.702**	19.542**	−16.401**	19.463**
	(8.004)	(8.947)	(8.009)	(9.010)
Controls	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	3864	3917	3864	3917
Adjusted R^2	0.180	0.184	0.181	0.181
Comparison coefficients	Observed difference = −0.083		Observed difference = −0.042	
	p-value = 0.006		p-value = 0.047	

This table shows the relationship between Institutional investors' CSR Communication and CSR performance for firms with high and low levels of institutional shareholding, partitioned based on the median value of institutional ownership. The last row shows the significance of the differences in coefficient estimates between two groups. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

general can be an alternative mechanism for institutional ownership and affect CSR performance by assessing whether it can produce a marginal effect in firms with low levels of institutional ownership and promote the improvement of their CSR performance.

We divided the sample based on the median of institutional ownership and compared the impact of CSR communication on high- and low-institutional shareholding groups. Table 12 reports the results. Columns (1) and (2) show that the coefficients on CSR communication are positive and significant at the 1 % level in both the high and low ownership groups. This indicates a robust relationship between CSR communication and performance, regardless of institutional ownership characteristics. The coefficient is higher for the low institutional ownership group (0.204) than for the high group (0.121), and the difference is significant at 1 %. This suggests CSR communication has a greater impact with a lower level of institutional shareholding. Engagement through private communication may be the investors' primary channel of influence when the stakes are lower. Those with higher stakes can also leverage shareholder rights. Overall, these results support the substitutive relationship between institutional ownership and communication.

6.4. Underlying mechanism

Thus far, we have documented that CSR-related communication during site visits can lead to subsequent improvements in firms' CSR performance. We further investigate potential channels for our results. One potential mechanism is that through effective communication, these investors can enhance the transparency of CSR information, which then leads to CSR performance improvements. Chen et al. (2018) suggest that firms face pressure to strengthen their CSR commitments when they are required to disclose their CSR activities. Investor–manager interactions regarding CSR provide additional relevant disclosures to the market. This may garner further attention from stakeholders (e.g., suppliers, consumers and communities) and the government, thereby disciplining firms and improving their corporate behavior. Research also shows that a firm's information environment affects its sustainability (Burke, 2022). Therefore, institutional investors can support firms' CSR efforts by increasing public access to information. The attention given to CSR issues by institutional investors can strengthen external oversight, attract the attention of other stakeholders and promote the firm's sustainable development by improving corporate information transparency.

Alleviating financing constraints may serve as another potential mechanism through which communication can improve CSR practices. Previous studies highlight that financial constraints have a major impact on sustainable corporate policies. Financially unconstrained firms are more likely to invest in a sustainable

Table 13
Underlying Mechanism.

	(1)	(2)	(3)	(4)
	Information environment	*Information environment*	*Financing Constraint*	*Financing Constraint*
CSR_cosine	0.007***		−0.008***	
	(0.001)		(0.003)	
CSR_freq		0.006***		−0.007***
		(0.001)		(0.002)
Constant	−2.607***	−2.581***	4.620***	4.584***
	(0.227)	(0.226)	(0.607)	(0.608)
Controls	Yes	Yes	Yes	Yes
Industry Fixed	Yes	Yes	Yes	Yes
Year Fixed	Yes	Yes	Yes	Yes
Observations	7781	7781	7781	7781
Adjusted R^2	0.141	0.143	0.522	0.522

This table presents results examining the underlying mechanisms institutional investors' CSR communication may influence corporate sustainability. Columns (1) and (2) analyze the effect on firms' information environments, proxied by annual information disclosure ratings. Columns (3) and (4) examine impacts on financing constraints, measured using the KZ Index. CSR communication is captured by the variables *CSR_cosine* and *CSR_freq*. Standard errors reported in parentheses are clustered at the firm level. Variable definitions are provided in Appendix A. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

development strategy, as this requires a solid financial base and resource input (Xu and Kim, 2022). Chapman et al. (2022) suggest that direct and ongoing dialogue between management and investors is essential for mutual understanding and trust. Effective communication between firms and investors can reduce investors' uncertainty about CSR initiatives, thus lowering perceived investment risks. Therefore, firms may face fewer financing constraints and can allocate more resources toward CSR activities.

We measured the information environment and financing constraint using two proxies to test these two potential channels. We first considered the annual information disclosure ratings of listed companies in the SZSE to measure their information environments. The SZSE categorizes firms into four annual grades (A, B, C and D) based on the quality of their information disclosure. A firm receiving an "A" grade is considered to have a high-quality information environment. We thus assigned these firms an *information environment* variable value of 1. We also used the KZ Index, as proposed by Kaplan and Zingales (1997), as a proxy for financial constraint. A high KZ index value indicates a great degree of financial constraint.

Table 13 presents the results of examining the mechanisms underlying the influence of institutional investors' CSR communication on corporate sustainability. Columns (1) and (2) give the results from analyzing the effect on firms' information environments, proxied by annual information disclosure ratings. Columns (3) and (4) give those from examining the impacts on financing constraints, measured using the KZ Index. The coefficients on *CSR_cosine* and *CSR_freq* in Columns (1) and (2) are positive and significant, indicating that firms with a higher level of CSR communication have superior information environments. This supports the mechanism of investor–manager CSR dialogue, which improves corporate transparency.

Similarly, the negative and significant coefficients on the CSR communication variables in Columns (3) and (4) suggest that CSR discussions between investors and managers reduce firms' financing constraints. This is consistent with the mechanism of CSR communication, which reflects investor support and the ease of access to resources. Overall, our empirical results provide evidence for both proposed channels. CSR communication with institutional investors enhances corporate information environments and relaxes financing constraints. This enables firms to devote greater resources and commitment to sustainability initiatives.

7. Conclusion

In this study, we explore the impact of institutional investors' communication on CSR. We examine site visit transcripts of Chinese A-share firms in the SZSE from 2013 to 2021 and conduct a machine learning-based textual analysis to quantify the level of attention institutional investors give to CSR issues when communicating with management. Our results provide strong evidence that CSR communication during site visits can effectively enhance subsequent CSR performance. We also analyze the cross-sectional characteristics affecting the strength of the impact of CSR communication. The effect is more pronounced for non-state-owned firms than for SOEs, indicating that institutional investors play a more crucial monitoring role in the absence of state oversight. In addition, the impact increases following the introduction of green policies such as the Green Investment Principles, suggesting that such frameworks encourage and empower investors to engage in sustainability. Firms with lower levels of institutional ownership also exhibit a more substantial effect, implying that private and direct communication has a more important role in improving CSR when shareholders' monitoring is limited.

Our study has meaningful implications for both policy and practice. It highlights the need for institutional investors to proactively communicate their CSR priorities during private interactions. Their monitoring role should involve going beyond financial metrics and engaging with CSR issues that can lead to long-term value creation. For market regulators, enabling institutional investor access and promoting CSR communication norms can lead to improved accountability regarding sustainability practices. Although we focus on Chinese firms, future research can examine other locations, as investors worldwide increasingly prioritize sustainability; thus, understanding how they can effectively foster CSR improvements remains an important topic.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Claude in order to improve language and readability. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Declaration of competing interest

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

Acknowledgements

This work is supported by the Major Program of National Fund of Philosophy and Social Science of China (No. 19ZDA105).

Appendix A. Variable Definitions

Variables	Definitions
TF-IDF weighted CSR-related Communication Measures	
$CSR\_cosine_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector of the CSR dictionary across all transcripts of the firm i in year t.
$CSR\_freq_{i,t}$	The sum of the relative frequencies of TF-IDF weighted CSR words in each transcript across all transcripts of the firm i in year t.
Dependent Variables	
$Score_{i,t}$	The annual average of the quarterly ESG rating scores of firm i in year t.
TF-IDF weighted CSR-related Communication Measures by Dimension	
$cosine\_share_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector for the "shareholders and creditors" dimension across all transcripts of firm i in year t.
$cosine\_emp_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector for the "employees" dimension across all transcripts of firm i in year t.
$cosine\_scc_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector for the "suppliers, customers and consumers " dimension across all transcripts of firm i in year t.
$cosine\_env_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector for the "environmental protection and sustainable development" dimension across all transcripts of firm i in year t.
$cosine\_social_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of each transcript and the TF-IDF weighted average vector for the "public relations and social welfare services" dimension across all transcripts of firm i in year t.
$freq\_share_{i,t}$	The sum of relative frequencies of TF-IDF weighted words for the "shareholders and creditors" dimension in each transcript across all transcripts of firm i in year t.
$freq\_emp_{i,t}$	The sum of relative frequencies of TF-IDF weighted words for the "employees" dimension in each transcript across all transcripts of firm i in year t.

Appendix A. (*continued*)

Variables	Definitions
$freq\_scc_{i,t}$	The sum of relative frequencies of TF-IDF weighted words for the "suppliers, customers and consumers" dimension in each transcript across all transcripts of firm i in year t.
$freq\_env_{i,t}$	The sum of relative frequencies of TF-IDF weighted words for the "environmental protection and sustainable development" dimension in each transcript across all transcripts of firm i in year t.
$freq\_social_{i,t}$	The sum of relative frequencies of TF-IDF weighted words for the "public relations and social welfare services" dimension in each transcript across all transcripts of firm i in year t.
Variables in Robustness Tests	
$adjustedScore_{i,t}$	$adjustedScore_{i,t}$ is defined as CSR rating score adjusted for the effect of firm size. It is constructed by subtracting the average CSR score of firms in the same size quintile from each firm's CSR rating score.
Unweighted CSR-related Communication Measures	
$CSR\_cosine_{i,t}$	The sum of cosine similarities between the unweighted word vector of each transcript and the unweighted CSR dictionary vector across all transcripts of company i in year t.
$CSR\_freq_{i,t}$	The sum of relative frequencies of unweighted CSR dictionary words in each transcript across all transcripts of company i in year t.
Questions-Only CSR Communication Measures	
$CSR\_cosine_{i,t}$	The sum of cosine similarities between the TF-IDF weighted word vector of the questions portion of each transcript and the TF-IDF weighted average vector of the CSR dictionary across all transcripts of the firm i in year t.
$CSR\_freq_{i,t}$	The sum of relative frequencies of TF-IDF weighted CSR words appearing only in the questions portion of each transcript across all transcripts of the firm i in year t.
Average CSR Communication Measures	
$CSR\_cosine_{i,t}$	The average cosine similarity between the vector representation of each transcript and the CSR dictionary vector across all transcripts of the company i in year t.
$CSR\_freq_{i,t}$	The average relative frequency of CSR dictionary words across all transcripts of the company i in year t.
CSR Communication Measures on CSR Report	
$Report\_cosine_{i,t}$	The cosine similarity between the TF-IDF weighted word vector of the CSR report and the TF-IDF weighted word vector of the CSR dictionary for the company i in year t.
$Report\_freq_{i,t}$	The relative frequency of TF-IDF weighted CSR dictionary words in the CSR report for company i in year t.
Communication Tone	
$Tone_{i,t}$	The average relative frequency of positive versus negative tone words in the site visit transcripts of the company i in year t.
Control Variables	
$Size_{i,t}$	Natural logarithm of total assets.
$Age_{i,t}$	Natural logarithm of the number of years firm i has been listed on a stock exchange at the end of year t plus one.
$Tangible_{i,t}$	The ratio of property, plant, and equipment to total assets.
$Lev_{i,t}$	The ratio of total debt to total assets.
$ROE_{i,t}$	The ratio of net income to shareholders' equity.
$Tobinq_{i,t}$	The ratio of the firm's market value to its book value.
$Inshold_{i,t}$	The proportion of shares held by institutional investors.
$SOE_{i,t}$	A dummy variable equals one if the firm is state-owned and 0 otherwise.
$Indp_{i,t}$	The percentage of independent directors on the board.
$Dual_{i,t}$	A dummy variable equals one if the firm's CEO also holds the position of chairman of the board of the same firm and 0 otherwise.

References

Barko, T., Cremers, M., Renneboog, L., 2022. Shareholder engagement on environmental, social, and governance performance. J. Bus. Ethics 180 (2), 777–812.

Borghesi, R., Houston, J.F., Naranjo, A., 2014. Corporate socially responsible investments: CEO altruism, reputation, and shareholder interests. J. Corp. Financ. 26, 164–181.

Bowen, R.M., Dutta, S., Tang, S., Zhu, P., 2018. Inside the "black box" of private in-house meetings. Rev. Account. Stud. 23 (2), 487–527.

Brown, S., Hillegeist, S.A., Lo, K., 2004. Conference calls and information asymmetry. J. Account. Econ. 37 (3), 343–366.

Burke, J.J., 2022. Do boards take environmental, social, and governance issues seriously? Evidence from media coverage and CEO dismissals. J. Bus. Ethics 176 (4), 647–671.

Bushee, B.J., Gerakos, J., Lee, L.F., 2018. Corporate jets and private meetings with investors. J. Account. Econ. 65 (2–3), 358–379.

Chapman, K.L., Miller, G.S., Neilson, J.J., White, H.D., 2022. Investor relations, engagement, and shareholder activism. Account. Rev. 97 (2), 77–106.

Chen, T., Dong, H., Lin, C., 2020. Institutional shareholders and corporate social responsibility. J. Financ. Econ. 135 (2), 483–504.

Chen, X., Harford, J., Li, K., 2007. Monitoring: Which institutions matter? J. Financ. Econ. 86 (2), 279–305.

Chen, Y.-C., Hung, M., Wang, Y., 2018. The effect of mandatory CSR disclosure on firm profitability and social externalities: evidence from China. J. Account. Econ. 65 (1), 169–190.

Cheng, Q., Du, F., Wang, X., Wang, Y., 2016. Seeing is believing: analysts' corporate site visits. Rev. Account. Stud. 21 (4), 1245–1286.

Cheng, Q., Du, F., Wang, B.Y., Wang, X., 2018. Do corporate site visits impact stock prices? Contemp. Account. Res. 36 (1), 359–388.

Dhaliwal, D.S., Li, O.Z., Tsang, A., Yang, Y.G., 2011. Voluntary nonfinancial disclosure and the cost of equity capital: the initiation of corporate social responsibility reporting. Account. Rev. 86 (1), 59–100.

Dikolli, S.S., Frank, M.M., Guo, Z.M., Lynch, L.J., 2022. Walk the talk: ESG mutual fund voting on shareholder proposals. Rev. Account. Stud. 27 (3), 864–896.

Dimson, E., Karakaş, O., Li, X., 2015. Active ownership. Rev. Financ. Stud. 28 (12), 3225–3268.

Du, S., Bhattacharya, C.B., Sen, S., 2010. Maximizing business returns to Corporate Social Responsibility (CSR): the role of CSR communication. Int. J. Manag. Rev. 12 (1), 8–19.

Dyck, A., Lins, K.V., Roth, L., Wagner, H.F., 2019. Do institutional investors drive corporate social responsibility? International Evidence. J. Financ. Econ. 131 (3), 693–714.

Fernando, C.S., Sharfman, M.P., Uysal, V.B., 2017. Corporate environmental policy and shareholder value: following the smart money. J. Financ. Quant. Anal. 52 (5), 2023–2051.

Gibson Brandon, R., Glossner, S., Krueger, P., Matos, P., Steffen, T., 2022. Do responsible investors invest responsibly?*. Rev. Financ. 26 (6), 1389–1432.

Gillan, S.L., Starks, L.T., 2000. Corporate governance proposals and shareholder activism: the role of institutional investors. J. Financ. Econ. 57 (2), 275–305.

Glo$ner, S., 2019. Investor horizons, long-term blockholders, and corporate social responsibility. J. Bank Financ. 103 (6), 78–97.

Guiso, L., Sapienza, P., Zingales, L., 2008. Trusting the stock market. J. Financ. 63 (6), 2557–2600.

Han, B., Kong, D., Liu, S., 2018. Do analysts gain an informational advantage by visiting listed companies? Contemp. Account. Res. 35 (4), 1843–1867.

Heath, D., Macciocchi, D., Michaely, R., Ringgenberg, M.C., 2023. Does Socially Responsible Investing Change Firm Behavior?*. Rev. Financ.

Hirschman, A.O. 1970. Exit, voice, and loyalty: Responses to decline in firms, organizations, and states, vol. 25. Harvard university press.

Hockerts, K., Moir, L., 2004. Communicating corporate responsibility to investors: the changing role of the investor relations function. J. Bus. Ethics 52 (1), 85–98.

Hoepner, A.G.F., Oikonomou, I., Sautner, Z., Starks, L.T., Zhou, X.Y., 2023. ESG shareholder engagement and downside risk. Rev. Financ.

Hong, X., Zhuang, Z., Kang, D., Wang, Z., 2019. Do corporate site visits impact hedge fund performance? Pac.-Basin Financ. J. 56, 113–128.

Hu, Y., Shan, J., Zhan, P., 2020. Institutional investors' corporate site visits and firms' sustainable development. Sustainability 12 (17).

Hwang, C.Y., Titman, S., Wang, Y., 2022. Investor tastes, corporate behavior, and stock returns: an analysis of corporate social responsibility. Manage. Sci. 68 (10), 7065–7791.

Jiang, Y.H., Wang, C., Li, S., Wan, J., 2022. Do institutional investors' corporate site visits improve ESG performance? Evidence from China. Pac.-Basin Financ. J, 76.

Jiang, X., Yuan, Q., 2018. Institutional investors' corporate site visits and corporate innovation. J. Corp. Financ. 48 (1), 148–168.

Kaplan, S.N., Zingales, L., 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? Quart. J. Econ. 112 (1), 169–215.

Kim, I., Wan, H., Wang, B., Yang, T.N., 2019. Institutional investors and corporate environmental, social, and governance policies: evidence from toxics release data. Manage. Sci. 65 (10), 4901–4926.

Kim, S., Yoon, A., 2022. Analyzing active fund managers? Commitment to ESG: evidence from the united nations principles for responsible investment. Manage Sci.

Lee, C.M.C., Zhong, Q., 2022. Shall we talk? The role of interactive investor platforms in corporate communication. J. Account. Econ. 74 (2–3).

Li, K., Mai, F., Shen, R., Yan, X., 2021. Measuring corporate culture using machine learning. Rev. Financ. Stud. 34 (7), 3265–3315.

Lin, Y.J., Fu, X.Q., Fu, X.L., 2021. Varieties in state capitalism and corporate innovation: evidence from an emerging economy. J. Corp Financ., 67
Lins, K.V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: the value of corporate social responsibility during the financial crisis. J. Financ. 72 (4), 1785–1823.
Liu, S., Dai, Y., Kong, D., 2017. Does it pay to communicate with firms? Evidence from firm site visits of mutual funds. J. Bus. Finan. Account. 44 (5–6), 611–645.
Liu, J.C., Xiong, X., Gao, Y., Zhang, J., 2023. The impact of institutional investors on ESG: evidence from China. Account. Financ. 63, 2801–2826.
Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. J. Financ. 66 (1), 35–65.
McCahery, J.A., Sautner, Z., Starks, L.T., 2016. Behind the scenes: the corporate governance preferences of institutional investors. J. Financ. 71 (6), 2905–2932.
Mikolov, T., Chen, K., Corrado, G., Dean, J., 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
Nofsinger, J.R., Sulaeman, J., Varma, A., 2019. Institutional investors and corporate social responsibility. J. Corp. Financ. 58 (5), 700–725.
Nofsinger, J., Varma, A., 2014. Socially responsible funds and market crises. J. Bank Financ. 48, 180–193.
Piotroski, J. D., Wong, T. 2012. Institutions and information environment of Chinese listed firms. In: Capitalizing China, University of Chicago Press, pp. 201–242.
Price, S.M., Doran, J.S., Peterson, D.R., Bliss, B.A., 2012. Earnings conference calls and stock returns: The incremental informativeness of textual tone. J. Bank Financ. 36 (4), 992–1011.
Reiter, N., 2021. Investor communication and the benefits of cross-listing. J. Account. Econ. 71 (1) 101356.
Renneboog, L., Ter Horst, J., Zhang, C.D., 2008. Socially responsible investments: Institutional aspects, performance, and investor behavior. J. Bank Financ. 32 (9), 1723–1742.
Rennekamp, K.M., Sethuraman, M., Steenhoven, B.A., 2022. Engagement in earnings conference calls. J. Account. Econ. 74 (1).
Shen, H.T., Lin, H.H., Han, W.Q., Wu, H.Y., 2023. ESG in China: a review of practice and research, and future research avenues. China J Account Res 16 (4).
Soltes, E., 2014. Private interaction between firm management and sell-side analysts. J. Account. Res. 52 (1), 245–272.
Wu, D., 2023. Text-based measure of supply chain risk exposure. Manage Sci.
Xu, Q., Kim, T., 2022. Financial constraints and corporate environmental policies. Rev. Financ. Stud. 35 (2), 576–635.
Ziek, P., 2009. Making sense of CSR communication. Corp. Soc. Responsib. Environ. Manag. 16 (3), 137–145.