Paying Attention to ESG Matters: Evidence from Big Data Analytics

Wendi Huang, G. Andrew Karolyi and Alan Kwan†

Abstract

Environment, social and governance (ESG) issues are increasingly important to investors. Yet, ESG ratings feature data quality issues with considerable dispersion among data providers. We propose a new measure based on *attention* to ESG issues using novel data of a firm's internet search intensity around business topics. Increases in firm's attention to ESG-related topics predict improvements in real outcomes as well as in a firm's ESG ratings. Investor attention to ESG predicts changes in investors' voting patterns and investment positions. Studying co-movement between attention to ESG by firms and their investors offers new evidence of how investors influence firms.

Key words: ESG; Big data analytics; intent data; institutional investors.
JEL Classification Codes: C55, G11, G23, G24, Q01.
Version : July 2021.

† Huang is a PhD student, Hong Kong University Business School, Karolyi is Professor of Finance and Harold Bierman Jr. Distinguished Professor of Management at the Cornell S.C. Johnson College of Business, Cornell University and Kwan is Assistant Professor, Hong Kong University Business School. The authors thank their discussants at CICF (Nan Li), SFS Cavalcade (Kai Li), Asian Bureau of Finance and Economic Research (Philipp Krüger), seminar participants at Hong Kong University, Cheung Kong Business School, City University Hong Kong, as well as Brad Barber, Rajna Gibson, Harrison Hong, Oguzhan Karakas, Pedro Matos, Justin Murfin, Roberto Rigobon, Lukas Roth, Laura Starks, Russ Wermers, William Xiong, Ruoke Yang, and Ayako Yasuda for valuable comments. They are grateful to The Company for sharing their business-to-business intent data, especially their executive team members who answered many questions about the data science underlying the demographic and firmographic data at scale. The authors have secured access to this proprietary data under a non-disclosure agreement with The Company; there are no relevant and material financial relationships that bear on their research. Karolyi serves as a consultant to Avantis Investors. Huang may be reached at hwdown@connect.hku.hk, Karolyi, at gak56@cornell.edu and Kwan, at apkwan@hku.hk.

1. Introduction.

Businesses today are deeply intertwined with environmental, social, and governance (ESG) concerns. The "E" in ESG includes the energy a firm takes in and the waste it discharges, and encompasses carbon emissions and climate change; "S" is about the relationships and reputation a firm has in the communities and countries in which it does business, including labor relations and diversity, equity and inclusion; and, "G" is the internal system of practices, controls and procedures a firm adopts to govern itself to comply with the law and meet needs of external stakeholders. In August 2019, the U.S. Business Roundtable strongly affirmed a new set of business principles featuring ESG issues as a linchpin.[1] ESG-oriented investing has experienced a major rise: assets managed by investors who incorporate ESG criteria into their investing process now top $30 trillion globally. In the U.S., Congress, with the proposed H.R. 1187 *ESG Disclosure Simplification Act of 2021*, is now moving toward making corporate ESG disclosures mandatory.

Despite the large allocation of societal resources toward promoting social responsibility by corporations, in a recent survey for the CFA Institute, Matos (2020) suggests that, despite the booming interest in corporate social responsibility, there is a commonly held view that ESG implementation has been hampered to date because it has not been defined consistently. Matos further notes data quality issues in ESG ratings from commercial data providers, which makes it difficult for investors to identify which businesses adhere to ESG tenets and for academics to study the impact of ESG investing on performance and real corporate change. A final issue he points to is "greenwashing," a false or exaggerated representation regarding how well aligned investments really are with sustainability goals.

In this paper, we take a novel approach to understanding ESG and its consequences for corporate change by introducing novel data on the attention paid to ESG by employees of the firms and by their investors. The proprietary data come from a data analytics firm (which we will call "The Company") that specializes in measuring online/digital organization-level

[1] See "Statement on the Purpose of a Corporation" (August 2019) as signed by dozens of corporate CEOs around the world. See also Larry Fink's open letter to CEOs, "A Fundamental Reshaping of Finance" (January 14, 2020), or Japan's Government Pension Investment Fund (GPIF), the world's largest pension fund, announcing revisions in 2017 to incorporate ESG issues, entitled "Fiduciary Duty in the 21st Century" (2019). The Global Sustainable Investment Review 2018 collates results from market studies of regional sustainable investment forums from Europe (Eurosif), U.S. (SIF), Japan (Sustainable Investment Forum), Canada (RIA Canada), Australia, and New Zealand (Responsible Investment Association Australiasia). The report is sponsored by Hermes Investment Management, RBC, and UBS.

interest in business-related topics. The data cover a comprehensive subset of the universe of U.S. firms and a wide range of digital content spanning thousands of topics across various themes of business, including human resources, business strategy and operations, finance, marketing, enterprise and consumer technology, biotech, engineering, construction and manufacturing. Several thousands of partner media publishers contribute, among which are household names such as *Bloomberg* and *Wall Street Journal* in finance and *Laptop Mag* in technology. While many publishers are news outlets, some contribute technical content, policy-oriented white papers or video content. Participation among partners is rewarded with access to some of the data provider's analytics products and services.

For each of the billions of web content interactions observed monthly across the publisher network, the firm categorizes the relevant topic, the location of the IP address and the organization associated with the user, where possible. The intended use of the dataset is to facilitate sales teams in finding leads among business customers. Presumably, a prospective customer displaying elevated interest in the form of content interactions suggests the prospective customer might be likelier to buy a related product or service. This data is part of advertising and marketing analytics called "intent data," and The Company is arguably a category leader.[2] We possess two versions of the data, one of which is not sold and was made available solely for our paper. The first is a weekly topic-firm interest index with firm scores dating back to May 2015. Like Google Trends, it is an index score between 0 (low reading intensity) to 100 (high reading intensity) produced at a weekly level. Unlike Google Trends, our data connect the reading level intensity *to a specific firm*. Second, we have a more granular, daily topic-firm count of content interactions dating back to May 2016.

The key innovation of our paper is to leverage The Company's intent data to construct a new measure of firm (and investor) attention to ESG-related issues. We hand-select from over 5,000 topics a subset of 323 ESG-relevant topics. We place them into one of nine categories aimed to match common classifications from ESG ratings.[3] The largest categories

[2] A recent paper by Tong, Luo, and Xu (2020) includes a review of research across major marketing journals that study mobile marketing phenomena and consumer behavior changes using consumer hyper-context information (e.g. location, time, environment) to design personalized targeting ads. The earliest work includes that by Balasubramanian, Peterson, and Jarvenpaa (2002) and Barwise and Strong (2002).

[3] These ratings are Compliance, Corporate Governance, Customer Relations, Cybersecurity, Data and Sensitive Information Protection, Environment, Equality and Diversity, Labor Relations and Corporate Social Responsibility. This final category covers topics almost exclusively on social issues, other than those focused on labor, customer and diversity. It includes sub-

are Labor Relations (63 topics), Environment (46 topics) and Corporate Governance (29 topics). Examples of environmental topics include "Air Pollution," "Global Warming," "Climate Change," and "Emissions."

Preliminary diagnostics suggest that our measures meaningfully capture a firm's attention related to ESG. In particular, we find that, in the category of "Environment," firms in the Utilities and Mining, Quarrying, and Oil and Gas Extraction sectors spend the most time thinking about environmental issues. In the "Labor" category, the most time is spent by firms in the Educational Services and Health Care and Social Assistance sectors. This screen supports the idea that firms in industries most exposed to these issues read more, a reasonable assumption for a measure of attention.

We next relate our measure of ESG attention to firms' and investors' ESG outcomes. Such an exercise not only helps us understand whether such a measure could directly have useful applications for ESG-oriented investors, but also sheds new insights on the economic contents of the ESG ratings used in practice. We first examine how ESG attention by *firms* links to ESG-related outcomes. Second, we examine reading by *investors* and ESG-related voting and investing. Finally, we examine the co-movement between investors' reading and firms' reading. It is often argued that investors prefer firms that are more socially responsible and even influence firms directly on such matters.[4] Beyond confirming existing findings, our approach allows us to "quantify" the landscape of ESG investors.

We begin first by examining whether our ESG attention indicators are meaningfully associated with *real* firm outcomes. Before describing our research goals, it is important to emphasize what our research goals are not. Our aim is not to argue that it is solely the attention we observe "causes" specific actions but merely that we *measure* an economically meaningful process within the firm. Also, we believe attention is the most accurate word for our findings, but our goal is not to isolate attention from other mechanisms such as "learning" or another activity within the firm. Under any such interpretation, our general claims will hold regarding

topics like corporate philanthropy and community engagement. See, among others, the report by Koller, Nuttall, and Henisz, "Five Ways that ESG Creates Value," *McKinsey Quarterly* (November 14, 2019). Also, Table 1 in Matos (2020) offers a useful classification.

[4] Please see, for example, (Dimson, Karakas, and Li, 2015, Krüger 2015, Fernando, Sharfman, and Uysal, 2017, Jagannathan, Ravikumar, and Sammon, 2017, Riedl and Smeets, 2017, Starks, Venkat, and Zhu, 2017, Dyck, Lins, Roth, Towner, and Wagner, 2018, Barber, Morse, and Yasuda, 2019, Dyck, Lins, Roth, and Wagner, 2019, Hartzmark and Sussman, 2019, Chen, Dong, and Lin, 2020).

the usefulness of our indicator or what we learn about ESG indicators more broadly. With this caveat in mind, there are three possible relationships between ESG attention and ESG performance. First, rank-and-file employees of an organization may consume information about ESG merely as a hobby, and attention to ESG issues may have no bearing on the ESG-related or overall performance of the organization. There are two alternatives to this null hypothesis. First, we may find a negative relationship between ESG reading intensity and ESG performance. This could arise if organizations read about ESG ahead of impending negative environmental or social news. Second, we may find a positive relationship if organization members read about ESG either ahead of publicly-released positive news, or if they consume information that they might need (legal advice, investor relations, or just topical knowledge) ahead of actions they take to improve the firm's ESG performance.

In a firm-year panel, we construct three variables: (a) the count of production facilities upgraded with green technologies as reported to the Environmental Protection Agency's (EPA) Toxics Release Inventory (TRI) program; (b) the benefits offered to employees in the Internal Revenue Service's (IRS) Annual Report of Employee Benefit Plans (via IRS Form 5500); and (c) the number of penalties enacted by Occupational Safety and Health Administration (OSHA). Where possible, our analysis includes private firms, reflective of the coverage low-cost digital trace data provides. We also study alternatives of these outcome variables, such as pollution production or OSHA inspections. We show that spikes in attention to ESG-related topics are positively correlated with improvements in real outcomes, whereas attention on topics unrelated to ESG is insignificant. Further, our measures win in a horse-race against existing ESG measures in predicting real outcomes.

We next examine ESG ratings and firm attention. We presume an ESG attention measure allows us to gauge whether the ESG rating captures attention within the firm at all. In a firm-time panel, we test whether firms with high reading activity tend to see *improvements* in ESG scores. We examine three popular ESG ratings. We first connect our ESG attention measure to Refinitiv's ASSET4 ratings. We find a one standard deviation increase in ESG reading is associated with an average increase in the ESG Combined score of 17.75%, within firm and within time. Furthermore, we find non-ESG reading intensity tends to either be not at all or negatively correlated with a firm's future ESG performance. This suggests our selection of ESG topics is meaningful. Refinitiv also has two sub-components of ESG Combined score: the overall "ESG" score, in which Refinitiv scores the firm based on perceived ESG practices, and a "Controversies" prediction score, which refers to the perception of the firm in global media sources. We find that the positive association of ESG attention and ESG scores arises primarily from the "Controversies" score.

Next, we conduct the same exercise using data from RepRisk AG. Here, we seek to predict the "RepRisk Index" score, which measures the severity of ESG-related risk in the current month based on textual analysis of news and regulatory filings. At the firm-month level, we find Spikes in ESG-related reading correlate with lower RepRisk scores, with a one standard deviation increase in ESG reading is associated with an average decrease in the "RepRisk Index" score of 5.22%, within-firm and within-time. Interestingly, we find the RepRisk Index is somewhat correlated with non-ESG attention as well, suggesting the news-based index may be less clean.

Our final firm-level test examines ESG reading intensity and Morgan Stanley Capital International (MSCI) KLD ratings. We first use the adjusted KLD score most widely used in the literature (Albuquerque, Koskinen, and Zhang, 2019, among others). We show there is no relationship when conducting analysis within-firm and within-year. One possibility, as noted in Chatterji, Levine, and Toffel (2009), is that KLD's ratings do not reflect reality well. Another possibility is that the convention adopted in practice is sub-optimal. Thus, we present two alternative weightings of the KLD rating which empirically are more closely linked to firm attention.

The second part of our paper focuses on institutional investor-related tests. We build measures of reading intensity for major asset management firms that parallel those for firms. Here, we examine the two ways investors engage with firms: voice and exit (for a recent survey, see Becht, Franks, and Wagner, 2019). We first examine whether investors vote in a more pro-social manner if their intensity of reading on ESG issues spikes prior to a shareholder meeting. Indeed, we find that investors' ESG reading intensity predicts that investors vote in a more ESG-friendly way, even after controlling for the ISS recommendations or management recommendations. We next examine whether investors that read more about ESG issues are likelier to compose their portfolio of ESG-performance-linked stocks. Indeed, in portfolio-level results, we find a positive correlation between investors' ESG reading intensity and the shares of stocks with high ESG ratings. Turning to an investor-stock-quarter analysis, we find that when institutional investors read more intensively about ESG issues, they are more likely to pick up and less likely to sell-off completely stocks that have better ESG ratings. The result holds for specifications saturated by investor-by-quarter fixed effects, a variety of stock characteristics and controlling for investor's reading intensity levels that are unrelated to ESG.

In our final tests, we study the relationship between investors' and firm's ESG-related reading intensity, what we call their joint dynamics. The association between investors' and firm's ESG reading at a topic-quarter level shows that, when top five investors experience an increase in reading intensity on a specific ESG topic, there is a 5.9% higher likelihood that the firm's reading also jumps on that topic, compared to 2.1% for the other investors. We then provide plausibly causal evidence that investors engage and even influence firms on ESG reading. We show that, when investors get distracted in unrelated industries (as in Kempf, Manconi, and Spalt, 2017), the positive association between investors' and firm's ESG reading is attenuated. Here, we calculate each investor's "influence" based on the strength of the co-movement in their attention to ESG and the portfolio firm's attention to ESG. Contrary to the view that only the largest asset managers/owners matter (BlackRock, Vanguard and State Street), we find that several other major funds appear partially responsible for influencing corporate attention to ESG issues. Of special note, a substantial fraction of investors do not impact their portfolio company's ESG attention positively.

These findings contribute importantly to the literature on ESG investing. In an important review, Matos (2020) documents: (1) the rise of interest in ESG issues, particularly climate change, and (2) a lack of consensus among academics as to what the core issues of ESG are and how it should be measured. Many studies note the disagreement across ratings.[5] Our measure contains a new source of information not contained in existing ratings, and allows us to reinterpret ratings by their link to a firm's *attention* to ESG issues. As ESG ratings are also subject to firm's "greenwashing" (Yang 2019, Raghunandan and Rajgopal, 2020), our

[5] Following this latter thrust, Gibson, Krüger, Riand, and Schmidt (2019) find that the average correlation between ESG ratings from six different data providers was less than 50%. Chatterji, Durand, Levine, and Touboul (2016) attribute part of the observed disagreement in ESG ratings to providers' different definitions of firms being socially responsible. Berg, Kolbel, and Rigobon (2020) point to scope, measurement, and weights as sources of divergence among ESG ratings.

ESG measure can shed some light on this issue.[6] Recently, Cao, Titman, Zhan, and Zhang (2020) advance a measure of how investors pay attention to ESG as implied by their holdings of ESG-linked stocks. Though different in approach, our paper and theirs are closest in purpose. However, our novel attention-based measure offers some advantages.[7] While currently not implemented by industry practitioners to the best of our knowledge, it can one day be integrated into practice. Finally, more broadly than ESG investing, having a methodology of measuring business sector attention to social issues is important as society faces dramatic challenges such as climate change (Nordhaus 2019) which require broad awareness and coordination. Our methodology can be used to assess the attentional impact of major policy announcements or initiatives.

We also contribute to the broader literature on the link between agents' attention and their actions. Ocasio (1997) proposes the attentional view of the firm, in which a firm's actions are the result of how they channel and distribute the attention of their decision-makers. Modeling attention allocation has emerged as a critical thrust of recent advances in theory in organizational economics (see Dessein, Galeotti, and Santos, 2016, Dessein and Prat, 2016, Kacperczyk, Van Nieuwerburgh, and Veldkamp, 2016, among others). Empirically, this view has been expressed in many fields of study, including firm growth (Joseph and Wilson, 2018) and innovation (Li, Maggitti, Smith, Tesluk, and Katila, 2013). Choi, Gao, and Jiang (2020) shed light on investors' attention and actions about global warming, although they examine retail investors in aggregate using Google Trends. Iliev, Kalodimos, and Lowry (ILK, 2020) study the Securities and Exchange Commission's (SEC) EDGAR log and how it is linked to corporate governance. Compared to ILK (2020), we focus on both firms and investors and study a broader set of "E" and "S" content beyond just "G." To the best of our knowledge, there is no other approach to study firm-level ESG attention and information acquisition on a broad basis. Our approach may be useful as a testing ground for economic theories across various fields within economics.

[6] Furthermore, a recent paper by Tang, Yan, and Yao (2020) points to the bias of rating agencies, which is that firms connected to rating agency through institutional ownership receive higher ESG ratings.
[7] We do measure ESG attention directly and, in doing so, validate their approach, while pushing the empirical envelope through new data and extending scope to study firm – and not only investor – attention and the joint dynamics in ESG attention between firms and their investors.

2. Data.

2.1 Intent data.

For this study, we obtained proprietary data on internet research activity from a business-to-business (B2B) “intent” data provider, The Company, that was also used in Kwan (2020) and Kwan and Zhu (2020). Intent data refers to a recent development in data analytics aimed at gauging a business’s buying interest based on patterns of web content consumption. The premise of intent data is that if an economic agent consumes more web content related to a particular topic, this agent might have an elevated interest or “intent” in procuring a related product or service. The Company who supplied their data tracks organizational-level (tracked by web domain, such as microsoft.com) interest in specific topics at specific locations (for example, Microsoft’s interest in “Python” in Redmond, WA).

The Company whose data we use is a leading provider of intent data. In principle, any business can generate its own intent data by observing the behavior of its website audience. Beyond a single firm, The Company orchestrates a cooperative of contributors under a “give-to-get” model. Co-op members consist of thousands of mainstream business media sites such as *Wall Street Journal*, *Bloomberg*, *Forbes*, *Business Insider* and *CBSi*, along with more specialized groups of sites such as *1105Media*, *ITCentral Station*, and *Questex*. Most sites are anonymous but span a wide range of business functions, such as technology, finance, marketing, legal, human resources, engineering and manufacturing. Co-op members receive some of the services The Company sells.

Co-op members contribute to the pooled dataset via a technology mechanism which shares information about web content consumption, including the external IP address of the network originating the HTTP request and the URL of content accessed. This data is then filtered into domain, location, and topic. To do so, The Company performs two major steps. First, for each visitor of a website, a user profile is generated, via consent-based and anonymized third-party cookies, which, in combination with the external IP address and other proprietary methods, allows the data provider to associate a domain with the profile, when such an association can be inferred. This mechanism is capable of making inferences even when a person works from home or is on a mobile network (see Kwan (2020) for details).[8] Second, the content is tagged for topics. The Company operates a supervised learning algorithm using a hand-picked set of training manuals which have been labeled for topics The Company aims to study. For example, to chart interest in "Cloud Computing," they have assembled a set of 80 to 100 articles that have been labeled as being pertinent to cloud computing. Topics are chosen by The Company either as a result of publisher and customer requests or according to The Company's view on business-relevant topics and issues.

Table 1 presents summary statistics on our intent dataset coverage. Over the course of the five years in our period of analysis from 2015 through 2019, the number of topics identified by The Company has nearly tripled from 2,462 to 6,765. Panel A exhibits a similar pace of growth that arises in the number of web-domains – or business "addresses" - that The Company tracks from 1.7 million in 2015 to 6.9 million by the end of our sample. To get a proper sense of the data, we also report the number of domain-mapped business-related interactions per day which reaches a peak of 686 million in 2017 across the 4.3 million domains and for 3,589 different topics as of that year. Panel B lists the topic taxonomy as of 2019 by themes as defined by The Company. This range of topics indicates the breadth of internet research interests across the domains The Company tracks.

Most of our analysis uses what The Company calls their "Spike" score. [This is our name for The Company's index score and not that of The Company itself.] A firm's Spike score is a weekly index aggregate which measures a firm's topic-level interest and it dates back to May 2015. We document the construction of this measure in Appendix A.1. Scores are produced for those topic-domains in which there is a threshold number of observations in the first and last 3 weeks of a 12-week rolling window. The score runs from 0 to 100, with a score of 50 representing no increase in interest, and scores below or above reflecting falling or rising interest, respectively. The score measures the change in interest in the 3 weeks prior to the preceding twelve weeks, accounting for other firms' relative increases in this same topic over

[8] Whereas some profile associations are made on IP address (Domain Name Service Records), many are not. IP addresses tend to identify relatively large firms and most smaller firms have a DNS record registered by their internet service provider. For smaller companies, associations between profile and company are made through third-party cookies or other identification methods, providing much larger coverage than through IP addresses alone. As cookies persist in a browser, this follows a device through various IP addresses which permits a stable association between website visitor and firm when employees are working from home or traveling.

this same period. This last step is important because aggregate interest in a topic might have increased due to mechanical changes in the topic taxonomy or due to an increase in the supply of publisher content without an actual increase in unique, firm-specific interest. This, along with a variety of other proprietary adjustments, facilitates the detection of genuine bursts in reading interest, rather than mere mechanical increases.

We note two potential concerns with the data. One may wonder how sensitive our analysis is to the construction of the Spike score. Our results are robust to any potential idiosyncrasies of The Company's construction of this index score, as our main analyses hold over a shorter sample when using the "daily aggregates," which are simply counts of articles read related to a specific topic. We focus on this Spike score for two reasons: (1) it has the longest history available, which is crucial for alternative data which typically have shorter sample periods; and, (2) it is the version commercially available and similar in spirit to offerings by other intent data providers. In this sense, our analysis generalizes beyond a single data provider. Another question is whether changes in the data – benign improvements in the topic taxonomy or improvements to algorithms over time – affect our analysis. All our analyses include time fixed effects which should affect all firms or investors in a similar fashion. Moreover, to the best of our knowledge, The Company has remained focused on its core business of sales enablement. Therefore, our results should improve if The Company were to explicitly design its data for financial applications. Finally, we note that our goal is not to validate the specific way this company constructed their Spike measure, but merely to assert the viability of intent data for measuring ESG attention.

2.2 Defining ESG-related topics.

From the several thousands of topics provided, we hand-select topics most relevant to ESG. We find 323 ESG topics in total. In our main analysis, we decompose topics into four ESG categories: Environment, Labor (including Labor Relations, Equality and Diversity), Social (including Customer Relations, Corporate Social Responsibility) and Governance (including Compliance, Corporate Governance, Cybersecurity, Data and Sensitive Information

Protection). Appendix A.1 outlines how we assign ESG topics to the intent data.[9] Table 2 describes the number and examples of topics.

In Panel A of Table 2, we show the number of topics within each category we classify. The category "Labor Relations" contains the highest number of topics while "Compliance" contains the lowest number of topics. The relative differences between compliance and labor relations could be ascribed to one of three possibilities: (1) the composition of The Company's cooperatives (the publishers who contribute content), (2) the composition of content on the internet, (3) or the composition of The Company's topic engine, which might simply have more topics tracked in these areas. In Panel B of Table 2, we show ten examples of topics from the proprietary dataset within each of the four ESG dimensions to illustrate the categorization approach. For example, "Labor" covers topics in both "Labor Relations" and "Equality and Diversity" categories, including "Diversity Recruiting," "Employee Safety," "Equal Employment Opportunity," "Equal Pay/ Comparable Worth," and "Gender Equality."[10]

We report the number of topics, domains, and domain-mapped business-related interactions per day among ESG topics in Panel C of Table 2. The counts can be compared against those of the universe of all topics reported in Panel A of Table 1. ESG topics as we have defined them represent a substantial fraction of all interactions per day ranging from a low of 6.74% in 2016 to a high of 9.40% in 2017.

In Panel D of Table 2, we list the ten NAICS-2 industries which have the highest percentage of reading intensity across each of the four ESG dimensions during our sample period, with reading intensity here defined as the ratio of total record of topics in that dimension to the record of all topics. The industries topping these lists are intuitive. Turning to "Environment" topics, the industry with the highest reading intensity is Utilities, followed closely by Mining, Quarrying, and Oil and Gas Extraction. The second dimension we show is "Labor." The industry that reads most intensely on these topics is Educational Services,

[9] Alternatively, we also assign topics to nine categories: Compliance, Corporate Governance, Customer Relations, Cybersecurity, Data and Sensitive Information Protection, Environment, Equality and Diversity, Labor Relations and Corporate Social Responsibility. Our nine categories are in the spirit of the common classifications in industry ESG ratings, though we recognize that there is some discretion in our choices.

[10] Three of these topics seem somewhat related; diversity, gender equality, and equal employment opportunity all refer to equity regardless of race or gender. The Company accounts for correlations between topics through its topic engine. If an article pertains to both diversity recruiting and equal employment opportunity (EEO) equally, the article will be given a weight of 50% for both topics. In this way, double-counting concerns are mitigated.

followed by Health Care and Social Assistance, and Accommodation and Food Services, industries which are all labor-oriented. The third dimension is "Compliance," of which the Finance and Insurance and Professional, Scientific, and Technical Services industry reads with the most intensity. The last dimension we show is "Data and Sensitive Information Protection." The top industries are Finance and Insurance, Professional, Scientific, and Technical Services, and Health Care and Social Assistance which are all sensitive to data and information privacy issues. The results show that our ESG measures can reliably rank industries in terms of their attention to ESG.

Finally, we conduct an analysis to investigate alternative interpretations regarding our ESG intensity measures. First, while we believe ESG reading intensity captures reading by employees about ESG issues, one might conjecture our measures instead capture reading about their own company or their company's peers in the news, in general. If so, one might wonder what added value reading measures capture beyond the company news itself. A second consideration that affects our interpretation is related to who within the firm reads. We cannot know whether such attentional activity emanates from the grassroots efforts of employees or from a top-down directive by management.[11] We can, however, investigate what departments within the firm is associated with this reading intensity, whether sales and marketing or investor relations, for example.

To shed light on these concerns, we conduct supplementary analysis using a short sub-sample of granular "event" data made available by The Company. An event is defined as an individual instance of an article being consumed, while tracking the anonymized reader across websites. It is a clinical analysis given our sample period (2020 through May 2021) is too short to represent for our full sample tests. With respect to the question of firm-specific news, we are permitted the ability to understand (1) the websites that generate the ESG content, and (2) within the site, whether the specific article was relevant to a specific firm. We first calculate the portion of ESG reading coming from a financial newspaper, where presumably firm-specific news is most likely. We find that 23% of ESG content comes from financial newspapers. Of content in financial newspapers, a quarter (or 5.6% of the total) is a firm-

[11] Indeed, the case of Google and the cancellation of Project Maven, a U.S. Department of Defence artificial intelligence project to facilitate drone strikes, suggests that calls for ESG can come from the bottom-up, employee-driven effort.

specific article. Therefore, we conclude that this notion can explain only a limited fraction of the reading intensity. We conduct an additional analysis on the estimated functional area of the reader based on the types of websites the reader typically visits. The functional area is conceptually similar to a department within the firm. We find that no single functional area dominates ESG reading. Rather, ESG reading is conducted by the department within a firm most relevant. We also ask how many "profiles" (e.g. internet users) within the same firm tend to read an ESG topic. We find that the average ESG topic has more users reading it than the average topic, suggesting that ESG is read broadly within the firm.

Additional details about these supplementary experiments are available in Appendix A.2.

2.3 *ESG ratings and other data.*

Our first source of ESG rating data is *Refinitiv ASSET4* (formerly, Thomson Reuters' Asset4 ESG database). Refinitiv collects ESG-related information of publicly traded firms from public sources, such as annual reports, Corporate Social Responsibility (CSR) reports and non-governmental organization (NGO) websites. Then Refinitiv captures and calculates over 450 company-level ESG metrics and combines them into ten main categories. The weighted average of ten category scores finally formulates an ESG score which reflects the firm's annual relative ESG performance. Refinitiv also provides an ESG Combined score which is discounted for significant ESG controversies. In this paper we rely on ESG Combined, ESG, and ESG Controversies scores to measure the firm's ESG performance. Higher scores indicate better ESG performance. We use data from May 2015 through year-end of 2018 for our analysis. We start from 2015 as this is the first year covered by The Company.

We next obtain monthly ESG-related risk data from *RepRisk*. RepRisk differs from Refinitiv and KLD (discussed below) in that it relies more on a computerized, systematic approach. RepRisk scours the internet for regulatory filings and news articles in multiple languages, scouring tens of thousands of sources. When its algorithms screen an event damaging the firm's ESG reputation, it applies a human analyst to verify the information and enter it into its database. The data are then used to compute a monthly RepRisk index per firm. We mainly use monthly current RepRisk Index (what we will call *Current RRI*) to measure the firm's ESG-related risk and a higher index means more exposure to ESG-related risk. The sample period starts from May 2015, which is the first month covered by The Company. It ends in August 2018, the last month we can obtain data from RepRisk as of writing.

Finally, we obtain firm-level rating data by MSCI ESG's *KLD*, which mainly rates firms based on a wide range of strengths and concerns across seven categories: community, diversity, employee relations, environment, governance, human rights, and product. Specifically, each KLD category has a variety of areas, and if a firm has strengths/concerns it will be given one point in that area. We use data from 2015 to 2018, the last year covered by Wharton Research Data Services (WRDS) as of writing.

We also obtain several real ESG performance data. First, we use EPA's Toxic Release Inventory to measure emissions and Pollution Prevention (P2) database to identify firms that have made significant efforts to upgrade their technology and become more environmentally friendly. We have this data from 2015 to 2018. For the latter, we obtain the number of facilities reporting newly implemented source reduction activities each firm-year and include both public and private firms in the analysis. Second, we obtain the IRS's Form 5500 Information from Axiomatic Data from 2015 to 2019. Specifically, on an annual basis, we observe employer contributions to employee pension plans as well as indicator variables across several types of employee benefits for Russell 3000 firms. Finally, we gather inspection and penalties data from OSHA during the period from 2015 to 2019. Specifically, we count the number of penalties or unprogrammed inspections for each firm-year and include both public and private firms in the analysis.

We also obtain ESG disclosure data for public firms from several sources, including Global Reporting Initiative (GRI) Sustainability Disclosure Database, CSRwire, Corporate Register, and Responsibility Reports. In particular, we construct an indicator variable as to whether the firm has issued a sustainability report on CSR/ESG for a given year from 2015 to 2019. Our institutional holdings data comes from FactSet Ownership. Our mutual fund voting data comes from Institutional Shareholder Services (ISS) Voting Analytics. We obtain company identifiers and financial data from Standard and Poor's Compustat and stock market data from the Center for Research on Security Prices (CRSP).

3. Assessing the ESG reading intensity of firms.

In this section, we discuss our Spike measure and assess whether it captures economically meaningful firm attention to ESG. We first present graphical anecdotal evidence. Building off of prior studies, such as Akey and Appel (2019), Li and Raghunandan (2020), and Heath, Macciocchi, Michaely, and Ringgenberg (2021), we then analyze three sets of real outcomes: firm's efforts at pollution prevention, employee benefits and OSHA penalties.

Our central hypothesis is that if attention to ESG issues is reflective of the firm's heightened attention toward corporate social responsibility, we should expect associative changes in these real outcomes. To be clear, our argument is not a statement of causality in the sense that acts of reading specific content captured in this Intent dataset are causing the firm to act, but rather that the Intent data captures economically meaningful process of attention paid within the firm. First, if firms are learning or endeavoring to improve their pro-social image, these changes will be positive. Alternatively, if firms are reading in relation to impending controversy, we expect a negative relation. Empirical and experimental evidence suggests that agents asymmetrically pay attention when taking positive actions (for example, Akepanidtaworn, Di Mascio, Imas, and Schmidt, 2019) and when news is good (please see Karlsson, Loewenstein, and Seppi, 2009, Sicherman, Loewenstein, Seppi, and Utkus, 2016, Olafsson and Pagel, 2017, Pagel 2018). Hence, we believe that it is likelier than not that firms pay more attention when they are aiming to conduct change rather than reading ahead of bad news. Finally, if firm employees are reading for leisure, we may expect no relation at all. Of course, all three of the above motivations may be at play when consuming ESG-related internet content, so our analysis speaks to the average relation between reading activity and real outcomes.

After establishing the empirical relationship between our measures and real ESG outcomes, we ask whether ESG ratings (1) predict the same outcomes, and (2) capture ESG attention within the firm. Presuming that one is convinced by our evidence that our ESG measure reflects actual attention by employees within the firm, this exercise helps us assess the economic content of ESG ratings. While we believe our reading intensity measures have numerous advantages in their own right, to the extent we observe a comovement between reading intensity and firms' ratings, we interpret that as suggestive evidence in support of third-party ratings. More importantly, this exercise helps us understand what types of variation in attention within the firm the ratings capture.[12]

3.1. Distribution of the Spike Score across firms and years.

In Table 3, we report the distribution of weekly Spike scores for all CRSP firms by year. In each year, we report a variety of quantiles of the distribution of Spike scores. Panel A shows the distributions for all topics, and Panel B shows those for ESG topics only. Recall that The Company regards a Spike score above 60 as a significant increased interest in a topic-domain, as measured based on the prior three weeks relative to the 12 weeks before that with some adjustments. From the distributions, we can see that a Spike score in higher ends of distributions (95^{th} or 99^{th} percentiles) are stable across different years and that they are similar for all topics or ESG topics. A Spike score of 80 as a threshold for high interest in a topic in a given firm-week lies between the 95^{th} and 99^{th} of the firm-week distribution. While admittedly an arbitrary threshold, the number of Spike scores that are at least 80 seem to capture high reading intensity or attention. Hereafter, a Spike refers to a Spike score above 80. Our usage of an arbitrary threshold is similar to Ben-Rephael, Da, and Israelsen (2017).

3.2 Two illustrative examples.

We present two pieces of anecdotal evidence to build the interpretation that firms rationally pay attention to ESG-related issues because of their business-related exposure. First, we study how firm attention to ESG changed around the date on which BlackRock's Larry Fink released an open letter on January 14, 2020 to CEOs, which discusses the need for firms to focus improve their ESG performance. The letter primarily focuses on environmental issues and, to a lesser extent, on labor issues. In Figure 1, we plot the attention of firms with the highest percentage share by Blackrock as their investor versus firms with the lowest share or no ownership by Blackrock. Blackrock-owned firms read more about the environment (and to a lesser extent labor) in the next two days, relative to firms with low or zero Blackrock ownership. The top-left and top-right panels of Figure 1 illustrate these effects with the red line representing the higher reading intensity among BlackRock-owned firms and the blue line for those with firms with zero BlackRock holdings. Governance and data security issues in the

[12] We acknowledge that we do not cover all third-party ratings. We also acknowledge that third party ratings may have variations of their own rating (an environment-specific component, for example). In this paper, we focus on the versions of the ratings as used by academic literature.

bottom-left and bottom-right panels, by contrast, exhibit no divergence between firms in either group, as expected.

Second, we analyze the association between a firm's Environment-related reading and their stock market performance during climate change shocks. We obtain the climate change news data from Engle, Giglio, Kelly, Lee, and Stroebel (2020). We discuss their data methodology in Appendix B. After estimating the climate beta, we calculate the percentage of reading that is related to Environment topics during the sample period. We use both weekly Spike scores and daily aggregates from The Company in the analysis. Finally, we sort the percentage reading into quintiles and plot the average climate beta for each quintile in Figure 2. In an un-tabulated analysis, we also try to sort Environment-related reading into deciles and find similar patterns. As seen in Figure 2, the average climate beta increases with the percentage of Environment reading. This pattern is robust to whether we use Spike data over a longer period or the underlying daily aggregate data over a shorter period. That the firm's attention is correlated with their stock price sensitivity to climate change news supports the interpretation that reading about an ESG topic is related to business needs.

3.3. ESG reading intensity by firm and real ESG performance.

We now turn to our tests of ESG attention on real outcomes (pollution prevention, worker benefits and OSHA penalties/inspections). We interpret improvements in these measures as improvements in ESG policies at the firm. We then discuss robustness checks related to these analyses. Our analysis generally aligns reading in year *t* to data the public learns of through regulatory filings in year *t+1* but which are filed for the year *t*. We later discuss sensitivity analysis to this empirical design choice, where we consider both one year ahead to make our analysis fully predictive and one year "behind" analyses as a placebo.

In Panel A of Table 4, we test whether and how a firm's ESG-related reading intensity is associated with their environmental performance. First, we obtain pollution prevention data from the EPA Toxic Release Inventory (TRI) Program. Our outcome variable is *Log(1 + # Green facilities),* where "Green facilities" is the number of facilities reporting newly implemented source reduction activities for any firm-year. A higher count reflects a higher number of facilities upgraded by a firm in a given year for improved environmental performance. We add firm fixed effects across all models to control for trends in improvements over our sample that are occurring anyway, as well as year fixed effects to control for any year effects. In Models (1) and (2), we include both public and private firms. The estimated coefficient of *Log(1+Spikes^{ESG})* is positive and statistically significant with a *p*-value of less than 1%. The economic magnitude is notable: a one standard deviation increase in ESG reading is associated with an increase in a firm's greenness of 0.0355 (1.224 × 0.029), which represents 12.24% of its standard deviation within-firm and within-time (0.290). When we analyze only public firms in Model (3), the estimated coefficient of *Log(1+Spikes^{ESG})* remains positive. Although our reduced sample leads to lowered statistical significance, the implied economic significance gets even larger.

Finally, in Models (4), (5) and (6) we control for three annual ESG ratings we study in this paper. Coefficients are standardized for ease of interpretation. They are also aligned in time such that the information in the rating should be contemporaneous to the data reported to the EPA, even though that risks a look-ahead bias in favor of the third-party ESG rating. For example, for the year 2018, we use the 2018 KLD rating, which is released in 2019. Despite this disadvantage, the results suggest that the estimated coefficient of *Log(1+Spikes^{ESG})* remain positive and statistically significant at the 5% level, and the ESG reading intensity has both greater statistical and economic significance than ESG ratings. Our finding implies that ESG ratings are not linked to this real outcome, while our ESG attention measure is. To mitigate the concern that the number of facilities may change over time, we use the percentage or indicator of green facilities as alternative outcome variables and obtain similar results (see Appendix Table C.1).

In Panel B, we repeat the above analysis using pension contributions from the IRS Form 5500 provided by Axiomatic Data, which we have for publicly listed firms only. Our outcome variable is the log-amount of employer pension contributions scaled by the sample period's average assets by firm. Again, our findings are similar, proving that not scaling does not hurt our results. The estimated coefficient of *Log(1+Spikes^{ESG})* is positive and statistically significant.

Finally, in Panel C, we aggregate OSHA penalties in a firm-year, scaled by the number of establishments for a given firm, as the outcome variable. We also try the log-number of OSHA penalties or dollar amounts of penalties (scaled by firm size) and get similar results in

Appendix Table C.1. The estimated coefficient of *Log(1+Spikes^{ESG})* is negative and statistically significant. A similar story emerges as in Panel A and B. In Models (1) to (3), spikes in ESG-related reading are associated with reduced OSHA penalties. The last three columns present a horse-race between our measure and other ESG metrics. Our ESG metrics are statistically and economically significant—while existing ESG ratings are not—in explaining OSHA penalties.

3.4. Robustness checks on the link between ESG reading and ESG real outcomes.

We conduct multiple variations of this analysis, all of which are reported in Appendix C. First, our analysis is predictive or possibly similar to "now-casting" in that it relates the year *t* reading to the year *t* outcome (which the public learns in year *t+1*). This exercise is useful because it suggests that people read during the time that they may be learning or taking actions. However, we alter the timing of our analysis so that the analysis is one year ahead, finding similar results with marginally reduced statistical significance. We also align year *t+1* reading with year *t* outcomes, finding null results for our three real outcomes.

Second, for each of our outcomes, we perturb the outcome to either align with existing literature or as a robustness check. For the EPA analysis, we also conduct analysis on total emissions per sales.[13] That is, conditional on the level of production, how much pollution does a firm produce? Controlling for chemical-year joint fixed effect, we find that pollution-per-sale drops when firms read more about ESG. For our analysis on employee contributions to benefit plans, we also examine the total number of worker benefits. We find that the number of benefits offered by the firm tends to increase when firms read more about ESG, controlling for non-ESG reading. Third, for OSHA, we also look at unprogrammed or irregular inspections. We show that unprogrammed inspections, like penalties, also drop. All analyses reveal qualitatively similar findings that ESG-specific reading is associated with our real outcomes.

Third, we conduct two tests that show further refinements of our measure lead to better results. First, we horse-race our measure of spikes on ESG, defined at the threshold of 80, against the same measure calculated at lower thresholds. We find that a Spike score of 70 is also positively linked to real ESG outcomes but it loses in a horse-race to our Spike score of

[13] Studies differ in the functional form of the outcome variable. Some studies (See Heath, Macciocchi, Michaely, and Ringgenberg, 2021 for example) focus on the firm-year and others (for example, Akey and Appel, 2019) focus on the firm-chemical-year as a unit of observation. We conduct analysis at both units, finding qualitatively similar inferences.

80. This would be expected if our reading intensity score is correlated with firm action – the higher the score, the more strongly linked to outcomes. Also, we further subdivide the ESG reading into nine components. Our expectation is that, if our measure of ESG attention isolates meaningful activity at the firm level, ESG reading most relevant to the associated ESG outcome should drive this empirical relationship. In line with our expectation, we find that environmental performance is best explained by environmental reading, while worker benefits and OSHA penalties are strongly correlated instead with labor-related reading intensity. Unrelated topics are generally statistically less relevant or empirically unrelated. We also conduct our analysis using the daily aggregates. We lose one year of data availability, but when we simply count the number of articles deflated by total assets, we obtain a similar result that ESG-relevant reading is positively linked to the ESG outcome while non-ESG reading is not.

Fourth, we examine how firms' sustainability disclosures relates to real outcomes. We form a *Disclosure* dummy variable indicating whether or not the firm has an ESG sustainability report each year. In Appendix D, we present analysis relating disclosure to our real outcomes, and disclosure interacted by our measure of reading intensity with our real outcomes. We find that disclosure by itself is not positively correlated with improvements in firms' ESG outcomes. This non-result echoes concerns of consumers and critics of ESG ratings in that disclosure – the mere claim of a policy – may not result in actual change at the firm. However, for most outcomes, interacting $Log(1+Spikes^{ESG})$ with *Disclosure* is positive, suggesting a stronger relationship of disclosure with real outcomes when firms exhibit heightened levels of attention. This suggests our metric may be a useful channel-check against firms' ESG disclosures.

3.5 ESG reading intensity by firm and ESG rating by Refinitiv.

To examine the relation between our measure of ESG attention and real ratings, we first obtain ESG ratings data of all covered firms from Refinitiv. Our match delivers 6,759 firm-years among 2,164 unique firms. Panel A of Table 5 shows summary statistics of three original ESG scores by Refinitiv and the main variables used in this regression analysis. The combined scores for Refinitiv averages at 45.710 on the 0 to 100 scale with a standard deviation of 16.407. However, we will also compare our magnitudes to *within*-firm variation in some specifications as this distribution captures mostly cross-sectional variation.

In Panel B of Table 5, we test whether and how a firm's ESG-related reading intensity is associated with their ESG rating by Refinitiv. In this panel, ESG-related reading intensity is captured by the number of Spike scores that reach at least 80 in a given firm-year and we take the logarithm of that count as the main variable, or $Log(1{+}Spikes^{ESG})$. The regression specifications add year fixed effects across all columns, industry fixed effects in one and firm fixed effects across most others.

The summary statistics for $Log(1{+}Spikes^{ESG})$ imply a mean of 3.937 and an associated interquartile range of 3.497 to 4.820. This mean represents the equivalent of 50.3 Spike scores of 80 among the 2,164 firms in a given year with an interquartile range from 32 to 123. For the first four columns of Panel B, we use the level of ESG Combined score ($Refinitiv^{Combined}$) as the outcome variable of interest. Recall that the Combined score is the firm's relative ESG performance across more than 450 metrics along with its ESG Controversies overlay. In Model (1), we add industry fixed effects to control any industry invariants. The estimated coefficient of $Log(1{+}Spikes^{ESG})$ is positive and statistically significant at the 1% level. When adding firm fixed effects in Model (2), the estimated coefficient of $Log(1{+}Spikes^{ESG})$ remains positive, abates somewhat in magnitude, and is statistically significant at the 1% level. In Model (3), we further control non-ESG reading intensity $Log(1{+}Spikes^{Not\ ESG})$ and the estimated coefficient of $Log(1{+}Spikes^{ESG})$ becomes larger. The economic magnitude implied by this estimated coefficient of 1.031 is also large: a one standard deviation increase in ESG reading is associated with an increase in the ESG Combined score of 1.365 (1.324 × 1.031), which represents 17.75% of its standard deviation within-firm and within-time (7.690). In Model (4), we decompose ESG reading intensity by Environment, Labor, Social and Governance to understand what drives the positive association. The result shows that, for the ESG Combined score, the Social category dominates.

Models (5) and (6) examine the level of the ESG core ($Refinitiv^{ESG}$) and ESG Controversies scores ($Refinitiv^{Contro}$) as outcome variables, respectively. The coefficient of $Log(1{+}Spikes^{ESG})$ remains positive but it is only statistically significant and at the 5% level for the ESG Controversies score. These results imply that higher ESG-related reading intensity is associated with improvements in the ESG Combined score and that it is, in turn, concentrated in the association with the ESG Controversies score. Moreover, the component of the Spike linked to the Social category is most prominently associated with the ESG Combined score.

3.6 ESG reading intensity and RepRisk risk ratings.

We next obtain the monthly RepRisk index (*Current RRI*) score for all covered firms with the identifier RepRisk ID merged with public firms. After merging, we have 59,413 firm-months and 1,735 unique US firms. In Panel A of Table 6, we show summary statistics of *Current RRI* and main variables used in the analysis. The mean *Current RRI* is 12.827 with a standard deviation of 12.097. The distribution of scores across firm-months is left-skewed with lots of zero values (at least 25% of the observations) and a maximum value of 55. In Panel B of Table 6, we test whether and how a firm's ESG-related reading intensity is associated with a firm's ESG-related risk. We use *Current RRI* as the outcome variable and add month fixed effects across the first four columns. Models (5) and (6) are firm-year specifications with year fixed effects and are directly comparable with those in Table 5. In Model (1), we only add industry and year fixed effects and the coefficient of $Log(1+Spikes^{ESG})$ is positive and statistically significant at the 1% level. Our explanation is that RepRisk is a firm-specific index for ESG-related risks and firms with relatively high ESG risks may pay more attention and efforts to ESG, which drives the positive cross-sectional association. In Model (2), we add firm fixed effect instead and the coefficient of $Log(1+Spikes^{ESG})$ is negative and statistically significant at the 1% level. This is consistent with the positive sign in Model (6) for the Refinitiv Controversies score. The economic magnitude implied is notable: A one standard deviation increase in ESG reading is associated with an average decrease in *Current RRI* of 0.337 (1.037 × 0.325), which is 5.22% of its residual standard deviation (6.457). In Model (3), we control for non-ESG reading intensity and the result is similar. In Model (4), we decompose ESG reading intensity by Environment, Labor, Social and Governance to understand what drives the negative association. The result shows that "E" and "S" categories matter the most.

To make a head-to-head comparison with annual Refinitiv or KLD rating, in Models (5) and (6) we aggregate the firm-month observations to firm-year by using the year-end *Current RRI*. The result is similar except that when controlling non-ESG reading, the coefficient of $Log(1+Spikes^{ESG})$ becomes insignificant. This is somewhat surprising given the resilience of the measure in Table 5. It does imply that there is important information in the higher-frequency monthly RepRisk scores that is lost by constructing a lower-frequency annual level of granularity.

Another question is whether employees read about ESG topics because their firm is in the news for negative ESG-related news. Such a scenario would be of no harm to our central claim that our ESG attention measure meaningfully predicts other measures. That said, we also conduct the same analysis controlling for ESG news about the firm using data from Ravenpack, a database that collects news from media and press releases and that provides tone and topics for a series of news topics. We define ESG news as news pertaining to specific topics identified by Ravenpack as being ESG-relevant (see Hafez and Gomez, 2019, Cui and Docherty, 2020). We find similar results when we control for the number of negative news articles explicitly, or if we simply remove those firm-months from the sample. We conclude firms are not merely reading about negative or positive news events about their own firm.

Overall, the above results show that ESG-related reading is negatively associated with the level of RepRisk Index (*Current RRI*), which means ESG-related reading can mitigate a firm's ESG-related risk. The mitigation mainly comes from Environmental and Social categories of reading intensity.

3.7 ESG reading intensity by firm and ESG rating by KLD.

We use KLD (also known as MSCI *ESG KLD STATS*) as our third source of ESG rating data. For these tests, we have 2,462 unique firms. Recall that KLD reports a wide range of strengths and concerns across seven categories: community, diversity, employee relations, environment, governance, human rights, and product. Each KLD category has a variety of sub-areas, and if a firm has strengths/concerns it will be given one point in that area. We sum all strengths and concerns for any firm-year to obtain a count of total strengths (*Str*) and total concerns (*Con*). Following papers such as Albuquerque, Koskinen, and Zhang (2019), our "adjusted" KLD score *KLD1* equals *(Str-Con)/(n_Str+n_Con).* The variables *n_Str* and *n_Con* are the maximum possible number of strengths and concerns across categories and sub-areas, respectively, which may change over time as KLD adds or removes data categories.

We also construct *KLD2* and *KLD3*, which perturb the construction of *KLD1* by only counting a strength or concern when a strength or concern is non-zero, as counting performance against all issues could bias towards large firms (for which KLD may simply have more information gathered and therefore more non-zero ratings) or firms of certain business models where ESG is of concern. *KLD2* defaults a firm with zero strengths or concerns to be 0, while *KLD3* deems the observation as missing. We note that our upcoming analyses hold whether one is looking at the net difference between strengths and concerns or simply at concerns or strengths separately. Our finding is that firms reading more concerns drive our results, which is consistent with some prior literature (Chatterji, Levine, and Toffel, 2009, among others) which argues that concerns measured by KLD are more economically meaningful or KLD records more negative than positive events (Krüger 2015).

In Panel A of Table 7, we show summary statistics of three versions of the KLD score and main variables used in analysis. As expected, the mean relative strength measures for *KLD2* or *KLD3* are much higher at 0.558 and 0.562, respectively, than those for *KLD1* at 0.024. The standard deviations are similarly much higher across firm years. In Panel B of Table 7, we test whether and how a firm's ESG-related reading intensity is associated with adjusted KLD scores. For the first three columns, the dependent variable is *KLD1*. In Model (1), when we add industry and year fixed effects, the estimated coefficient of $Log(1+Spikes^{ESG})$ is positive and statistically significant at 1%. When adding firm fixed effects instead of industry in Models (2) and (3), the coefficient of $Log(1+Spikes^{ESG})$ becomes insignificantly negative. In Models (4) and (5), we repeat the analysis with *KLD2* and *KLD3*. The coefficient of $Log(1+Spikes^{ESG})$ is positive but only statistically significant at the 1% level when using *KLD3* as the outcome variable. The economic magnitude here implied is large: A one standard deviation increase in ESG reading is associated with an average increase in *KLD3* of 5.990 (1.116 × 5.367), which represents 20.09% of a standard deviation of regression residuals (29.808). In Model (6), we decompose ESG reading as in Table 5 and Table 6, where we see the Governance category dominates the overall result.

The KLD rating has received considerable criticism from academic literature. Chatterji, Levine, and Toffel (2009) argue that KLD's ratings are not optimally using publicly available data. Berg, Koelbel, and Rigobon (2020) document that KLD rating has the highest measurement divergence among all ratings they examine. Our results do not invalidate the criticism of a naïve aggregation of KLD ratings, but suggest that some variations of the KLD score may capture specific components ESG attention reasonably well.

3.8 Additional Robustness checks on ESG rating analysis.

Finally, we also conduct analysis using what The Company calls its Daily Aggregates file, which is the input file that underlies the weekly Spike score. As mentioned before, we do not use this for our main analysis because we have a shorter sample period, it is not available commercially, and the Spike score takes certain quality control steps to normalize non-fundamental changes to the data such as the number of contributing publishers. However, in this file, we can count the exact number of records read by the organization pertaining to a particular topic. Using this alternative version of the data, we confirm our main analysis. We present and discuss our results in Appendix C. We also conduct our analysis decomposing KLD into strengths versus concerns. In general, we find evidence that higher ESG reading predicts both strengths and concerns in our *KLD2* and *KLD3* measures, although we find a slightly stronger association with concerns, in line with Chatterji, Levine, and Toffel (2009) and Krüger (2015).

4. Assessing the ESG reading intensity of investors.

We now turn to study ESG reading and the actions of investors. First, we examine whether asset management firms are more likely to vote in an ESG-friendly way when they have greater ESG reading intensity. Our second analysis evaluates links between ESG reading and averages of ESG scores among the stocks held by the asset management firm at the investor-quarter level. We ask whether, during quarters in which there are jumps in ESG reading intensity, we observe changes in the percentages of stocks held by the firm with high ESG scores. We disaggregate this further into the investor-stock-quarter level.

4.1 Investor's ESG reading and fund voting.

We first obtain mutual fund voting data from ISS Voting Analytics during the period from 2015 to 2018. Based on N-PX filings, ISS provides each fund's voting record for each proposal during the shareholder meeting. Using CIK number in N-PX filings, we identify investor domains through CRSP, FactSet and Capital IQ. Based on ISS descriptions of proposals, we only keep those relevant to ESG issues. Then we merge voting data with The Company and align the investor reading 1 month before the shareholder meeting. Our final sample comprises 2,529 unique ESG proposals of 616 firms, and 5,964 funds owned by 373 asset management holding firms (we call these investors).

In Table 8, we present our baseline results. For all models we control for meeting and investor fixed effects. In Models (1) to (5), the outcome variable "Fund vote for ESG" is a dummy variable indicating whether the fund votes for ESG-friendly proposals or votes against anti-social proposals. In Model (1), the coefficient of *Log(1+Spikes^{ESG})* is positive and statistically significant at the 1% level, suggesting that when investors have strong interest in ESG, they are more likely to vote in an ESG-friendly way. In Models (2) and (3), when we further control investors' non-ESG reading intensity or similar "*Fund Vote for ESG*" dummies for ISS or Board recommendations, the statistical significance of ESG reading intensity remains and the economic significance gets larger. In Models (4) and (5), we re-run the analysis for E and S or Governance (G) proposals separately. Surprisingly, we find that when investors pay much attention on G issues, they are more likely to vote for governance proposals, but are less likely to vote for E and S proposals. In Models (6) to (8), the outcome variable *Fund vote with ISS* is a dummy variable indicating whether the fund votes are the same as ISS's recommendation. As we only keep "For" and "Against" for fund voting records, the number of observations drop slightly. The results suggest that when investors have strong ESG interest before voting, they are more likely to vote with ISS's recommendation. This reflects the notion that ISS generally favours good ESG practices.

As some studies use an investor's portfolio ESG holdings to classify socially responsible investors (Hwang, Titman, and Wang, 2015; Brandon, Krüger, and Mitali, 2020; and, Cao, Titman, Zhan, and Zhang, 2020), we control for an investor's portfolio ESG performance at the previous quarter of shareholder meeting in the analysis. Our sample size gets diminished in Appendix E, but the coefficient of *Log(1+Spikes^{ESG})* remains positive and statistically significant at the 1% level. We also find that an investor's portfolio ESG performance measured by KLD or Refinitiv does not predict fund voting in the same way as their ESG reading intensity does (pointing in the wrong direction, even). However, we find some evidence that investors holding stocks with a lower RepRisk index (e.g. better ESG scores) on an equal-weighted basis are more likely to vote ESG-friendly way when they have higher average ESG risks in their portfolio. However, the magnitude of *Log(1+Spikes*^{ESG}*)* is much larger than that of portfolio-weighted ESG scores.

4.2 Investor's ESG reading and portfolio-level ESG performance.

We first obtain holdings data from FactSet which provide the websites for the overall holding firms. After we get investor-level domains, we merge holdings data with The Company to get an investor's quarterly ESG reading intensity. We use all common stocks in CRSP for this analysis. Our final sample starts from the third quarter of 2015 (first complete quarter covered by The Company) and ends at the fourth quarter of 2019 (last quarter of FactSet). The ESG ratings of stocks are calculated before the start of the quarter in which the investors conduct ESG reading.

Based on this sample, we first assess the relationship between an investor's ESG reading intensity *Log(1+Spikes*^{ESG}*)* and the ESG performance of their portfolio. We take an equal-weighted approach to calculate portfolio-level ESG ratings because we do not want the results to be mainly driven by large-cap stocks. We present the results in Table 9. Across all models, the main explanatory variable is investor's ESG reading intensity *Log(1+Spikes*^{ESG}*)*. In each case, we control for that investor's non-ESG reading intensity, as well as quarter and investor fixed effects. We standardize the outcome variables and multiply them by 100. In Models (1) to (5), we use equal-weighted ESG ratings following those that were featured in Table 5, Table 6 and Table 7. Model (1) indicates that an investor's ESG reading intensity is positively associated with the adjusted KLD score (the version commonly used in literature). Models (2) to (4) feature the ESG ratings by Refinitiv. The coefficients of *Log(1+Spikes*^{ESG}*)* are statistically significant at the 1% level for the Refinitiv ESG Combined score and Refinitiv ESG score. The magnitude is noteworthy: based on Model (2), a one standard deviation increase in reading intensity is associated with a 2.333 (1.237 × 1.886) higher average Refinitiv Combined score, or 2.33% of its standard deviation. Surprisingly, we do not find statistically significant results for the Refinitiv ESG Controversies score as we did in Table 5 at the firm-year level of analysis.

Model (5) also reveals that the large increase in ESG reading by the asset manager is not associated with the *Current RRI* from RepRisk. In Model (6), however, we use an equal-weighted *Peak RRI*, which is the maximum of *Current RRI* in the last two years. One can think of this as a high-water mark on the "reputation" in ESG risks that can carry forward over time. The coefficient on *Log(1+Spikes^{ESG})* is negative and statistically significant at the 10% level. In Model (7), we transform the outcome variable to be a percentage of stocks that have positive *Peak RRI* among all those stocks held by the asset management firm that quarter. This threshold measure is negatively associated with *Log(1+Spikes^{ESG})* in a way that is consistent with our findings in Table 6 – asset management firm employees increase ESG reading intensity during quarters in which they have a relatively high fraction of stock holdings with high ESG risks. In Model (8), the outcome variable is transformed to be the percentage of stocks that have *Peak RRI* that is larger than or equal to 50. We use the threshold 50 because it is used by RepRisk to classify stocks with high ESG risk. This is the equivalent to the measure in Model (7) but with an additional threshold. The results indicate that the coefficient of *Log(1+Spikes^{ESG})* remains negative, but it is only statistically significant at the 1% level in Model (7).

4.3 Investor's ESG reading and trading decisions.

To give more direct evidence than our portfolio-level tests, we next turn to investor-stock-quarter levels of analysis using the same sample as above. We define several variables for this analysis. A quarterly measure of investment, $Invest_{ijq}$, is defined as log change in the dollar value of investor *i* holdings for stock *j* in quarter *q* and can be computed as:

$$Invest_{ijq} = log(1 + Holdings_{ijq}) - log(1 + Holdings_{ijq-1}(1 + Ret_{jq})), \quad (1)$$

where the term $Holdings_{ijq}$ is the dollar holdings by investor *i* for stock *j* at end of quarter *q*. A firm's stock return is represented by Ret_{jq}. This captures changes in positions of existing holdings, but we also define different types of investment. "Selloff" is defined as a liquidation of all shares in an existing position. "Decreases" is defined as more than one percent decreases in dollar holdings of a stock. A "Hold" is defined as any change in the dollar value of a position within a one percent change of dollar holdings as of the beginning of the quarter. "Increases" is defined as any change greater than a one percent increase of dollar holdings. Finally, a "Pickup" is defined as a *de novo* investment in a stock that was not held by the investor last quarter. Each of these are defined as indicator variables equal to one if the change in position is equivalent to the definition above and zero, otherwise.

In Table 10, we examine how an investor's ESG reading and a stock's ESG rating affect an investor's overall investment as well as different types of investment. We add various stock characteristics by last quarter-end that could influence an investor's trading decisions: the trailing quarterly stock return and volatility, its market capitalization, momentum, gross profitability, and book-to-market ratio. Volatility is computed from daily returns. Momentum is the return over the past year, skipping one month. Gross profitability is the net income over assets, while book-to-market deflates book value by market capitalization. After requiring stock-level control variables and at least one ESG rating to be non-missing and the stocks to be held by the investor last quarter, there are 7,561,289 investor-stock-quarter observations. The sample diminishes after being merged with ESG ratings data but allows for 3,626 different institutional investors and 2,862 unique stocks.

Panel A exhibits summary statistics of main variables in the analysis. Across the nearly 7.6 million investor-stock-quarter observations, the average position change in dollar value of the typical investor in a typical stock is negative at -1.392 percent per quarter during 2015-2019. There is considerable variation with a standard deviation of 4.022 percent per quarter. We standardize all acronyms for ESG ratings for this analysis and name them $Score^{ESG}$ in Panel B – the heading in each of the four columns represent the ESG rating score that applies. We run regressions for the overall investment. For these tests, the stocks are only those held by the investors last quarter. The variable of interest is $Score^{ESG} \times Log(1+Spikes^{ESG})$. Across all columns, we control for non-ESG reading intensity by investors, stock characteristics and investor-quarter fixed effects. In Model (1), we present our results for adjusted KLD score (*KLD1*). The coefficient of $Score^{ESG} \times Log(1+Spikes^{ESG})$ is positive at 1.200 and statistically significant at the 5% level. The result is similar when we use ESG Combined score by Refinitiv in Model (2). When we use *Current RRI* or change of *Current RRI* in Models (3) and (4), although the coefficient of $Score^{ESG} \times Log(1+Spikes^{ESG})$ becomes insignificant, there is some evidence that investors try to reduce investments in stocks with high ESG risks; especially those with increases in ESG risks – Model (4) has a coefficient of -2.385 of $Score^{ESG}$ for ΔRRI with a robust *t*-statistic of 3.33. The weak link with the Spike score is perhaps indicative that by some measure of revealed preference, investors do not use RepRisk scores as they do KLD or Refinitiv.

In Panel C, we use different types of investment as outcome variables to understand what drives the result. In these regression tests, we use ESG^{Zscore} which is the sum of the standardized adjusted KLD score and ESG Combined Score.[14] Models (1) to (4) present results for stocks held by the investors last quarter. Out of 5,760,778 observations, there are 570,279 "Selloff" events, 2,696,547 "Decreases," 1,333,612 "Hold" events, and 1,730,619 "Increases." The results suggest that "Increases" and "Selloff" dominate. For "Increases," the coefficient of $Score^{ESG} \times Log(1+Spikes^{ESG})$ is positive and statistically significant at 5%, indicating that when stocks have better ESG performance *and* investors have strong ESG interest (proxied by high ESG reading intensity), investors are more likely to increase their positions on these stocks.

In Model (5), we present result for "Pickups" of stocks that were not held by investors last quarter. Among 54,965,093 qualifying observations, there are 610,048 "Pickup" events. The coefficient of $Score^{ESG} \times Log(1+Spike^{ESG})$ is positive and statistically significant at the 1% level, suggesting that when stocks have better ESG performance and investors have strong ESG interest, investors are more likely to establish *de novo* positions in those stocks.

We conduct a few robustness checks for these specifications. Our analyses are robust to interacting ESG attention with these stock characteristics, to the extent one is concerned that these firm characteristics or changes thereof drive changes in the ratings. We can also exact more taxing fixed effect specifications. Pooling all investor-quarter-stock observations, we impose investor-stock fixed effects, finding qualitatively similar results.

From the results above, we conclude that when investors exhibit a strong interest in ESG which is proxied for by high ESG reading intensity, they are more likely to invest in or less likely to sell (especially completely sell off) stocks that have better ESG performance. These results are consistent with Brandon, Krüger, and Mitali (2020) which indicate growing investor preferences for sustainable investing and the resulting price pressure that institutions exert on stocks with good ESG scores. We provide direct evidence between investor preferences and trading on stocks with good ESG scores. Furthermore, among the three ESG ratings we use, investors care most about KLD and Refinitiv when conducting the trading decisions, but there is some evidence of concern for ESG risks (via *RepRisk*).

[14] In un-tabulated results we run analysis for adjusted KLD score and ESG Combined Score separately. The results are consistent with what we show in Panel C of Table 10 except for two differences. First, for KLD sample, the coefficient of $Score^{ESG} \times Log(1+Spikes^{ESG})$ is negative and statistically significant at 5% for "Selloff". Second, for ESG Combined Score, the coefficient of $Score^{ESG} \times Log(1+Spikes^{ESG})$ is negative and statistically significant at 5% for "Decreases".

5. Joint dynamics of investor and firm ESG reading intensity.

In this section, we study the relationship between investors' and firm's ESG-related reading intensity, what we call their joint dynamics. The goal of this analysis is to assess whether or not and the extent to which institutional investors have an influence on the ESG issues of firms they hold. We also discuss other analyses to help bolster our interpretation of ESG attention.

5.1 Capturing the joint dynamics of investor and firm ESG reading intensity.

We first calculate the relative rank for investors to each firm. Following the work by Kempf, Manconi, and Spalt (2017), we rank investors based on their relative importance to both firms and other investors. Specifically, at any quarter-end we first rank investors based on their dollar holdings for each firm (*Investor Rank*), and separately based on the relative fraction of a stock within each investor's portfolio based on dollar holdings (*Firm Rank*). We then calculate the most important investors based on an equal-weighted rank of *Investor Rank* and *Firm Rank.* We split the result into *Top5* and those ranked outside the top five we call *Rest*. At least 10 investors for each firm in any quarter is required.

We present two analyses in Table 11. Panel A presents a topic-firm-quarter analysis. The dependent variable *1{SpikeFirm}[times 100]* indicates whether the firm itself has a Spike score that is larger than or equal to 80 for any topic-quarter. Similar dummies are constructed for top-five investors and all-but-the-top-five investors, respectively. These regressions include nearly 14 million firm-quarter-topic observations with 4,015 unique firms, 18 quarters (from Q3 2015 to Q4 2019), and across 323 different ESG topics. We add firm and quarter fixed effects for all specifications. Our results obtain under a more saturated fixed effect specifications, including combinations of investor × quarter, firm × quarter and investor × firm fixed effects, but we present simpler specifications to facilitate parsimonious interpretations of economic magnitudes.

In Model (1), we show that an investor's ESG-related reading is positively associated with a firm's ESG-related reading intensity and although the coefficients for both $1\{Spike^{Top5\ Inv}\}$ and $1\{Spike^{Rest\ Inv}\}$ are statistically significant at the 1% level, the economic magnitudes differ. The coefficient of top 5 investors is 5.864, which is more than twice the size of that by investors beyond the top 5 in rank. The economic magnitudes imply that when top 5 investors increase reading dramatically on ESG topics, there is a 5.864% higher likelihood that the firm also spikes on ESG topics. Other investors are only associated with a 2.094% higher likelihood. In Model (2), we further add topic fixed effect and the results are consistent with Model (1). In Models (3) to (6), we repeat our regression tests for each of the four ESG categories: Environment, Labor, Social and Governance. Across all columns the coefficients are positive and statistically significant at the 1% level. The economic magnitudes suggest that the coefficients are highest in Governance category, followed by Environment, Labor, and Social. In Appendix C.9, we show a more granular reading of the top five investors and other specifications that are the same as those in Panel A of Table 11. What is notable is that the coefficient of investors decreases monotonically with the investor rank. This analysis lends more confidence to our results in Panel A of Table 11.

Next, in Panel B of Table 11, we test for a causal interpretation by adopting the strategy of Kempf, Manconi, and Spalt (KMS) (2017). We disaggregate our sample further to the investor-stock-topic-quarter. For each investor-stock-quarter, we calculate distraction: for a focal stock *f*, how distracted is the investor by other stocks in other industries? Per KMS (2017), we exclude investor-stock-quarter observations when the stock is in the industry with the highest or lowest return. Given the sheer enormity of this panel, in Models (1) and (2), we keep the top 5 investors in each firm-quarter in the analysis and in Models (3) and (4), we keep all investors and conduct random sampling of the firm-quarter to make the sample size manageable.[15] Consistent with what we find in Panel A of Table 11, the coefficient of investor ESG reading $1\{Spike^{Inv}\}$ is positive and statistically significant at the 1% level. The interaction term $1\{Spike^{Inv}\} \times Distraction$ is negative and statistically significant at the 1% level, which suggests that when investors are distracted by abnormal events in unrelated industries, they are less likely to influence firms on ESG issues.

One might wonder whether our results imply that ESG-related news about a firm drives both firm and investor attention. Even in this scenario, it is unclear why firms would read less about their own firm when investors are distracted unless through the influence of the investor on the firm. That said, we assess the implications of this scenario for our results by removing

[15] The number of firm-quarter observations are 2% of the original one, and we conduct random sampling multiple times and get consistent results.

any firm-quarter where there is any ESG-related news identified by Ravenpack. This suggests that this particular finding is not being driven by ESG news about the firm. Moreover, our clinical analysis in section 2 suggests that the vast majority of ESG-related reading is not firm-specific news.

Overall, the results in Table 11 suggest that investors exert much influence on the firm's ESG reading, and the degree of influence increases when investors have a higher rank in the firm. The Governance category has the highest relative importance, followed by Environment, Labor, and Social.

5.2 Rank of each investor's ESG influence on firms

After examining the average influence of investors on firms' ESG reading, we next examine heterogeneity in investors in terms of their influence. This thought experiment allows us to ascertain what percentage of investors are pro-social and have influence on firms. We operationalize this ranking by forming a measure of the co-movement between a firm's reading and a given investor's reading. For each investor, we run the following regression. We refer to β as the elasticity of the firm's reading with respect to the investor's reading,

$$I\{Spike_{jqs}\} = \alpha_j + \alpha_q + \alpha_s + \beta_1 * 1\left\{\{Spike_{qs}^{Inv}\} * 1\{IO_{jq-1}^{Inv} \geq 0.01\}\right\} + \beta_2 * 1\{\{Spike_{qs}^{Inv}\} + \beta_3 1\{IO_{jq-1}^{Inv} \geq 0.01\} + \varepsilon_{jqs,} \quad (2)$$

$Spike_{jqs}$ is the reading intensity of firm j at quarter q on topic s, $Spike_{qs}^{Inv}$ is the reading intensity of the investor analyzed at quarter q on topic s, IO_{jq-1}^{Inv} is an investor's institutional ownership of firm j at quarter $q-1$. We do the analysis conditional on a minimum threshold level of institutional ownership because investors are more likely to engage firms in which they have a stake in. We require the investors to hold at least 10 stocks with at least 1% institutional ownership and to have had at least 50 reading spikes among ESG topics during the sample period. After this filtering, we have 808 qualifying investors. While we interpret β_1 as a measure of "elasticity," or how sensitive firm ESG attention is to investor ESG attention, we also multiply the elasticity measure by the average number of stocks that investor holds at least 1%. The product of these numbers is what we call "influence." While "elasticity" measures how effective an ESG engagement might be, this alternative measure speaks to the idea that larger investors may have a wider ESG footprint across their portfolios of holdings. Neither measure is clearly better, so we show results of both influence and elasticity in Table 12. Note that the coefficient β_3 also has some potential meaning as a measure of selectivity – namely, whether some investors simply invest in firms that pay more attention to ESG issues.

In Panel A of Table 12, we show the top 20 investors ranked either by influence or elasticity. The ranking results suggest that Blackrock has the highest ESG influence, followed by State Street and Northern Trust. In fact, the two (State Street and Blackrock) far outstrip the third in importance. Interestingly, Northern Trust does not rank as a member of the "big three" (Blackrock, State Street and Vanguard) identified by Azar, Duro, Kadach and Ormazabal (2021). Vanguard is interestingly not on the list of the top 20 by influence. When we rank investors by elasticity alone, Investors Group and Capital Group rank at the top with around 14% elasticity. Blackrock falls in ranking (5^{th}, 11.2% elasticity). This finding suggests that some firms may actually be more vocal about ESG issues, though smaller in influence.

Another potential takeaway is that the distribution of the association investors have with firms' ESG attention spikes, a correlate of their actual ESG activity, is positively skewed with a few firms accounting for most of the influence exerted at large. In Panel B of Table 12, we plot the histogram of estimates of investors' ESG influence. The *y*-axis shows the frequency counts across investors for which the values occur within the *x*-axis intervals of estimated influence. Because the influence measure is highly right-skewed, we winsorize the upper end at 10 for the figure. The histogram indicates that more than half of the influence falls between -1 and 0, and there are few investors with influence more than 5. Many firms cluster close to 0, suggesting they are ineffective in influencing firms' reading activities. In Panel C of Table 12, we plot the histogram of estimates of investor's elasticity instead. The *y*-axis shows the frequency counts associated with 0.02-wide intervals of elasticity along the *x*-axis. This histogram is fairly symmetric and highly peaked around zero. This suggests the majority of institutional investor reading intensity is weakly associated with that of the firms they hold when they also think about the same issues. The 95^{th} percentile of the elasticity is around 0.1, indicating that investors have high unconditional ESG influence when they have above 10% elasticity. The above analysis suggests that a majority of investors have almost no influence or a negative influence on the firm.

6. Conclusions.

In this paper, we leverage big data analytics from an unnamed firm's proprietary intent data to produce a new measure of ESG attention predicated on the internet research activity of the employees of firms across the web. The analysis suggests a meaningful statistical association between firms' and investors' ESG reading intensities and their future ESG performance. Rather than reading passively in anticipation of negative news, firms and investors read intensely in order to take actions they anticipate would improve their ESG performance. Firm ESG performance tends to improve with ESG reading intensity across various measures we examine, while investors seem to trade or vote in a more ESG-friendly way as their ESG reading intensity increases. The magnitudes of these relationships are economically significant. By some measures of ESG, a standard deviation increase in ESG attention is related to a 20% of a standard deviation of improvement in ESG performance.

Our analysis makes several novel contributions. First, for a literature surrounding ESG fraught with concerns about imperfections in existing ESG index-based measures (among others, Gillan, Koch, and Starks, 2021), our findings suggest that these measures do in fact correspond to attention to ESG matters among employees within the firm and by investors. However, the statistical association varies across ratings. The version of the widely used KLD score that the literature studies appears to be somewhat weakly associated with firms' ESG attention, but appears to be the metric used the most by investors, as revealed by the strength of association between KLD ratings, ESG reading intensity among employees of investors, and their future investment actions. This discordance may represent a potential concern with industry best practices, as we show that alternative constructions of the KLD metric are more closely linked to ESG attention. Even more importantly, our new finding on the intensity of co-movement among a firm's attention to ESG with an investor's attention to ESG provides new evidence of generally difficult-to-observe interactions between firms and investors on ESG issues. Our tests suggest that the top-five owners are at least twice as likely to matter as others, and this influence is strongest for governance-related, less so environment-related, issues. We believe our reading intensity measure from proprietary intent data can provide a new valuable tool to investors and firms in the measurement of a variety of issues, including ESG and beyond.

References

Albuquerque, R., Koskinen, Y., Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, *65*(10), 4451-4469.

Akepanidtaworn, K., Di Mascio, R., Imas, A., & Schmidt, L. (2019). Selling fast and buying slow: Heuristics and trading performance of institutional investors. *Available at SSRN 3301277*.

Akey, P., Appel, I. (2019). Environmental externalities of activism. *University of Toronto Working Paper. Available at SSRN 3508808*.

Azar, J., Duro, M., Kadach, I., Ormazabal, G. (2021). The big three and corporate carbon emissions around the world. *Journal of Financial Economics*, forthcoming.

Balasubraman, S., Peterson, R. A., Jarvenpaa, S. L. (2002). Exploring the implications of m-commerce for markets and marketing. *Journal of the Academy of Marketing Science*, *30*(4), 348-361.

Barber, B. M., Morse, A., Yasuda, A. (2021). Impact investing. *Journal of Financial Economics*, *139*(1), 162-185.

Barwise, P., Strong, C. (2002). Permission-based mobile advertising. *Journal of Interactive Marketing*, *16*(1), 14-24.

Becht, M., Franks, J., Wagner, H. (2019). Corporate governance through voice and exit. *London Business School Working Paper. Available at SSRN 3456626.*

Ben-Rephael, A., Da, Z., & Israelsen, R. D. (2017). It depends on where you search: Institutional investor attention and underreaction to news. *The Review of Financial Studies*, *30*(9), 3009-3047.

Berg, F., Koelbel, J. F., Rigobon, R. (2020). Aggregate confusion: the divergence of ESG ratings. *MIT Sloan School Working Paper. Available at SSRN 3438533*.

Brandon, R. G., Krüger, P., Mitali, S. F. (2020). The sustainability footprint of institutional investors: ESG driven price pressure and performance. *Swiss Finance Institute Working Paper. Available at SSRN 2918926*.

Cao, J., Titman, S., Zhan, X., Zhang, W. E. (2020). ESG Preference, Institutional Trading, and Stock Return Patterns. *University of Texas at Austin Working Paper. Available at SSRN 3353623*.

Chatterji, A. K., Durand, R., Levine, D. I., Touboul, S. (2016). Do ratings of firms converge? Implications for managers, investors and strategy researchers. *Strategic Management Journal*, *37*(8), 1597-1614.

Chatterji, A. K., Levine, D. I., Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics and Management Strategy*, *18*(1), 125-169.

Chen, T., Dong, H., Lin, C. (2020). Institutional shareholders and corporate social responsibility. *Journal of Financial Economics*, *135*(2), 483-504.

Choi, D., Gao, Z., Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, *33*(3), 1112-1145.

Cui, B., Docherty, P. (2020). Stock price overreaction to ESG controversies. *Monash University Working Paper. Available at SSRN 3559915*.

Dessein, W., Galeotti, A., Santos, T. (2016). Rational inattention and organizational focus. *American Economic Review*, 106(6), 1522-36.

Dessein, W., Prat, A. (2016). Attention in organizations. Bramoullé, Y, Galeotti, A., Rogers, B. W., eds. *The Oxford Handbook of Economics of Networks*, 675-697.

Dimson, E., Karakaş, O., Li, X. (2015). Active ownership. *The Review of Financial Studies*, *28*(12), 3225-3268.

Dyck, A., Lins, K. V., Roth, L., Towner, M., Wagner, H. F. (2018). Entrenched Insiders and Corporate Sustainability (ESG): How Much Does the "G" Matter for "E" and "S" Performance Around the World? *University of Toronto Working Paper*.

Dyck, A., Lins, K. V., Roth, L., Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, *131*(3), 693-714.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, *33*(3), 1184-1216.

Fernando, C. S., Sharfman, M. P., Uysal, V. B. (2017). Corporate environmental policy and shareholder value: Following the smart money. *Journal of Financial and Quantitative Analysis*, *52*(5), 2023-2051.

Gillan, S. L., Koch, A., Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, forthcoming.

Gibson, R., Krüger, P., Riand, N., Schmidt, P. S. (2019). ESG rating disagreement and stock returns. *Swiss Institute of Finance Working Paper. Available at SSRN 3433728.*

Hafez, P., Gomez, F., 2019. Socially responsible investing: Combining ESG ratings with news sentiment generates alpha. *RavenPack Working Paper. Available at SSRN 3782432.*

Hartzmark, S. M., Sussman, A. B. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance*, *74*(6), 2789-2837.

Heath, D., Macciocchi, D., Michaely, R., Ringgenberg, M. C. (2021). Does Socially Responsible Investing Change Firm Behavior? *Hong Kong University Business School Working Paper*.

Hwang, C. Y., Titman, S., Wang, Y. (2015). Investor tastes, corporate behavior and stock returns: An analysis of corporate social responsibility. *University of Texas at Austin Working Paper.*

Iliev, P., Kalodimos, J., Lowry, M. (2020). Investors' attention to corporate governance. *Review of Financial Studies, forthcoming*.

Jagannathan, R., Ravikumar, A., Sammon, M. (2017). Environmental, social, and governance criteria: why investors are paying attention. *National Bureau of Economic Research working paper #w24063.*

Joseph, J., Wilson, A. J. (2018). The growth of the firm: An attention-based view. *Strategic Management Journal*, *39*(6), 1779-1800.

Kacperczyk, M., Van Nieuwerburgh, S., Veldkamp, L. (2016). A rational theory of mutual funds' attention allocation. *Econometrica*, 84(2), 571-626.

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, *38*(2), 95-115.

Kempf, E., Manconi, A., Spalt, O. (2017). Distracted shareholders and corporate actions. *The Review of Financial Studies*, *30*(5), 1660-1695.

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, *115*(2), 304-329.

Kwan, A. (2020). Measuring Remote Knowledge Work Using Big Data. *Hong Kong University Business School Working Paper*.

Kwan, A., Zhu, C. (2020). Does Internet Research by Sophisticated Investors Cause Adverse Selection? *Hong Kong University Business School Working Paper.*

Li, Q., Maggitti, P. G., Smith, K. G., Tesluk, P. E., Katila, R. (2013). Top management attention to innovation: The role of search selection and intensity in new product introductions. *Academy of Management Journal*, *56*(3), 893-916.

Li, X., Raghunandan, A. (2020). Institutional ownership and labor-related misconduct: Evidence from US federal violations. *London School of Economics Working Paper*. *Available at SSRN 3460126*.

Matos, P. (2020). *ESG and Responsible Institutional Investing Around the World: A Critical Review*. CFA Institute Research Foundation, Charlottesville, VA.

Nordhaus, W. (2019). Climate change: the ultimate challenge for economics. *American Economic Review*, 109(6), 1991-2014.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic Management Journal*, *18*(S1), 187-206.

Olafsson, A., & Pagel, M. (2017). The ostrich in us: Selective attention to financial accounts, income, spending, and liquidity. *National Bureau of Economic Research working paper #w23945*.

Pagel, M. (2018). A News-Utility Theory for Inattention and Delegation in Portfolio Choice. *Econometrica*, *86*(2), 491-522.

Raghunandan, A., Rajgopal, S. (2020). Do the Socially Responsible Walk the Talk? *Columbia University Working Paper. Available at SSRN 3609056.*

Riedl, A., Smeets, P. (2017). Why do investors hold socially responsible mutual funds?. *The Journal of Finance*, *72*(6), 2505-2550.

Sicherman, N., Loewenstein, G., Seppi, D. J., & Utkus, S. P. (2016). Financial attention. *The Review of Financial Studies*, *29*(4), 863-897.

Starks, L. T., Venkat, P., Zhu, Q. (2017). Corporate ESG profiles and investor horizons. *University of Texas Working Paper. Available at SSRN 3049943.*

Tang, D. Y., Yan, J., Yao, C. Y. (2020). The Determinants of ESG Ratings: Rater Ownership Matters. *Hong Kong University Business School Working Paper*.

Tong, S., Luo, X., Xu, B. (2020). Personalized mobile marketing strategies. *Journal of the Academy of Marketing Science*, *48*(1), 64-78.

Yang, R. (2019). What Do We Learn From Ratings About Corporate Social Responsibility (CSR)?. *Columbia Business School Research Paper*. *Available at SSRN 3165783*.

Figure 1: Firm Attentional Responses to BlackRock's Larry Fink letter on January 14, 2020.

In this figure, we plot the reading intensity of firms surrounding the day of January 14, 2020 on which Larry Fink of Blackrock issued the letter to CEOs regarding sustainability, emphasizing climate issues (although not exclusively). The y-axis is the percentage of total reading that are related to any ESG category for firms in the top 10% of Blackrock holdings versus the firms in the bottom 10% (including those firms where Blackrock had zero holdings), normalized to be 0 on January 6. The ESG category chosen are Environmental, Governance, Labor, and Data and Sensitive Information Protection.

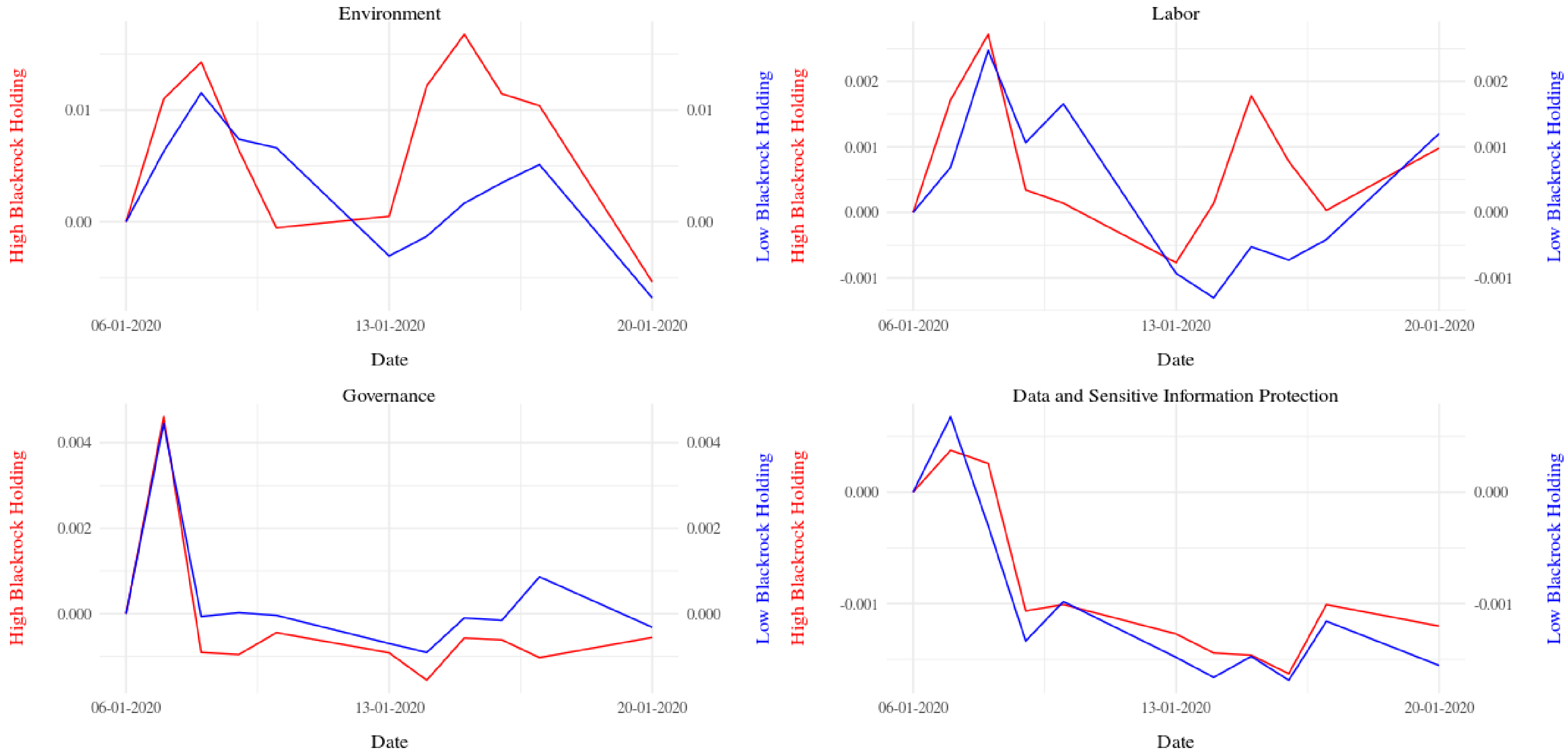


Figure 2: Average climate beta on firm's environment-related reading

In this figure, we plot the average climate beta on firm's environment-related reading. The estimation of the climate beta is detailed in Appendix B. The y-axis is the average climate beta of each bin/portfolio. The x-axis shows the five bins into which a firm's environment-related reading intensity is sorted: Bin 1 indicates the lowest reading while Bin 5 indicates the highest reading. The heading of each plot indicates whether we use Spike data or daily data of The Company to calculate firm's environment-related reading.

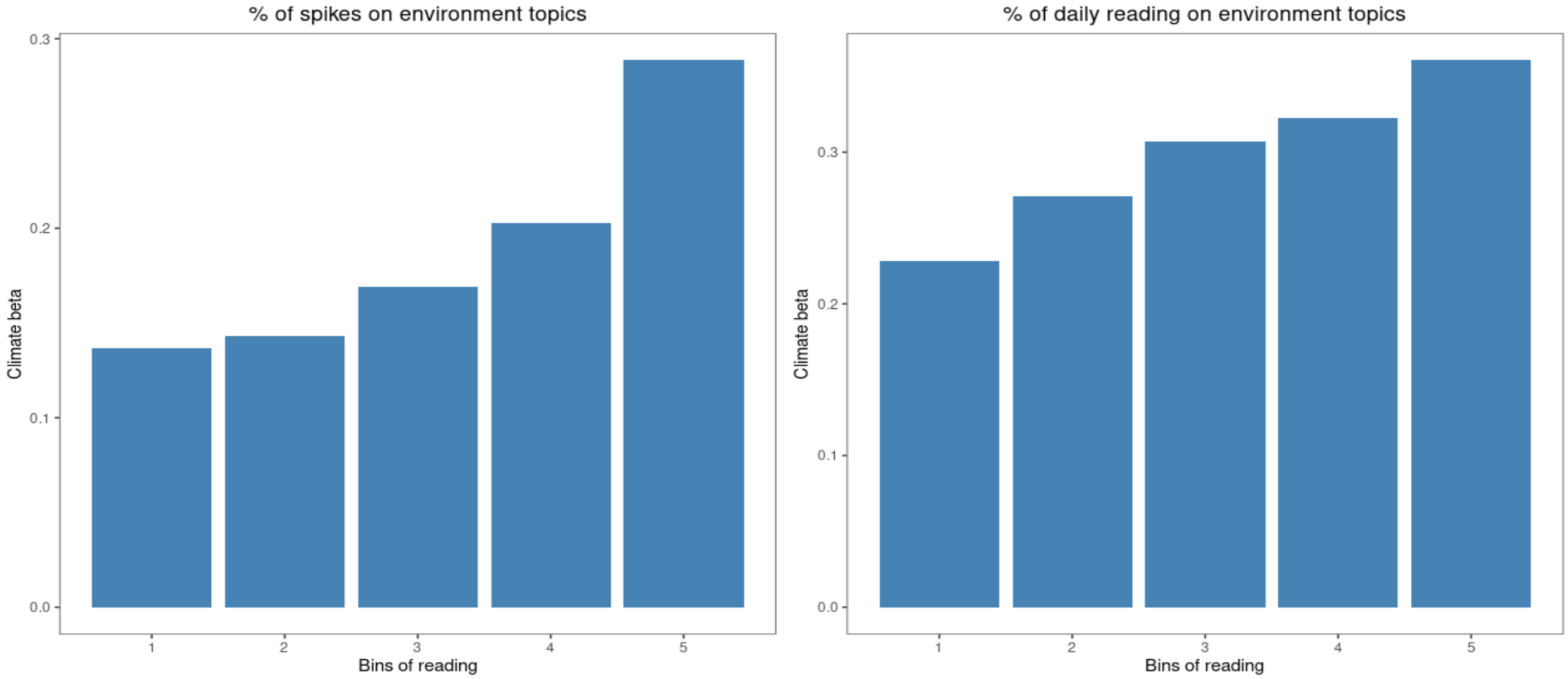


Table 1: Intent Dataset Coverage.

This table presents summary statistics on our intent dataset coverage. Over the course of the five years in our period of analysis from 2015 through 2019, the number of topics identified by The Company has nearly tripled from 2,462 to 6,765. Panel A exhibits the pace of growth arises in the number of web-domains – or business "addresses" - that The Company tracks from 1.67 million in 2015 to 6.9 million by the end of our sample. We also report the number of domain-mapped business-related interactions per day which reaches a peak of 686 million in 2017 across the 4.3 million domains and for 3,589 different topics as of that year. Panel B lists the topic taxonomy as of 2019 by themes as defined by The Company.

Panel A: Number of The Company Topics by Year

Year	# Topics	# Domains	Domain-mapped Business-Related Interactions Per Day
2015	2462	1677494	(not available)
2016	2962	1819851	506892107
2017	3589	4303994	686276353
2018	5433	5473714	623016344
2019	6765	6907293	272946594

Panel B: Number of topics in each topic category

Theme name	# topics	Theme name	# topics
Events and Conferences	60	Human Resources	321
Government	87	Healthcare	327
Biotechnology	106	Energy/Construction/Manufacturing	339
Consumer Technology	149	Marketing	523
Business	301	Finance	561
Legal	305	Company	1025
		Non-consumer technology	1849

Table 2: ESG topics

Panel A: Number of topics within each ESG category

This panel shows number of topics within each ESG category we classify. There are 323 ESG topics classified to 9 categories.

Category Name	**# topics**
Compliance	18
Corporate Governance	29
CSR	28
Customer Relation	23
Cybersecurity	39
Data and Sensitive Information Protection	40
Environment	46
Equality and Diversity	21
Labor Relation	63
Total	323

Panel B: Example topics within each dimension

This panel shows 10 example topics within each 4 dimension, which we select for demonstration purpose: Environment, Labor, Compliance, and Data and Sensitive Information Protection.

Environment	**Labor**	**Compliance**	**Data and Sensitive Information Protection**
Air Pollution	Diversity Recruiting	Accounting Compliance	Data Privacy and Protection
Alternative-Fuel Vehicles	Employee Safety	Business Law	Data Security
Carbon Footprint	Employee Satisfaction	Compliance	Enterprise Application Security
Carbon Management	Equal Employment Opportunity (EEO)	Compliance Management	Information Security
Climate Change	Equal Pay / Comparable Worth	Compliance Training	Internet Security
Emissions	Gender Equality	Global Employment Law	Intrusion Prevention
Global Warming	Labor Relations	Know Your Customer (KYC)	Security Monitoring
Greenhouse Gas	Labor Union	Minimum Wage	Sensitive Data
Renewable Energy	Wellness Benefits	Regulatory Compliance	Strong Encryption
Water Pollution	Workers' Compensation	Tax Compliance	Threat Prevention

Table 2: ESG topics (continued)

Panel C: Number of The Company ESG Topics by Year

Year	# Topics	# Domains	Domain-mapped Business-Related Interactions Per Day	% of Interactions across all topics
2015	172	1520884	(not available)	(not available)
2016	188	1668113	34182462	6.74%
2017	226	4111852	64512222	9.40%
2018	323	5130072	47857872	7.68%
2019	323	6161740	20245142	7.42%

Panel D: Top Industries for Select Categories

This table shows 10 industries which have highest percentage reading within each 4 dimension: Environment, Labor, Compliance, and Data and Sensitive Information Protection. Labor includes both topics of Labor Relation and Equality and Diversity. The percentage of reading of each dimension is defined as total record of topics in that dimension divided by total record of all topics. We define industry as first two digits of NAICS code.

Environment	Labor	Compliance	Data and Sensitive Information Protection
Utilities	Educational Services	Finance and Insurance	Finance and Insurance
Mining, Quarrying, and Oil and Gas Extraction	Health Care and Social Assistance	Professional, Scientific, and Technical Services	Professional, Scientific, and Technical Services
Educational Services	Accommodation and Food Services	Administrative, Support Waste Management and Remediation Services	Health Care and Social Assistance
Construction	Management of Companies, Enterprises	Mining, Quarrying, and Oil and Gas Extraction	Accommodation and Food Services
Management of Companies, Enterprises	Administrative, Support Waste Management and Remediation Services	Accommodation and Food Services	Information
Professional, Scientific, and Technical Services	Arts, Entertainment, and Recreation	Construction	Arts, Entertainment, and Recreation
Agriculture, Forestry, Fishing and Hunting	Wholesale Trade	Educational Services	Real Estate and Rental and Leasing
Accommodation and Food Services	Agriculture, Forestry, Fishing and Hunting	Management of Companies, Enterprises	Wholesale Trade
Administrative, Support Waste Management and Remediation Services	Professional, Scientific, and Technical Services	Health Care and Social Assistance	Mining, Quarrying, and Oil and Gas Extraction
Manufacturing	Manufacturing	Real Estate and Rental and Leasing	Management of Companies, Enterprises

Table 3: Distributions of Spike Scores

Panel A All topics

This panel shows distributions of The Company's Spike Score for CRSP firms by year across all The Company topics. We report 25th, 50th, 75th, 90th, 95th and 99th percentiles of the distributions.

Year	25th	50th	75th	90th	95th	99th
2015	14	25	30	58	76	86
2016	16	25	45	60	73	85
2017	38	48	57	67	73	83
2018	40	48	57	66	71	81
2019	41	51	60	70	75	85

Panel B ESG topics only

This panel shows distributions of The Company's Spike Score for CRSP firms by year across ESG topics. We report 25th, 50th, 75th, 90th, 95th and 99th percentiles of the distributions.

Year	25th	50th	75th	90th	95th	99th
2015	14	25	35	55	71	84
2016	16	25	47	59	71	84
2017	38	48	57	66	72	83
2018	40	48	57	65	70	80
2019	41	50	60	69	74	84

Table 4: Firm's ESG reading and real ESG outcomes

Panel A: EPA pollution prevention

This panel shows how firm's ESG reading is associated with subsequent changes in the firm's (log) number of facilities reporting newly implemented source reduction activities ("Green facilities"). We get the outcome variable from EPA TRI program and show summary statistics of it in Appendix F.1. Models (1) to (2) present results for both public and private firms, while Models (3) to (6) present results for public firms only. In Models (4) to (6), we standardize all independent variables to compare their economic significance, and the heading in each of the three columns represent the ESG rating score that applies. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Log(1+ # Green facilities)					
Sample:	**All**	**All**	**Public**	**Refinitiv^{Combined}**	**Current RRI**	**KLD1**
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes ^{ESG})	0.020***	0.029***	0.065**	0.042**	0.040**	0.047**
	(0.005)	(0.008)	(0.026)	(0.020)	(0.017)	(0.019)
Log(1+Spikes$^{Not\ ESG}$)		-0.010	-0.038			
		(0.008)	(0.026)			
Score^{ESG}				0.003	-0.007	-0.001
				(0.014)	(0.020)	(0.021)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,590	14,590	3,597	1,756	2,498	2,375
Adjusted R^2	0.585	0.585	0.448	0.463	0.426	0.443

Panel B: Employee benefits

This panel shows how firm's ESG reading is associated with subsequent changes in employee benefits provided by the firm. The outcome variable is company contributions to employee's pension plans scaled by firm size, and we take logarithm of it. We get the outcome variable from Axiomatic Data and show summary statistics of it in Appendix F.1. Models (1) to (5) present results for public firms only. In Models (3) to (5), we standardize all independent variables to compare their economic significance, and the heading in each of the three columns represent the ESG rating score that applies. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Contributions divided by firm size				
Sample:	**All**	**All**	**Refinitiv^{Combined}**	**Current RRI**	**KLD1**
	(1)	(2)	(3)	(4)	(5)
Log(1+Spikes ^{ESG})	0.026***	0.071***	0.036***	0.015**	0.035***
	(0.006)	(0.015)	(0.008)	(0.008)	(0.009)
Log(1+Spikes$^{Not\ ESG}$)		-0.048***			
		(0.015)			
Score^{ESG}			0.004	0.007	0.022
			(0.014)	(0.009)	(0.014)
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	10,430	10,430	3,889	4,116	4,796
Adjusted R^2	0.937	0.937	0.946	0.933	0.944

Table 4: Firm's ESG reading and real ESG outcomes (continued)

Panel C: OSHA penalties

This panel shows how firm's ESG reading are associated with subsequent changes in the number of OSHA penalties. We divide it by number of establishments as the outcome variable. We get the outcome variable from OSHA and show summary statistics of it in Appendix F.1. Models (1) to (2) present results for both public and private firms, while Models (3) to (6) present results for public firms only. In Models (4) to (6), we standardize all independent variables to compare their economic significance, and the heading in each of the three columns represent the ESG rating score that applies. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	OSHA penalties per establishments					
Sample:	**All**	**All**	**Public**	**Refinitiv^{Combined}**	**Current RRI**	**KLD1**
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes ^{ESG})	-0.016***	-0.038***	-0.030**	-0.023**	-0.029***	-0.022**
	(0.005)	(0.011)	(0.013)	(0.010)	(0.009)	(0.009)
Log(1+Spikes$^{Not\ ESG}$)		0.026**	0.008			
		(0.012)	(0.017)			
Score^{ESG}				0.010	0.010	-0.010
				(0.007)	(0.009)	(0.008)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,403	10,403	4,715	1,815	2,561	2,586
Adjusted R^2	0.249	0.250	0.399	0.531	0.485	0.457

Table 5: Firm's reading and ESG performance (Refinitiv)

Panel A: Summary Statistics (Annual)

Statistic	N	Mean	St. Dev.	Min	25^{th}	75^{th}	Max
$Refinitiv^{Combined}$	6,759	45.710	16.407	18.317	33.730	56.315	86.039
$Refinitiv^{ESG}$	6,759	51.287	18.504	19.240	35.825	66.930	89.422
$Refinitiv^{Contro}$	6,759	48.200	20.048	0.000	52.580	58.930	66.670
$Log(1+Spikes^{ESG})$	6,759	3.937	1.324	0.000	3.497	4.820	5.789
$Log(1+Spikes^{Environ})$	6,759	2.144	1.037	0.000	1.609	2.890	3.951
$Log(1+Spikes^{Labor})$	6,759	2.874	1.122	0.000	2.398	3.664	4.700
$Log(1+Spike^{Social})$	6,759	1.867	0.999	0.000	1.099	2.639	4.025
$Log(1+Spikes^{Govern})$	6,759	3.097	1.324	0.000	2.485	4.060	5.017
$Log(1+Spikes^{Not\ ESG})$	6,759	6.522	1.627	0.000	6.067	7.556	8.769

Panel B: Level of Refinitiv ESG Score

This panel shows how firm's ESG-related reading is associated with ESG rating by Refinitiv. ESG score measures firm's relative ESG performance across more than 450 metrics, and ESG Combined Score is the ESG score with the ESG controversies overlay. For all three Refinitiv scores, higher levels indicate better performance. Models (1) to (4) present results for level of ESG Combined score, Models (5) and (6) show results for ESG score and ESG Controversies Score separately. We define industry by first two digits of SIC code. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	$Refinitiv^{Combined}$				$Refinitiv^{ESG}$	$Refinitiv^{Contro}$
	(1)	(2)	(3)	(4)	(5)	(6)
$Log(1+Spikes^{ESG})$	1.230^{***}	0.483^{***}	1.031^{**}		0.102	1.582^{**}
	(0.212)	(0.166)	(0.422)		(0.203)	(0.686)
$Log(1+Spikes^{Not\ ESG})$			-0.514		0.072	-1.037^{*}
			(0.357)		(0.174)	(0.595)
$Log(1+Spikes^{Environ})$				-0.199		
				(0.296)		
$Log(1+Spikes^{Labor})$				-0.249		
				(0.305)		
$Log(1+Spike^{Social})$				0.754^{***}		
				(0.285)		
$Log(1+Spikes^{Govern})$				0.381		
				(0.292)		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes					
Firm FE		Yes	Yes	Yes	Yes	Yes
Observations	6,759	6,759	6,759	6,759	6,759	6,759
Adjusted R^2	0.072	0.676	0.677	0.677	0.942	0.395

Table 6: Firm's reading and ESG performance (RepRisk)

Panel A: Summary Statistics (Monthly)

Statistic	N	Mean	St. Dev.	Min	25th	75th	Max
Current RRI	59,413	12.827	12.097	0.000	0.000	21.000	55.000
Log(1+Spikes^{ESG})	59,413	1.677	1.037	0.000	1.099	2.398	4.316
Log(1+Spikes^{Environ})	59,413	0.450	0.619	0.000	0.000	0.693	2.303
Log(1+Spikes^{Labor})	59,413	0.876	0.821	0.000	0.000	1.386	3.497
Log(1+Spike^{Social})	59,413	0.371	0.544	0.000	0.000	0.693	2.079
Log(1+Spikes^{Govern})	59,413	1.080	0.954	0.000	0.000	1.792	3.466
Log(1+Spikes$^{Not\ ESG}$)	59,413	4.075	1.378	0.000	3.466	4.977	6.802

Panel B: RepRisk Index

This panel shows how firm's ESG-related reading is associated with ESG-related risk by RepRisk. The outcome variable is current RRI, which measures firms' current exposure to ESG risks. In Models (1) to (4) the unit of observation is firm-month while in Models (5) to (6) the unit of observation is firm-year. We define industry by first two digits of SIC code. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Current RRI					
	Monthly				Annual	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes^{ESG})	2.060***	-0.325***	-0.203***		-0.450***	-0.250
	(0.165)	(0.071)	(0.065)		(0.103)	(0.317)
Log(1+Spikes$^{Not\ ESG}$)			-0.133**			-0.178
			(0.064)			(0.271)
Log(1+Spikes^{Environ})				-0.177***		
				(0.061)		
Log(1+Spikes^{Labor})				-0.171***		
				(0.063)		
Log(1+Spike^{Social})				-0.181**		
				(0.070)		
Log(1+Spikes^{Govern})				-0.109*		
				(0.062)		
Month FE	Yes	Yes	Yes	Yes		
Year FE					Yes	Yes
Industry FE	Yes					
Firm FE		Yes	Yes	Yes	Yes	Yes
Observations	59,413	59,413	59,413	59,413	6,610	6,610
Adjusted R^2	0.100	0.706	0.706	0.706	0.653	0.653

Table 7: Firm's reading and ESG performance (KLD)

Panel A: Summary Statistics (Annual)

Statistic	N	Mean	St. Dev.	Min	25th	75th	Max
KLD1	7,396	0.024	0.036	-0.107	0.000	0.037	0.217
KLD2	7,396	0.448	0.646	-1.000	0.000	1.000	1.000
KLD3	5,894	0.562	0.678	-1.000	0.273	1.000	1.000
$Log(1+Spikes^{ESG})$	7,396	4.198	1.116	0.000	3.738	4.934	6.749
$Log(1+Spikes^{Environ})$	7,396	2.335	0.939	0.000	1.946	2.996	4.844
$Log(1+Spikes^{Labor})$	7,396	3.098	0.984	0.000	2.565	3.784	5.889
$Log(1+Spike^{Social})$	7,396	2.033	0.937	0.000	1.386	2.773	4.304
$Log(1+Spikes^{Govern})$	7,396	3.334	1.203	0.000	2.773	4.163	5.905
$Log(1+Spikes^{Not\ ESG})$	7,396	6.842	1.284	0.000	6.353	7.640	9.541

Panel B: Adjusted KLD Score

This panel shows how firm's ESG-related reading is associated with ESG rating by KLD. Models (1) to (3) present results for the first version of KLD score, which is (Str – Con)/(n_Str + n_Con). In Models (4) to (6), we focus on the intensive margin, which is defined as (Str – Con)/(Str + Con). If there are no strengths or concerns for any firm-year, the KLD score would be 0 in Model (4) but missing in Models (5) and (6). Str and Con are number of strengths and concerns the firm have in each year respectively. n_Str and n_Con are number of maximum strengths and concerns the firm could have respectively. We multiply outcome variables by 100 in all models. We define industry by first two digits of SIC code. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	KLD1			KLD2	KLD3	
	(1)	(2)	(3)	(4)	(5)	(6)
$Log(1+Spikes^{ESG})$	0.749***	-0.032	-0.132	3.577*	5.367**	
	(0.055)	(0.039)	(0.082)	(1.981)	(2.089)	
$Log(1+Spikes^{Not\ ESG})$			0.099	2.006	0.284	
			(0.069)	(1.881)	(1.993)	
$Log(1+Spikes^{Environ})$						1.280
						(1.517)
$Log(1+Spikes^{Labor})$						-1.197
						(1.915)
$Log(1+Spike^{Social})$						1.128
						(1.372)
$Log(1+Spikes^{Govern})$						5.041***
						(1.547)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes					
Firm FE		Yes	Yes	Yes	Yes	Yes
Observations	7,396	7,396	7,396	7,396	5,894	5,894
Adjusted R^2	0.146	0.790	0.790	0.612	0.689	0.690

Table 8: Investor's ESG reading and ESG-friendly shareholder voting

This table shows how investor's ESG reading one month before the shareholder meeting is associated with fund voting. In Models (1) to (5), the outcome variable "Fund vote for ESG**"** is a dummy variable indicating whether the fund votes for ESG-friendly proposals or votes against anti-social proposals. Model (4) presents result for E and S proposals only and Model (5) presents result for governance proposals only. In Models (6) to (8), the outcome variable "Fund vote with ISS" is a dummy variable indicating whether the fund votes the same as ISS's recommendation on ESG-linked proposals. We multiply outcome variables by 100 in all models. Standard errors are clustered on shareholder's meeting. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Fund vote for ESG					**Fund vote with ISS**		
	All	**All**	**All**	**E and S**	**Gov**	**All**	**All**	**All**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(1+Spikes^{ESG})$	0.721***	1.183***	1.177***			0.939***	1.205***	1.247***
	(0.251)	(0.373)	(0.373)			(0.264)	(0.401)	(0.401)
$Log(1+Spikes^{Not\ ESG})$		-0.541*	-0.534*				-0.313	-0.388
		(0.291)	(0.291)				(0.308)	(0.312)
$Log(1+Spikes^{Environ})$				0.672*	0.495			
				(0.370)	(0.440)			
$Log(1+Spikes^{Labor})$				-0.398	0.682			
				(0.484)	(0.539)			
$Log(1+Spike^{Social})$				-0.491	0.636			
				(0.343)	(0.500)			
$Log(1+Spikes^{Govern})$				-0.922***	1.160***			
				(0.356)	(0.436)			
ISS Vote	No	No	Yes	Yes	Yes	No	No	Yes
Management Vote	No	No	Yes	Yes	Yes	No	No	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	457,454	457,454	457,454	224,333	233,121	426,144	426,144	426,144
Adjusted R^2	0.486	0.486	0.567	0.504	0.548	0.437	0.437	0.523

Table 9: Investor's ESG reading and portfolio-level ESG performance

This panel shows how investor's ESG reading is associated with the ESG related performance of an investor's portfolio. In Models (1) to (6), the outcome variable is the equal-weighted ESG rating of stocks. Peak RRI is the maximum of RepRisk Index in the last two years, which represents past ESG reputation. In Model (7), the outcome variable is percentage of stocks that have Peak RRI which are larger than 0. In Model (8), the outcome variable is percentage of stocks that have Peak RRI which are larger than or equal to 50. Other variables are defined in Tables 5, 6, 7. We standardize the outcome variables and multiply them by 100 in all models. Standard errors are clustered on investor. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	KLD1	**Refinitiv^{Combined}**	**Refinitiv^{ESG}**	**Refinitiv^{Contro}**	**Current RRI**	**Peak RRI**	**% Peak RRI 0**	**% Peak RRI 50**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(1+Spikes^{ESG})	1.330***	1.886***	1.206***	-0.083	-0.358	-0.843*	-1.958***	-0.812
	(0.416)	(0.559)	(0.400)	(0.473)	(0.510)	(0.510)	(0.722)	(0.512)
Log(1+Spikes$^{Not\ ESG}$)	0.178	-0.621	-0.511	0.216	0.305	0.219	0.381	0.021
	(0.366)	(0.514)	(0.345)	(0.432)	(0.387)	(0.389)	(0.641)	(0.393)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	40,000	39,787	39,787	39,787	32,427	32,427	32,427	32,427
Adjusted R^2	0.871	0.735	0.886	0.816	0.873	0.872	0.678	0.881

Table 10: Investor's ESG reading and trading decisions

Panel A: Summary Statistics (Quarterly)

Statistic	N	Mean	St. Dev.	Min	25th	75th	Max
Invest	7,561,289	-1.392	4.022	-15.792	-0.181	0.033	2.424
$Log(1+Spikes^{ESG})$	7,561,289	2.646	1.201	0.000	1.946	3.555	4.431
$Log(1+Spikes^{Not\ ESG})$	7,561,289	5.204	1.498	0.000	4.419	6.317	7.327
KLD1	7,053,543	0.000	1.000	-2.134	-0.714	0.451	3.758
$Refinitiv^{Combined}$	6,045,678	0.000	1.000	-1.868	-0.695	0.662	2.644
Current RRI	4,818,398	0.000	1.000	-1.292	-0.885	0.376	2.970
RRI Trend (ΔRRI)	4,818,398	0.000	1.000	-2.893	-0.536	0.206	4.675

Panel B: Overall investment

This panel shows how investor's ESG-related reading and stock's ESG performance is associated with overall investment activity. The outcome variable is *Invest*, which is log change of dollar holdings adjusted by quarterly stock return (defined in Section 4.2**)**. We present results for stocks held by the investors last quarter. $Score^{ESG}$ is the standardized measure of different ESG ratings. Model (1) presents results for adjusted KLD score and Model (2) presents results for ESG Combined Score. Model (3) shows results for Current RRI and Model (4) shows results for change of Current RRI (Δ*RRI*). In all models, we multiply the outcome variable, stock return and volatility by 100. Standard errors are clustered on investor. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Invest			
	KLD1	$\mathbf{Refinitiv^{Combined}}$	**Current RRI**	**ΔRRI**
	(1)	(2)	(3)	(4)
$Score^{ESG} \times Log(1+Spikes^{ESG})$	1.200**	1.048**	-0.171	-0.396
	(0.515)	(0.522)	(0.767)	(0.335)
$Score^{ESG} \times Log(1+Spikes^{Not\ ESG})$	-0.487	-0.562	0.408	0.546**
	(0.478)	(0.445)	(0.680)	(0.274)
$Score^{ESG}$	1.690	1.060	-0.995	-2.385***
	(1.540)	(1.294)	(2.348)	(0.716)
Return	0.244***	0.555***	0.070*	0.069*
	(0.040)	(0.039)	(0.038)	(0.038)
Volatility	-1.303***	-1.305***	-1.100***	-1.091***
	(0.117)	(0.118)	(0.104)	(0.105)
Log(Market Cap)	24.770***	23.832***	25.807***	26.066***
	(0.947)	(0.903)	(1.070)	(0.962)
Momentum	6.677***	9.167***	11.466***	11.305***
	(1.779)	(1.749)	(1.840)	(1.825)
Gross Profitability	-2.058	-6.510***	1.366	1.661
	(2.250)	(2.120)	(2.158)	(2.219)
Book-to-Market	8.129***	0.138	7.181***	7.425***
	(1.399)	(1.330)	(1.223)	(1.297)
Investor×Quarter FE	Yes	Yes	Yes	Yes
Observations	7,053,543	6,045,678	4,818,398	4,818,398
Adjusted R^2	0.262	0.284	0.259	0.259

Table 10: Investor's ESG reading and trading decisions (continued)

Panel C: Different types of investment

This panel shows different types of investment of Panel B. For demonstration purpose we use ESG^{Zscore}, which is the sum of the standardized KLD score and ESG Combined Score. Models (1) to (4) present results for stocks held by the investors last quarter while Model (5) presents results for stocks not held by the investors last quarter. In all models the outcome variables are dummy variables and we multiply them by 100. "Selloff" is defined as liquidation of all shares in an existing position. "Decreases" is defined as more than 1 percent decrease of dollar holdings. "Hold" is defined as within 1 percent change of dollar holdings. "Increases" is defined as more than 1 percent increase of dollar holdings. "Pickup" is defined as de novo investment in a stock that was not held by the investor last quarter. Standard errors are clustered on investor. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	1{Selloff}	**1{Decreases}**	**1{Hold}**	**1{Increases}**	**1{Pickup}**
	(1)	(2)	(3)	(4)	(5)
$ESG^{Zscore} \times Log(1+Spikes^{ESG})$	-0.054^{*}	-0.087	-0.045	0.132^{**}	0.036^{***}
	(0.031)	(0.069)	(0.058)	(0.055)	(0.007)
$ESG^{Zscore} \times Log(1+Spikes^{Not\ ESG})$	0.044	0.188^{***}	-0.089^{*}	-0.099^{**}	0.008^{*}
	(0.028)	(0.059)	(0.051)	(0.049)	(0.004)
ESG^{Zscore}	-0.203^{**}	-0.730^{***}	0.812^{***}	-0.082	-0.082^{***}
	(0.084)	(0.153)	(0.148)	(0.132)	(0.012)
Return	-0.041^{***}	-0.057^{***}	0.044^{***}	0.013^{**}	0.000
	(0.003)	(0.005)	(0.003)	(0.006)	(0.0003)
Volatility	0.116^{***}	0.048^{***}	-0.228^{***}	0.180^{***}	0.017^{***}
	(0.010)	(0.014)	(0.012)	(0.012)	(0.001)
Log(Market Cap)	-2.205^{***}	1.033^{***}	-3.408^{***}	2.375^{***}	0.648^{***}
	(0.074)	(0.118)	(0.141)	(0.124)	(0.018)
Momentum	-0.697^{***}	-3.764^{***}	2.237^{***}	1.527^{***}	-0.050^{***}
	(0.127)	(0.275)	(0.201)	(0.248)	(0.012)
Gross Profitability	0.184	1.804^{***}	-2.684^{***}	0.880^{***}	0.395^{***}
	(0.169)	(0.233)	(0.218)	(0.263)	(0.021)
Book-to-Market	-0.060	3.809^{***}	-3.051^{***}	-0.758^{***}	0.119^{***}
	(0.105)	(0.197)	(0.193)	(0.207)	(0.012)
Investor×Quarter FE	Yes	Yes	Yes	Yes	Yes
Observations	5,760,778	5,760,778	5,760,778	5,760,778	54,965,093
Adjusted R^2	0.293	0.148	0.221	0.151	0.131

Table 11: Investor's influence on firm's ESG Reading

Panel A: Top 5 vs other investors

This panel shows how investor's ESG-related reading is associated with firm's ESG-related reading at the topic-quarter level. Across all models the outcome variable is *1{Spike^{Firm}}*, which is a dummy variable indicating whether the firm has a Spike score that is larger than or equal to 80 for each topic-quarter. *1{Spike$^{Top5\ Inv}$}* and *1{Spike$^{Rest\ Inv}$}* are similar defined dummies for top 5 investors and rest investors of the firm. The heading from Models (1) to (6) represent the ESG categories used in the analysis. We multiply the outcome variable by 100. Standard errors clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	1{Spike^{Firm}}					
	All	**All**	**Env**	**Labor**	**Social**	**Gov**
	(1)	(2)	(3)	(4)	(5)	(6)
1{Spike$^{Top5\ Inv}$}	5.864***	5.283***	4.879***	4.593***	3.656***	6.446***
	(0.136)	(0.138)	(0.135)	(0.123)	(0.140)	(0.172)
1{Spike$^{Rest\ Inv}$}	2.094***	1.459***	1.138***	1.499***	0.966***	1.589***
	(0.073)	(0.070)	(0.091)	(0.072)	(0.069)	(0.079)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Topic FE		Yes	Yes	Yes	Yes	Yes
Observations	14,251,321	14,251,321	1,932,765	4,889,983	1,882,312	5,546,261
Adjusted R^2	0.037	0.045	0.049	0.038	0.041	0.059

Panel B: Distracted investors

This panel presents results for investors' ESG influence on firms when they are distracted by firms in other industries. The construction of variable *Distraction* follows Kempf, Manconi, and Spalt (2017). The unit of observation is investor-firm-quarter-topic. The heading from Models (1) to (4) represent the sample used in the analysis. Other variables are constructed in a similar way as in Panel A. We multiply the outcome variable by 100. Standard errors clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	1{Spike^{Firm}}			
	Top 5 Investors		**All investors with sampling**	
	(1)	(2)	(3)	(4)
1{Spike^{Inv}}	7.141***	7.728***	9.369***	10.902***
	(0.181)	(0.188)	(0.449)	(0.802)
1{Spike^{Inv}}× Distraction		-4.701***		-13.353***
		(0.586)		(4.927)
Distraction		-0.121		1.090
		(0.154)		(1.499)
Quarter + Firm + Investor + Topic FE	Yes	Yes	Yes	Yes
Observations	39,588,838	39,588,838	44,703,693	44,703,693
Adjusted R^2	0.040	0.040	0.120	0.120

Table 12: ESG influence by investor

Panel A: Investors with highest ESG influence

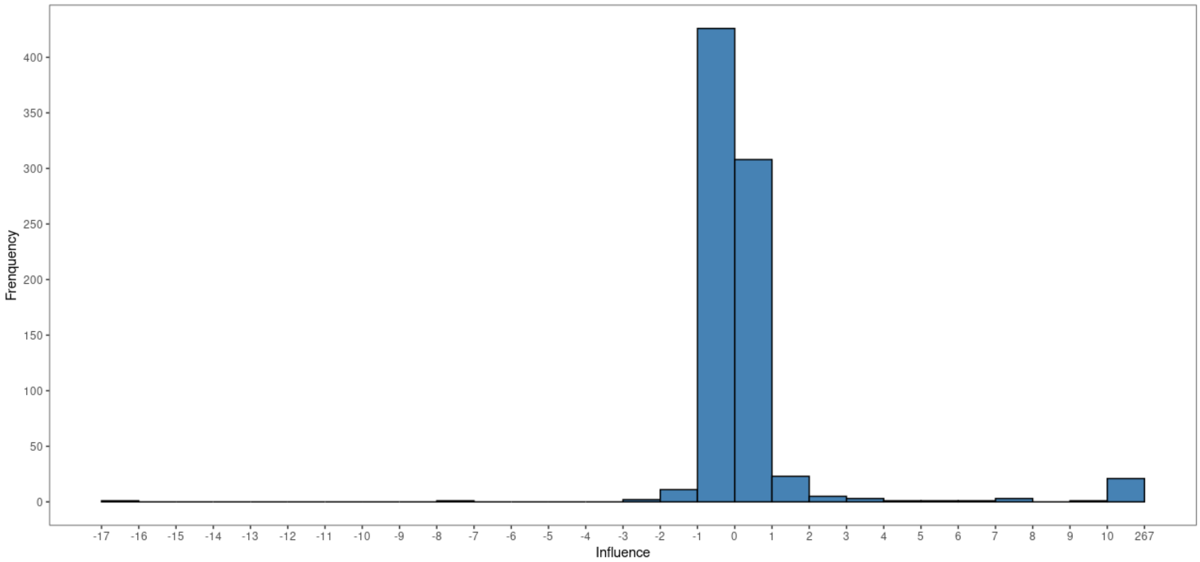
This panel shows top 20 investors ranked either by ESG influence or elasticity. Section 5.2 details the estimation of influence or elasticity. Specifically, elasticity of each investor is estimated from equation (2). Influence is elasticity multiplied by average number of firms with at least 1% institutional ownership by the investor.

Investors ranked by influence	**Influence**	**Investors ranked by elasticity**	**Elasticity**
Blackrock	266.268	Investors Group	0.141
State Street	250.812	Capital Group	0.125
Northern Trust	109.421	State Street	0.118
Fidelity Investments	75.591	Roosevelt Investments	0.115
Invesco	73.780	Blackrock	0.112
T. Rowe Price	55.050	Invesco	0.098
J.P. Morgan	52.320	J.P. Morgan	0.090
Wellington Management	47.507	Goldman Sachs	0.083
Goldman Sachs	47.012	Thompson Davis & Co	0.083
Capital Group	34.109	TIAA	0.081
TIAA	15.603	TOBAM	0.081
Columbia Threadneedle	13.751	Bronson Point Management	0.075
AQR	13.576	State of Texas	0.075
American Century Investments	13.057	T. Rowe Price	0.074
Morgan Stanley	12.663	Northern Trust	0.072
Franklin Templeton	11.319	Voya Financial	0.071
Dimensional Fund Advisors	11.045	Fidelity Investments	0.070
Principal	10.241	Fiduciary Trust International	0.070
AllianceBernstein	10.169	Bessemer Trust	0.067
Voya Financial	10.157	AQR	0.062

Table 12: ESG influence by investor (continued)

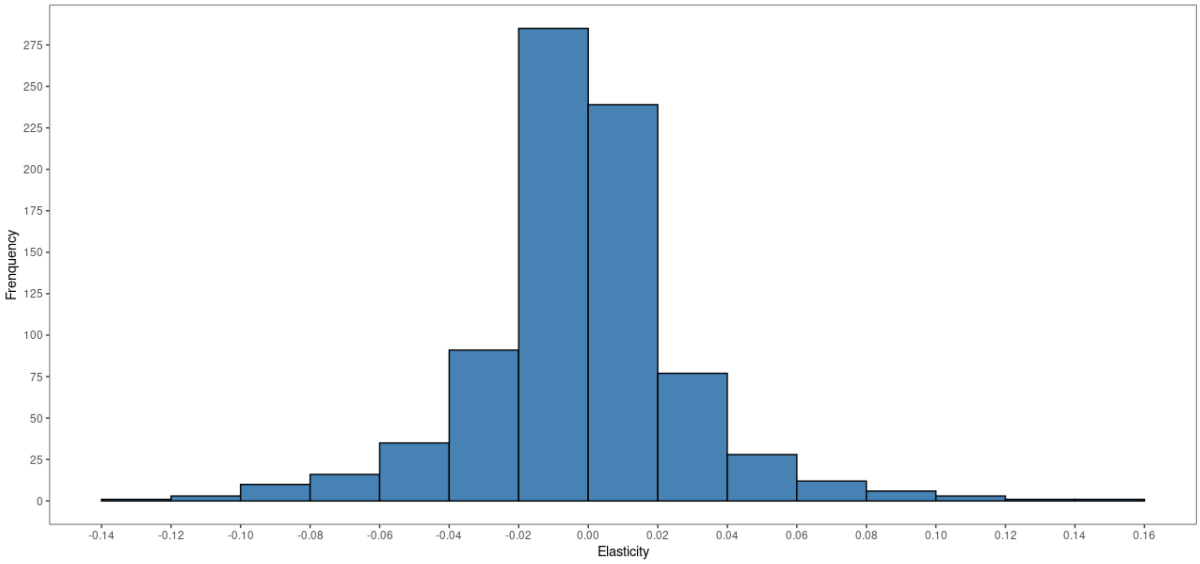
Panel B: Histogram of investor's influence

In this panel, we plot a histogram of estimates of investor's influence. The Y-axis shows the frequency that the values occur within the intervals set by the X-axis. The X-axis shows the intervals of estimated influence, and the width of each interval is 1 for influence between -17 and 10. For influence above 10, we aggregate the frequency of values in the rightmost interval.



Panel C: Histogram of investor's elasticity

In this panel, we plot a histogram of estimates of investor's elasticity. The Y-axis shows the frequency that the values occur within the intervals set by the X-axis. The X-axis shows the intervals of estimated elasticity, and the width of each interval is 0.02.



Internet Appendix

Appendix A

Appendix A.1 Intent Data Spikes and ESG Topics.

This appendix illustrates how we calculate the measures $Spikes^{ESG}$ (count of Spike scores of ESG topics which are at least 80) and $Spikes^{Not\ ESG}$ from the weekly Spike score in The Company's Topic Interest model. In the following example, we simplify by considering a domain-month with 4 weeks and 5 topics (2 ESG topics and 3 non-ESG topics) and show the Spike score of each topic-week, but it also applies to other more general circumstances. As seen in the following table, the Spike score can be missing.

Appendix Table A.1

	Charitable Giving	**Climate Change**	**Innovation**	**ROA**	**Brand Loyalty**
Week1	17	51	24	**80**	NA
Week2	20	69	56	61	31
Week3	21	**83**	78	49	36
Week4	29	**85**	**81**	36	**88**

The two ESG topics are "Charitable Giving" and "Climate Change," and the three non-ESG topics are "Innovation," "ROA," and "Brand Loyalty." We highlight the Spike score which is larger than or equal to 80. For the topic "Climate Change," there are two weeks with Spike score that is larger than or equal to 80, so the count would be 2. Doing the same calculations to other topics and doing aggregations, $Spikes^{ESG}$ is 2 and $Spikes^{Not\ ESG}$ is 3.

Appendix A (continued)

Appendix A.2: Understanding What and Who Reads?

To aid the interpretation of our ESG reading intensity measures, we conduct two types of analysis using more granular data at the level of individual content interaction events. For this analysis, unfortunately the data that we have access to is too short for our main analysis (due to storage costs, the company does not retain the full history), but it permits us to understand the types of sites generating ESG content in our sample and users that generate.

First, we examine what types of websites comprise ESG content. For a sub-sample of data from 2018-2020, we examine the top-level domains (e.g., www.google.com is the TLD of the URL https://www.google.com/search/). We cannot reveal the exact publisher names due to confidentiality agreements. However, the top-level domains permit to understand the nature of the sites that generate ESG content. We first ask whether ESG reading intensity consists of firms reading about ESG news of their own firm as opposed to reading generally about ESG. For example, reading intensity may be high at Apple on "Carbon Emissions" because Apple employees are reading an article in a financial newspaper which documents Apple's recent sustainability-oriented product enhancements. Thus, we first we break down the top-level domains into two types of sites: financial websites and others. Financial websites are likely ones where firm-specific news originates.

We find that financial websites (including general interest news websites that contain a great deal of financial content) generate 23% of the ESG content in our dataset. This suggests that at least 77% of the content we study comes from sources other than financial newspapers. Additionally, while we cannot disclose these top-level domains, examining the nature of the top-level domains, we conclude the top ten sites that disproportionately cover ESG news in our sample are legal reference sites, sites consisting of opinion articles on ESG issues (related to diversity, ethics in business or politics), or general interest news (with a slight left-leaning). Second, for these articles from financial newspapers, we obtain the original news headline, if still available, and merge with Ravenpack. Filtering on articles found in both datasets, only 20% of those articles (5% of the total) comprise articles that are highly relevant to a publicly listed firm (firm relevance score of 70 or above in Ravenpack). Hence, we conclude that the majority of ESG reading in our dataset is not firm-specific news but more likely to be general punditry or news coverage of ESG issues.

Second, we ask who within the firm reads about ESG. We pursue this question in two ways. First, we examine data at the "profile-level" as segmented by occupation. For each profile captured by The Company using a persistent browser session, The company classifies the functional area of the employee based the totality of websites the user visits is computed. The data cannot disentangle managers but can disentangle broad functional areas (Research, Legal, Human Resources, etc.). We have data since January 2020. In Panel A of Appendix Table A.2, we display the functional areas that have the highest intensity of reading about topics in a particular ESG category. First, for every publicly listed in firm in our sample, we take the average percentage of ESG-related total articles read by profiles designated as belonging to a functional area. We then average this percentage across all firms. The results suggest that the estimated functional area that reads content of a particular ESG activity tends to be relevant. For example, reading about Labor tends to come from people who work in HR. This buttresses the interpretation that employees within the firm who read tend to be those whom the content is most relevant for.

Second, we look at the fraction of the employees in a firm that reads topics in a given category. Panel B in Appendix Table A.2 displays these results. What we show is the number of browser profiles reading a topic on a given day divided by the profiles observed at the company that day. That is, what percentage of browsers (i.e. what percentage of employees) reads, relative to the average topic. In general, the results suggest that, within the same firm, ESG topics tend to be more broadly read, particularly for topics about Equality and Diversity.

We infer from this analysis that we are not capturing the sparse activity of a handful of managers but rather what appears to be an organization-wide endeavor to learn about ESG issues. Whether such attentional activity emanates from a grassroot interest within the firm or as part of a top-down directive remains unclear. The case of Google and the cancellation of Project Maven, a U.S. Department of Defense artificial intelligence project to facilitate drone strikes, suggests that calls for corporate social responsibility can be a bottom-up endeavor versus a top-down. Hence, even if we could identify managers, making this inference would be difficult as the two groups can influence one another. That being said, the fact that reading of ESG spans multiple different divisions within the firm suggests that fluctuation in ESG reading reflects not only the vagaries of management or the board of directors.

Appendix Table A.2

Panel A: Breakdown of Functional Area by Percentage of Time Allocation on ESG Topic

In this Panel, we present a breakdown of ESG reading by each individual, anonymized browser. For a subset of browser sessions with requisite data, the company estimates a functional area for that user. We retain the 12 most common functional areas. For each firm, we calculate the percentage of time allocated toward a ESG topics in the stated category, and then average across firms to form the rankings below. The functional areas that read the most on a given topic are listed first. The sample period of available data is 2020 to May 2021.

	Corporate Governance	**Customer Relation**	**Labor Relations**	**Environment**	**Equality Diversity**	**Data and Sensitive Information Protection**
1	HR	Marketing	HR	Research	Creative	Building and Grounds Maintenance
2	Legal	Sales	Operations	Building and Grounds Maintenance	HR	Information Technology
3	Marketing	Operations	Creative	Engineering	Legal	HR
4	Building and Grounds Maintenance	Research	Scientists	Operations	Scientists	Legal
5	Finance	Legal	Legal	Sales	Marketing	Scientists
6	Operations	Information Technology	Marketing	Finance	Sales	Sales
7	Scientists	Finance	Sales	Marketing	Operations	Research
8	Information Technology	Engineering	Building and Grounds Maintenance	Legal	Engineering	Marketing
9	Sales	HR	Research	Scientists	Information Technology	Engineering
10	Engineering	Building and Grounds Maintenance	Engineering	Information Technology	Research	Operations
11	Research	Creative	Information Technology	HR	Finance	Creative
12	Creative	Scientists	Finance	Creative	Building and Grounds Maintenance	Finance

Appendix Table A.2 (continued)

Panel B: Average Percentage of Browsers Reading a Topic on a Daily Basis

In this panel, we present a breakdown of the number of browsers within a firm reading a topic on a given day relative to the number of browsers observed in a firm that day. For each firm-ESG topic category, we calculate the average number of browsers reading a topic (zero if missing) relative to the number of users observed that day. We then average across firms. The browser share is a percentage of browsers that read a given topic relative to the browsers observed. The sample is May 2016 to December 2020, consisting of publicly listed firms.

ESG Category	**Browser Share**
CSR	0.451
Corporate Governance	0.331
Customer Relation	0.662
Data and Sensitive Information Protection	0.324
Environment	0.476
Equality Diversity	0.972
Labor Relation	0.831
Not ESG	0.392

Appendix B: Climate Change Betas

We calculate climate change betas for each stock in our sample and compare it to the amount of reading activity done on relevant environmental issues. As our baseline variable, we use their monthly variable $chneg\_ar1\_innovation$ which is the residual of AR(1) model of Crimson Hexagon's negative sentiment climate change news index. Crimson Hexagon (CH) is an AI-powered consumer insights company which has collected over one trillion news articles and social media posts. The authors construct the CH negative sentiment climate change news index as the share of news articles that are both about "climate change" and assigned to the "negative sentiment" category.

The variable we use can be viewed as innovations or shocks to negative climate change news in each month, which covers the period from May 2015 to May 2018. We multiply it by 100 in the analysis. Then, we estimate β from the following time-series regression for each individual stock and call it "climate beta:"

$$ret = \alpha + \beta * chneg\_ar1\_innovation + \emptyset * X + \varepsilon. \quad (3)$$

ret is a monthly stock return for a stock. X are control variables including monthly CRSP value-weighted stock index return, the stock's logarithm of market capitalization and book-to-market ratio. We control for these variables because it is argued in Engle, Giglio, Kelly, Lee, and Stroebel (2020) that they are associated with both climate change shocks and individual stock returns.

Appendix C. Alternative specifications

Appendix Table C.1 Firm's ESG reading and alternative real ESG outcomes

This table presents alternative measures for real ESG outcomes in Table 4. In Models (1) to (2), the outcome variable is the percentage or indicator of facilities that report newly implemented source reduction activities. Model (3) shows result for log-number of worker benefits in IRS Form 5500. In Model (4), the outcome variable is number of OSHA unprogrammed inspections per establishments. Model (5) shows result for log-number of OSHA penalties. In Model (6), we present results for dollar amounts of OSHA penalties scaled by firm size (proxied by firm sale). In Model (7), we present results for chemical-level toxic releases scaled by firm size (proxied by firm sale). Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	% Green facilities	**1{Green facilities}**	**Log(1 + # worker benefits)**	**OSHA inspections**	**Log(1 + # OSHA penalties)**	**OSHA penalty amount scaled by firm size**	**Toxic releases scaled by firm size**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(1+Spikes ^{ESG})	0.005***	0.030***	0.009**	-0.039***	-0.063**	-0.138**	-0.121*
	(0.001)	(0.008)	(0.004)	(0.012)	(0.032)	(0.064)	(0.070)
Log(1+Spikes $^{Not\ ESG}$)	-0.002	-0.016*	-0.008**	0.028**	0.069**	0.110*	0.061
	(0.002)	(0.009)	(0.004)	(0.012)	(0.031)	(0.063)	(0.055)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	
Chemical×Year FE							Yes
Chemical×Firm FE							Yes
Observations	14,590	14,590	10,430	10,403	10,403	10,403	242,699
Adjusted R^2	0.447	0.493	0.776	0.293	0.294	0.179	0.760

Appendix C. Alternative specifications (continued)

Appendix Table C.2 ESG category decomposition for real ESG outcomes

This table presents the relationship between firm's reading intensity on different ESG categories and various real ESG outcomes. The ESG categories we present are Environment, Labor Relation, Equality and Diversity, CSR, Customer Relation, Compliance, Corporate Governance, Data and Sensitive Information Protection, Cybersecurity. In Models (1) to (2), the outcome variable is the (log) number of facilities that report newly implemented source reduction activities. Models (3) and (4) show results for log-number of worker benefits in IRS Form 5500. In Models (5) and (6), the outcome variable is number of OSHA penalties per establishments. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Green Facilities		**Employee Contributions**		**OSHA Penalties**	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes^{Environment})	0.016***	0.013**		0.007		0.004
	(0.005)	(0.005)		(0.013)		(0.008)
Log(1+Spikes$^{Labor\ Relation}$)		0.007	0.027***	0.026**	-0.018***	-0.013**
		(0.009)	(0.006)	(0.012)	(0.005)	(0.006)
Log(1+Spikes^{Diversity})		0.003		0.0005		-0.011*
		(0.006)		(0.005)		(0.006)
Log(1+Spikes^{CSR})		-0.0003		0.005		-0.002
		(0.006)		(0.008)		(0.006)
Log(1+Spikes$^{Customer\ Relation}$)		-0.012**		-0.022***		0.012
		(0.005)		(0.008)		(0.010)
Log(1+Spikes^{Compliance})		-0.002		-0.005		0.0003
		(0.006)		(0.009)		(0.007)
Log(1+Spikes$^{Corporate\ Governance}$)		0.005		0.018**		-0.003
		(0.006)		(0.008)		(0.007)
Log(1+Spikes^{Data})		0.005		0.001		-0.011
		(0.006)		(0.011)		(0.007)
Log(1+Spikes^{Cybersecurity})		-0.003		-0.009		0.0004
		(0.006)		(0.011)		(0.008)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,590	14,590	10,430	10,430	10,403	10,403
Adjusted R^2	0.585	0.585	0.937	0.937	0.249	0.250

Appendix C. Alternative specifications (continued)

Appendix Table C.3 Using 70 as threshold to indicate high ESG interest (horserace)

This table shows how firm's ESG reading affect real ESG outcomes. *Log(1+Spikes* ^{ESG,70}*)* indicates that we use 70 (instead of 80) as threshold to indicate high interest in any ESG topic and calculate number of spikes. The outcome variables are the same as those in Table 4. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Green Facilities		**Employee Benefits**		**OSHA Penalties**	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes ^{ESG,70}*)*	0.017***	-0.040***	0.025***	-0.049**	-0.015**	0.029
	(0.005)	(0.012)	(0.007)	(0.019)	(0.006)	(0.022)
Log(1+Spikes ^{ESG,80}*)*		0.051***		0.066***		-0.031*
		(0.011)		(0.017)		(0.018)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	14,590	14,590	10,430	10,430	10,403	10,403
Adjusted R^2	0.585	0.586	0.937	0.937	0.224	0.231

Appendix Table C.4 1-year ahead prediction on real ESG outcomes

This table shows how firm's 1-year ahead ESG reading affect real ESG outcomes. The outcome variables are the same as those in Table 4. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Green Facilities		**Employee Benefits**		**OSHA Penalties**	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes ^{ESG}*)*	0.021***	0.030***	0.019***	0.029**	-0.012***	-0.031**
	(0.007)	(0.010)	(0.006)	(0.014)	(0.004)	(0.016)
Log(1+Spikes $^{Not\ ESG}$*)*		-0.010		-0.011		0.021
		(0.010)		(0.014)		(0.016)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,167	10,167	7,518	7,518	7,218	7,218
Adjusted R^2	0.537	0.537	7,518	7,518	0.359	0.360

Appendix C. Alternative specifications (continued)

Appendix Table C.5 Placebo test on real ESG outcomes

This table shows how firm's ESG reading affect real ESG outcomes, and ESG reading is aligned 1 year after real ESG outcomes. The outcome variables are the same as those in Table 4. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Green Facilities		**Employee Benefits**		**OSHA Penalties**	
	(1)	(2)	(3)	(4)	(5)	(6)
Log(1+Spikes ^{ESG})	0.024***	0.007	0.003	0.028	0.0003	0.017
	(0.006)	(0.013)	(0.009)	(0.021)	(0.007)	(0.022)
Log(1+Spikes$^{Not\ ESG}$)		0.019		-0.025		-0.017
		(0.014)		(0.020)		(0.022)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,181	10,181	7,515	7,515	7,589	7,589
Adjusted R^2	0.547	0.547	0.939	0.943	0.259	0.259

Appendix Table C.6 Firm's ESG reading and real ESG outcomes --- *Daily Aggregates*

This table shows how firm's ESG reading affect real ESG outcomes. In this table, we use daily reading counts data instead of the 'Spike' measure. Our sample period begins in May 2016, reducing our observations. Our measure is *Log(1+Reading ^{ESG}/Sale)*, in which *Reading ^{ESG}/Sale* is total records of ESG topics scaled by firm's sale last year. Across all columns we control firm's reading of non-ESG topics *Log(1+Reading $^{Not\ ESG}$/Sale)* and logarithm of firm sale (*Size*). Outcome variables are the same as those used in Table 4. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Green Facilities	**Employee Benefits**	**OSHA Penalties**
	(1)	(2)	(3)
Log(1+Reading ^{ESG}/Sale)	0.030**	0.038**	-0.040***
	(0.012)	(0.019)	(0.014)
Log(1+Reading $^{Not\ ESG}$/ Sale)	-0.019*	-0.028	0.021*
	(0.010)	(0.018)	(0.011)
Size	0.026**	0.350***	-0.025***
	(0.011)	(0.048)	(0.007)
Year FE	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes
Observations	8,917	7,223	6,293
Adjusted R^2	0.556	0.951	0.370

Appendix C. Alternative specifications (continued)

Appendix Table C.7 Firm's ESG reading and ESG rating --- *Daily Aggregates*

This table shows how firm's ESG-related reading is associated with ESG rating. In this table, we use daily reading counts data instead of the 'Spike' measure. Our sample period begins in May 2016, reducing our observations. Our measure is $Log(1+Reading^{ESG}/Asset)$, in which $Reading^{ESG}/Asset$ is total records of ESG topics scaled by firm's total asset last year. Across all columns we control firm's reading of non-ESG topics $Log(1+Reading^{Not\ ESG}/Asset)$ and logarithm of total assets (*Size*). Outcome variables are the same as those used in Tables 5, 6, and 7. In Models (5) and (6) we multiply outcome variables by 100. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Refinitiv^{Combined}	**RRI**	**KLD1**	**KLD2**	**KLD3**
	(1)	(2)	(3)	(4)	(5)
$Log(1+Reading^{ESG}/Asset)$	0.738*	-0.436**	0.171***	0.057***	0.052***
	(0.431)	(0.180)	(0.062)	(0.016)	(0.017)
$Log(1+Reading^{Not\ ESG}/Asset)$	-0.471	0.138*	-0.039	-0.018*	-0.024**
	(0.333)	(0.077)	(0.035)	(0.009)	(0.009)
Size	-0.179	-0.075	0.213	0.149***	0.165***
	(1.237)	(0.699)	(0.177)	(0.050)	(0.055)
Month FE		Yes			
Year FE	Yes		Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Firm FE		Yes			
Observations	4,368	41,188	5,424	5,424	4,343
Adjusted R^2	0.738	0.756	0.788	0.586	0.663

Appendix Table C.8 KLD strengths and concerns

This table shows how firm's ESG reading is associated with KLD strengths and concerns separately. *Str2* is defined as Str/(Str + Con) while *Con2* is defined as Con/(Str + Con), and they would be 0 if there are no strengths or concerns for any firm-year. *Str3* and *Con3 are* defined in the similar way, but they would be treated as missing if there are no strengths or concerns for any firm-year. Str and Con are number of strengths and concerns the firm have in each year respectively. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Str2	***Con2***	***Str3***	***Con3***
	(1)	(2)	(3)	(4)
$Log(1+Spikes^{ESG})$	1.354	-2.222**	2.683**	-2.683**
	(1.389)	(1.014)	(1.045)	(1.045)
$Log(1+Spikes^{Not\ ESG})$	1.542	-0.464	0.142	-0.142
	(1.329)	(0.971)	(0.996)	(0.996)
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	7,396	7,396	5,894	5,894
Adjusted R^2	0.566	0.582	0.689	0.689

Appendix C. Alternative specifications (continued)

Appendix Table C.9 Investor's influence on firm's ESG Reading --- Top 5 investors vs others

This panel shows how investor's ESG-related reading is associated with firm's ESG-related reading at the topic-quarter level. Across all columns the outcome variable is $1\{Spike^{Firm}\}$, which is a dummy variable indicating whether the firm has a Spike score that is larger than or equal to 80 for each topic-quarter. Similar dummy variables are constructed for top 5 investors and rest investors of the firm. The heading from Model (1) to 6 represent the ESG categories used in the analysis. We multiply the outcome variable by 100. Standard errors clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	$1\{Spike^{Firm}\}$					
	All	**All**	**Env**	**Labor**	**Social**	**Gov**
	(1)	(2)	(3)	(4)	(5)	(6)
$1\{Spike^{Inv1}\}$	7.157***	6.668***	6.053***	6.033***	4.995***	7.878***
	(0.329)	(0.330)	(0.333)	(0.301)	(0.359)	(0.386)
$1\{Spike^{Inv2}\}$	5.654***	5.197***	4.663***	4.538***	3.356***	6.451***
	(0.206)	(0.207)	(0.216)	(0.189)	(0.231)	(0.266)
$1\{Spike^{Inv3}\}$	5.224***	4.752***	4.382***	4.185***	3.017***	5.805***
	(0.162)	(0.163)	(0.184)	(0.158)	(0.189)	(0.208)
$1\{Spike^{Inv4}\}$	5.185***	4.702***	4.479***	4.263***	2.954***	5.541***
	(0.164)	(0.165)	(0.183)	(0.163)	(0.181)	(0.206)
$1\{Spike^{Inv5}\}$	5.072***	4.565***	4.248***	4.094***	3.183***	5.365***
	(0.155)	(0.156)	(0.179)	(0.153)	(0.172)	(0.196)
$1\{Spike^{Rest\ Inv}\}$	1.984***	1.370***	1.042***	1.409***	0.937***	1.492***
	(0.073)	(0.070)	(0.091)	(0.073)	(0.069)	(0.079)
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Topic FE		Yes	Yes	Yes	Yes	Yes
Observations	14,251,321	14,251,321	1,932,765	4,889,983	1,882,312	5,546,261
Adjusted R^2	0.040	0.047	0.051	0.041	0.042	0.062

Appendix D: Firm's disclosure and ESG performance

In Appendix D, we present our results on disclosure and on each of our three outcome variables. The results in Model (1) indicate that the coefficient of *Disclosure* itself is positive but not statistically significant. That disclosure by itself is not positively correlated with improvements echoes concerns of consumers and critics of ESG ratings in that disclosure – the mere assertion of a policy – may not result in actual change at the firm. In Model (2), we interact disclosure with our measure of ESG-related attention. Interestingly, Model (2) suggests that the complementarity of disclosure and attention is positively associated with the firm's improvement in greenness. Thus, our finding suggests that our ESG indicator may not only be useful in predicting real outcomes, but it may also be useful in predicting which firms that disclose follow-through and convert these disclosures to improvements in real outcomes. This may help investors and policymakers who are concerned that firms' disclosures serve to 'greenwash.' Models (3) and (4) repeat the analysis on IRS Form 5500 and present similar results. Models (5) and (6) report findings for OSHA-related outcomes. Model (5) suggests that like the prior 2 panels, disclosure is unrelated to the real outcome variable. Model (6) is insignificant, suggesting that firms which pay more attention to ESG and disclose are not differentially likely to avoid labor-related penalties.

Appendix D: Firm's disclosure and ESG performance (continued)

Appendix Table D.1

This table shows how firm's ESG disclosure and reading are linked to real ESG outcomes. Models (1) to (2) present results for log-number of facilities reporting newly implemented source reduction activities, Models (3) to (4) present results for log-number of company contributions to employee's pension plans scaled by firm size, Models (5) to (6) show results for OSHA penalties per establishments. The outcome variables are the same as those in Table 4. Disclosure is a dummy variable indicating whether the firm has a sustainability report in any year. Standard errors are clustered on firm. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

	Greenness		**Contributions**		**OSHA Penalties**	
	(1)	(2)	(3)	(4)	(5)	(6)
Disclosure	0.034	0.190	0.008	0.251**	0.001	-0.221
	(0.035)	(0.149)	(0.018)	(0.114)	(0.015)	(0.166)
Disclosure×Log(1+Spikes^{ESG})		0.094**		0.113***		-0.014
		(0.045)		(0.032)		(0.018)
Disclosure×Log(1+Spikes$^{Not\ ESG}$)		-0.080*		-0.103***		0.039
		(0.044)		(0.032)		(0.030)
Log(1+Spikes^{ESG})		0.034		0.050***		-0.024
		(0.028)		(0.016)		(0.017)
Log(1+Spikes$^{Not\ ESG}$)		-0.018		-0.032**		-0.005
		(0.028)		(0.015)		(0.021)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,597	3,597	10,430	10,430	4,715	4,715
Adjusted R^2	0.447	0.448	0.937	0.937	0.396	0.399

Appendix E: Comparing Investor's ESG reading and Against Portfolio Holdings of ESG Stocks

This table presents a horse-race between investor attention measures based on internet research versus attention based on investors holdings of highly ESG-rated stocks. The attention measure is either an equal-weighted or value-weighted portfolio ESG performance before the shareholder meeting. Four sets of columns reflect four different alternative ESG ratings. In Models (1) to (8), the outcome variable "Fund vote for ESG" is a dummy variable indicating whether the fund votes for ESG-friendly proposals or votes against anti-social proposals. We standardize all explanatory variables to compare their economic significance and we multiply the outcome variable by 100 in all models. Standard errors are clustered on shareholder's meeting. *, **, *** indicate statistical significance at the 10%, 5% and 1% level, respectively. All continuous variables are winsorized at 1% level.

Appendix Table E.1

	Fund vote for ESG							
	KLD1	**KLD1**	**Refinitiv^{Combined}**	**Refinitiv^{Combined}**	**Current RRI**	**Current RRI**	**Peak RRI**	**Peak RRI**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log(1+Spikes^{ESG})	2.586***	2.745***	2.651***	2.741***	2.816***	2.724***	2.821***	2.616***
	(0.769)	(0.770)	(0.767)	(0.770)	(0.772)	(0.766)	(0.772)	(0.757)
Log(1+Spikes$^{Not\ ESG}$)	-1.363**	-1.272*	-1.434**	-1.313*	-1.565**	-1.561**	-1.552**	-1.610**
	(0.690)	(0.685)	(0.686)	(0.681)	(0.685)	(0.684)	(0.686)	(0.684)
VW Score^{ESG}	-1.721***		-0.918*		0.016		0.555	
	(0.437)		(0.529)		(0.447)		(0.364)	
EW Score^{ESG}		-2.435***		-1.958***		-0.973**		-1.547**
		(0.427)		(0.403)		(0.494)		(0.641)
ISS Vote	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Management Vote	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Meeting FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Investor FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	289,870	289,870	290,151	290,151	287,688	287,688	287,688	287,688
Adjusted R^2	0.588	0.588	0.588	0.588	0.589	0.589	0.589	0.589

Appendix F: Additional summary statistics

Appendix Table F.1 Summary statistics on outcome variables in Table 4

Statistic	**N**	**Mean**	**St. Dev.**	**Min**	**25^{th}**	**75^{th}**	**Max**
Log(1+ # Green facilities)	14,590	0.276	0.539	0.000	0.000	0.693	2.639
Contributions divided by firm size	10,430	9.169	1.122	5.967	8.529	9.930	11.419
# OSHA penalties per establishments	10,403	0.088	0.237	0.000	0.000	0.046	1.500