

Internet Appendix for “Female Equity Analysts and Corporate Environmental and Social Performance”

Kai Li, Feng Mai, Gabriel Wong, Chelsea Yang, and Tengfei Zhang

Appendix IA.A Fine-tuning FinBERT Using Active Learning

To capture analyst monitoring through their equity research and access to management, we apply novel machine learning techniques to 2,434,739 analyst reports and 129,302 earnings conference calls. Specifically, we employ active learning, a human-in-the-loop machine learning approach, to develop two domain-specific E&S text classification models to capture analysts’ writing (in analyst reports) and questions (during earnings conference calls) about corporate E&S performance.

1. Preprocessing analyst reports and earnings calls

We download 2,434,739 reports over the period 2004-2020 from Thomson One’s Investtext database. The reports are in PDF format. We use GROBID (<https://github.com/kermitt2/grobid>), an open-source software, to extract structured information from PDF documents and transform this information into XML documents. The XML documents are then stripped of information identified as tables, annexes, notes, and author information; the main content is converted to plain text. We further split text into sentences using OpenNLP’s sentence segment module, a built-in function in GROBID.

We download 129,302 earnings call transcripts over the period 2007-2020 from Capital IQ’s Transcripts database. Given that E&S-related questions raised by an analyst during calls often involve multiple sentences, we opt to use an entire question as the unit of analysis for earnings conference call transcripts. This approach helps preserve valuable contextual information that would be lost through sentence-level analysis.

We hereafter refer to a sentence in analyst reports or a question in earnings conference call transcripts as a *passage* of text.

2. FinBERT: An introduction

Our approach builds on FinBERT (Huang, Wang, and Yang 2023), a state-of-the-art large language model pre-trained on financial text. The FinBERT model is based on the same transformer architecture of BERT (Devlin et al. 2019), a pre-trained language model that has achieved impressive results on a wide range of NLP tasks. The transformer architecture consists of multiple layers of self-attention mechanisms and feed-forward neural networks. This architecture improves the model’s ability to capture long-range dependencies between words in text and facilitates more efficient parallel computations, resulting in better performance than conventional neural network-based models.

The BERT model is pre-trained on a large corpus of text, in which it learns from two tasks that can be constructed from the corpus. The first task is masked language modeling. In this task, the model predicts the identity of words that have been randomly replaced with a mask symbol (e.g., [MASK]) in a sentence. This task is designed to help the model learn the meaning of individual words and how they fit into the context of a sentence. The second task is next sentence prediction. In this task, the model is trained with a training data set in which half of the times sentence B is the actual sentence that follows sentence A, and the other half of the times B is a randomly chosen sentence from the corpus. This task helps the model learn the larger document context and better understand the relationships between different sentences in the document.

The key difference between the BERT and FinBERT models is the training data used for pre-training. While BERT is trained on general corpora, such as books and Wikipedia, FinBERT is trained on a

specialized collection of financial text, including annual and quarterly reports, analyst reports, and earnings conference calls. These domain-specific training corpora allow FinBERT to better capture the unique language and terminology used in the financial domain.

After pre-training, the BERT (FinBERT) model can generate a contextualized embedding vector for each sentence, which can be further fine-tuned and used as classification features for other tasks, such as text classification. Because the model learns semantic (e.g., the meanings of words) and syntactic (e.g., the phrases and the compositions of sentences) information from a large corpus in the pre-training step, Huang, Wang, and Yang (2023) show that the fine-tuning step requires only a relatively small training sample to achieve a high accuracy of text classification. Their experiments also demonstrate that for domain-specific tasks, such as financial text sentiment classification, the FinBERT model outperforms the generic BERT model.

3. Constructing domain-specific training examples via active learning

Our goal is to train a three-class classifier that can take a passage of text, from either reports or calls, as input, and predict its probability of pertaining to environmental issues (E), social issues (S), or neither (Non-E&S).

To fine-tune the FinBERT model of Huang, Wang, and Yang (2023) using our two corpora, we employ *active learning*, a human-in-the-loop machine learning approach, to find domain-specific training examples. We then use these domain-specific training examples to fine-tune two different E&S classification models, one for analyst reports and the other for earnings conference calls.

Figure IA1 presents a flowchart of the active learning process. In Step 1, we use keywords related to E&S issues to generate a set of initial training examples. Passages containing these keywords are tentatively labeled as positive examples (E or S), and random passages are used as negative examples (Non-E&S). In Step 2, we use these initial training data to fine-tune the FinBERT model into a *Noisy E&S model*. In Step 3, we use the *Noisy E&S model* to classify the initial training examples. Given the *Noisy E&S model*'s output, a subset of important examples is labeled by human annotators. In Step 4, we use these labeled examples to fine-tune the *Noisy E&S model* and produce the *Final E&S model* (Cormack and Grossman 2014). We describe the four steps in detail below.

Step 1. Constructing the initial training data sets

In Step 1, we search for relevant passages from reports and calls on corporate E&S practices using a keyword list. To build our keyword list of corporate E&S performance, we start with one of the earliest ESG databases – the RiskMetrics KLD database (before it was acquired by MSCI and its methodology was updated). The KLD User Guide in 2010 includes descriptions of different E&S practices. The keyword list captures the essence of each broadly defined E&S category.

To search for relevant passages pertaining to E&S practices, we develop search queries to return results that match the keywords, while excluding queries that are too broad. We employ Apache Solr (<https://solr.apache.org/>) to index the full text and to conduct the search. Apache Solr is an open-source search platform that allows for powerful full-text search using queries that support exact term matching, the wildcard operator (e.g., the * operator represents unknown characters), and Boolean logic (e.g., AND/OR operators). For example, we drop the keyword “environment” as it is more often used to describe the macro-economic environment that is not directly related to E&S. As another example, under the E&S practices regarding product, “product recall” is a keyword. We develop the query “product* & recall,” such that 1) the query identifies passages that not only match the exact phrase “product recall,” but also capture sentences that include the two words separately, such as “the firm initiated a voluntary recall of some potentially contaminated products;” and 2) the query excludes irrelevant passages that only contain “recall,” such as “we’re generating unusually high recall rates for advertisers’ brands and unusually high recall rates for advertisers’ messages.”

Table IA3 in the Internet Appendix lists queries of corporate E&S practices.

Using these queries, we are able to find representative in-domain passages that are likely to be related to E&S issues with minimal human intervention. For analyst reports, we find 19,555 E-related and 4,817 S-related sentences. For earnings conference calls, we find 1,201 E-related and 123 S-related questions. To construct the initial training data set for each corpus, it is also necessary to include Non-E&S examples. To do this, we randomly select an additional 20,000 passages that did not match any E or S queries for each corpus, to serve as Non-E&S examples.

Step 2. Fine-tuning FinBERT into a Noisy E&S model

In Step 2, we use the initial training sample, including both the E&S and Non-E&S examples identified in Step 1, to fine-tune the pre-trained FinBERT model into a *Noisy E&S model*. The initial training data are randomly split into 80%/10%/10% train/validation/test subsets. We use the training set to fine-tune the model, the validation set to assess the performance of the model after each epoch (i.e., an iteration of the entire training data set), and the test set to evaluate the final performance of the model. The Receiver Operating Characteristic (ROC) curve is a probability curve that plots the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings and separates the “signal” from the “noise”. We use the area under the ROC curve (AUC) metric on the validation set to evaluate the performance of the model after each epoch. The training process is terminated when the AUC fails to improve after an epoch on the validation set. This approach, known as early stopping, can avoid overfitting.

The resulting model is considered a *Noisy E&S model* due to the training data containing both false negatives and false positives. Since search queries are restrictive, some negative examples (randomly drawn passages that do not match any E or S queries) in the initial training sample may be classified as relating to E or S. In addition, not all passages matching the queries in Table IA3 are definitely classified as relating to E or S.

Next, we use this *Noisy E&S model* to identify important examples for a human annotator to label.

Step 3. Identifying important examples for human annotation

In Step 3, we identify important examples for human annotation using the *Noisy E&S model*. To do this, we apply the *Noisy E&S model* from Step 2 to all examples in the initial training sample. This allows us to obtain a predicted probability vector for each example, indicating the probability that an example belongs to one of the three classes (i.e., E, S, or Non-E&S). These predicted probabilities can then be used to identify examples that are important for human annotation.

There are two common protocols for identifying important examples (Cormack and Grossman 2014). The first is continuous active learning (CAL), which entails labeling the examples that the model is most certain about (i.e., the examples with the highest predicted probabilities in either class). The second is simple active learning (SAL), which entails labeling the confusing examples that the model is unsure of (i.e., the examples with similar predicted probabilities across different classes, which can be measured using the entropy of predicted probabilities). Intuitively, when the model is trained on human-labeled examples that it has previously been most certain about, we strengthen its existing knowledge and help correct the most obvious errors, e.g., passages that match search queries but are not related to E&S given the context. On the other hand, labeling unsure examples can help the model identify the boundary between difficult cases. In the finance literature, the SAL approach is used by Kölbl et al. (2022) to construct a training sample for climate risk disclosures. Combining these two protocols allows the model to focus on the most informative examples rather than random examples in the training sample, which can improve the accuracy and efficiency of model training.

For CAL, we sort the examples based on the predicted probabilities provided by the *Noisy E&S model* and select the top 500 examples with the highest predicted probabilities belonging to E (S), resulting

in 1,000 examples. For SAL, we calculate the entropy of the predicted probability vector for each example and select the top 500 examples with the highest entropy. Entropy is a measure of the uncertainty of a probability distribution, and it is calculated as the sum of $-p \times \log(p)$ over all classes where p is probability. An example with $P(E) = P(S) = P(\text{Non-E\&S}) = 0.33$ would have the highest entropy and be at the top of the SAL list. In total, the human annotators (authors of this paper) manually label 1,500 examples for each corpus.

Table IA4 in the Internet Appendix lists some important examples identified by active learning protocols (CAL and SAL), illustrating how the *Noisy E&S model*'s predictions and human labels correspond in different contexts.

Step 4. Fine-tuning the Noisy E&S model into the Final E&S model

In the final step, the human-annotated examples are used to further fine-tune the *Noisy E&S model* into the *Final E&S model*. This step follows the methodology outlined in Step 2 and is therefore omitted in the interest of brevity.

4. Choosing a classification threshold

Given a passage, our *Final E&S model* produces a continuous probability in the interval $[0,1]$ for each of the three classes (E, S, and Non-E&S). To obtain a discrete label from these scores, we require a threshold t_c , and assign a label $C \in [E, S, \text{Non-E\&S}]$ to any passage with a predicted probability $P(C) \geq t_c$. Using discrete labels allows us to identify individual passages related to E&S issues for further analysis.

Choosing an appropriate threshold t_c requires balancing precision and recall. A low threshold will be too loose and will identify more passages as relevant that are only tangentially related to E&S, resulting in a high recall but low precision (a high false positive rate). On the other hand, a high threshold will be too strict and will identify only a small number of passages, resulting in a high precision but low recall (a high false negative rate). Picking a threshold is also necessary as our initial training sample is highly unbalanced, with the number of non-E&S examples dominating the other two classes.

To select the threshold t_c , we consult existing literature on classifying E&S issues using textual data so that the fraction of E&S passages identified by our *Final E&S model* with the chosen threshold t_c is in line with the reported values in the literature.

After removing reports from firms not included in the main sample and removing short sentences whose length falls below the bottom decile (8 words), our final sample comprises 965,377 analyst reports. For analyst reports, we set $t_E = 0.01$ and $t_S = 0.01$. After applying these thresholds, we find that the fraction of reports writing about environmental issues is 22.1%, and the fraction of reports writing about social issues is 13.4%. As far as we are aware, we are the first in the literature to examine E&S-related discussions in analyst reports; there are no comparable statistics in the literature.

After removing calls from firms not included in the main sample and removing short questions whose length falls below the bottom decile (11 words), our final sample comprises 51,872 earnings calls. For calls, we set $t_E = 0.020$, and $t_S = 0.015$. After applying these thresholds, we find that the fraction of calls discussing environmental issues is 12.4%, and the fraction of calls discussing social issues is 31.9%. These values fall within the range of the reported values in the literature whereby the fraction of calls discussing environmental issues ranges between 7% and 58%, and the fraction of calls discussing social issues ranges between 7% and 45% (Raman, Bang, and Nourbakhsh 2020; Chava, Du, and Malakar 2021).

5. Model interpretability analysis

Figure IA4 provides a model interpretability analysis to shed light on the qualitative differences between the two corpora. We utilize the integrated gradients method, a recent development in interpretability techniques for neural networks (Sundararajan, Taly, and Yan 2017). The method determines which input features – in our case, tokens in raw text – are most important in the fine-tuned FinBERT models' classification of each sentence.²⁶ We sample a total of 5,000 sentences from reports and 5,000 questions from calls that have been classified as either E or S and compute the importance score of the tokens. We then average the importance scores across the sample sentences and questions for each corpus, providing a corpus-specific measure of the importance of each token.

In Panel A, we show that among the E-related sentences, environmental damages and remediation-related issues (e.g., *remediation, hazardous, ozone*) are relatively more important in reports, whereas pollution, climate, and greenhouse-related issues (e.g., *pollution, renewables, climate, greenhouse*) are more important during calls. In Panel B, we show that among the S-related sentences, both corpora emphasize the significance of employee-related issues such as layoffs, safety, and strikes, but they diverge on broader topics. Community relations and discrimination-related issues (e.g., *community, sex, discriminations*) are given more emphasis in reports, while corporate wrongdoing (e.g., *corruption, indigenous, violation*), and cybersecurity incidents (e.g., *hackers*) are more important during calls. Overall, the analysis indicates that the fine-tuned FinBERT models possess high face validity for both corpora, and that the relative importance of tokens varies depending on the context. These findings support our decision to fine-tune separate machine learning models for analyst reports and earnings calls.

6. Structure topic modeling

This section provides additional details on the Structural Topic Modeling (STM) analysis used to investigate thematic differences in how male and female analysts discuss environmental and social (E&S) issues. The STM approach, developed by Roberts et al. (2014), enables the discovery of latent topics within a corpus of documents while accounting for document-level covariates such as analyst gender. We use the approach to explore how gender influences the prevalence of various E&S themes in analysts' writing and questions.

The STM approach offers several key advantages over traditional topic modeling methods like Latent Dirichlet Allocation (LDA). First, by explicitly modeling document-level covariates as variables affecting the prior distribution of topic proportions and topic-word distributions, STM allows researchers to directly estimate the relationships between metadata and topical prevalence as well as topical content. This provides a principled way to test hypotheses about the factors that drive both how much documents discuss certain topics and the language used to discuss those topics. In contrast, LDA does not incorporate covariate information into the model. Second, the inclusion of covariates in STM can significantly enhance the interpretability of topics. By identifying topics that are strongly associated with metadata attributes of interest, such as document source or author characteristics, we can gain additional context to understand the substantive meaning and significance of topics.

We apply STM using the R package *stm* (Roberts et al. 2019) to four corpora: (1) sentences in analyst reports discussing environmental issues, (2) sentences in analyst reports discussing social issues, (3) paragraphs in earnings call questions related to environmental issues, and (4) paragraphs in earnings

²⁶ The integrated gradients method utilized in our analysis is conceptually similar to the SHAP (Shapley Additive exPlanations) method used in Erel, Stern, Tan, and Weisbach (2021). The advantage of using this method in our context is that it is computationally more efficient with differentiable models such as neural networks. Furthermore, it is well suited for cases in which the input space is high-dimensional or continuous, which is common in natural language processing tasks such as ours.

call questions related to social issues. For each corpus, the gender of the analyst is included as a covariate. The STM is then estimated using a variational expectation-maximization algorithm until convergence. To interpret and label the discovered topics, we examine the most prevalent words and documents associated with each topic. Table 6 lists the top words for each topic in each corpus, along with our assigned labels. These labels are based on collections of words that are strongly associated with each topic, using the score criteria which divides the log frequency of a word in a topic by its log frequency in other topics.

To assess the relationship between analyst gender and E&S topic prevalence, STM calculates differences in topic proportions for each topic between male and female analysts. This is done by first fitting the STM model with analyst gender as a covariate to obtain a posterior distribution over the topic proportions. The estimateEffect function in the *stm* package then performs posterior inference by simulating from the variational approximation and estimating the expected topic proportion for each topic when setting the gender covariate to male versus female. Figure 1 displays the estimated difference in topic proportions between female and male analysts for each corpus, along with the 95% confidence interval. The differences are also summarized in the last column of Table 6. Positive values indicate topics that are more prevalent among female analysts, while negative values indicate topics that are more prevalent among male analysts.

References:

- Chava S, Du W, Malakar B (2021) Do managers walk the talk on environmental and social issues? Working paper, Georgia Tech Scheller College of Business.
- Cormack GV, Grossman MR (2014) Evaluation of machine-learning protocols for technology-assisted review in electronic discovery. *Proc. 37th International ACM SIGIR Conf.* 153–162.
- Erel I, Stern LH, Tan C, Weisbach MS (2021) Selecting directors using machine learning. *Rev. Financial Stud.* 34(7):3226–3264.
- Raman N, Bang G, Nourbakhsh A (2020) Mapping ESG trends by distant supervision of neural language models. *Mach. Learn. Knowl. Extr.* 2(4):453–468.
- Roberts ME, Stewart BM, Tingley D (2019) STM: An R package for structural topic models. *J. Stat. Softw.* 91:1–40.
- Sundararajan M, Taly A, Yan Q (2017) Axiomatic attribution for deep networks. *Proc. 34th International Conf. on Machine Learning.* 3319–3328.

Appendix IA.B

Variable definitions

All continuous variables are winsorized at the 1st and 99th percentiles. All values are reported in 2021 constant dollars.

Variable	Definition
Firm-year level	
E&S score	The average of the environmental performance score and the social performance score in a given year. The data is from Refinitiv.
E score	The environmental performance score in a given year. The data is from Refinitiv.
S score	The social performance score in a given year. The data is from Refinitiv.
Carbon emissions	Natural logarithm of one plus the sum of annual Scope 1 and Scope 2 carbon emissions (metric tons of CO ₂) in a given year following Sautner et al. (2023). The data is from S&P Global Trucost.
Environment-related penalties	Natural logarithm of one plus the total dollar amount of penalty incurred due to a firm's environmental violations in a given year. The data is from Violation Tracker.
Environment-related cases	Natural logarithm of one plus the total number of environmental violations in a given year. The data is from Violation Tracker.
Workplace safety-related penalties	Natural logarithm of one plus the total dollar amount of penalty incurred due to a firm's workplace safety or health violations in a given year. The data is from Violation Tracker.
Workplace safety-related cases	Natural logarithm of one plus the total number of workplace safety or health violations in a given year. The data is from Violation Tracker.
N_female	The number of female analysts who cover a firm in a given year. We determine whether an analyst is a female or not based on hand-collected information.
N_male	The number of male analysts who cover a firm in a given year. We determine whether an analyst is a male or not based on hand-collected information.
N_analyst	The number of analysts covering a firm in a given year.
Total assets	Book value of total assets (in billions of dollars).
Firm size	Natural logarithm of total assets.
Tobin's Q	The sum of market value of equity and book value of debt divided by total assets.
ROA	Operating income before interest and taxes divided by total assets.
Leverage	Book value of debt divided by total assets.
SG&A	SG&A expenses divided by total assets.
Cash holdings	Cash and short-term investment divided by total assets.
Tangibility	Net property, plant, and equipment divided by total assets.
Board independence	The fraction of independent directors on a board.
CEO duality	An indicator variable that takes the value of one if a CEO is chairperson of the board in a firm, and zero otherwise.
Institutional ownership	The fraction of shares outstanding held by institutional investors, set to missing if the ratio is larger than 1.
Female analyst ratio	The ratio of the number of female analysts to the total number of analysts covering a firm in a given year.

Having female analyst	An indicator variable that takes the value of one if there is at least one female analyst who covers a firm in a given year, and zero otherwise. Having male analyst is defined analogously.
$\ln(1 + \text{SG\&A expenses})$	Natural logarithm of one plus the SG&A expenses (in millions of dollars).
SRI fund ownership	The fraction of shares outstanding held by socially responsible investment (SRI) funds. The data is from Heath et al. (2023).
Female analyst ideology	The average ideology score (liberal relative to conservative) of female analysts covering a firm in a given year. We use the ideology score of the state in which an analyst's office is located as a proxy for her ideology. Male analyst ideology is defined analogously. The state-year ideology score data is from Berry et al. (1998).
$\ln(1 + N_{\text{E\&S incidents}})$	Natural logarithm of one plus the number of negative E&S-related news items in a given year. The data is from RepRisk.
$\ln(1 + N_{\text{E incidents}})$	Natural logarithm of one plus the number of negative environmental-related news items in a given year. The data is from RepRisk.
$\ln(1 + N_{\text{S incidents}})$	Natural logarithm of one plus the number of negative social-related news items in a given year. The data is from RepRisk.
$\ln(1 + N_{\text{E\&S sentences}})$	Natural logarithm of one plus the average number of E&S-related sentences in reports written by analysts covering a firm in a given year.
$\ln(1 + N_{\text{E sentences}})$	Natural logarithm of one plus the average number of environmental-related sentences in reports written by analysts covering a firm in a given year.
$\ln(1 + N_{\text{S sentences}})$	Natural logarithm of one plus the average number of social-related sentences in reports written by analysts covering a firm in a given year.
$\ln(1 + N_{\text{E\&S questions}})$	Natural logarithm of one plus the average number of E&S-related questions raised by analysts during a firm's conference calls in a given year.
$\ln(1 + N_{\text{E questions}})$	Natural logarithm of one plus the average number of environmental-related questions raised by analysts during a firm's conference calls in a given year.
$\ln(1 + N_{\text{S questions}})$	Natural logarithm of one plus the average number of social-related questions raised by analysts during a firm's conference calls in a given year.
Analyst report level	
Having E&S sentences	An indicator variable that takes the value of one if there is at least one E&S-related sentence in an analyst report, and zero otherwise.
Having E sentences	An indicator variable that takes the value of one if there is at least one environmental-related sentence in an analyst report, and zero otherwise.
Having S sentences	An indicator variable that takes the value of one if there is at least one social-related sentence in an analyst report, and zero otherwise.
$\ln(1 + N_{\text{E\&S sentences}})$	Natural logarithm of one plus the number of E&S-related sentences in an analyst report.
$\ln(1 + N_{\text{E sentences}})$	Natural logarithm of one plus the number of environmental-related sentences in an analyst report.
$\ln(1 + N_{\text{S sentences}})$	Natural logarithm of one plus the number of social-related sentences in an analyst report.
Recommendation	Recommendation in a report in I/B/E/S issued by an analyst for a firm following Huang, Zang, and Zheng (2014). Stock recommendation in a report uses a five-tier rating system where 2 represents "strong buy," 1 represents "buy," 0 represents "hold," -1 represents "underperform," and -2 represents "sell."
Target price	Target price in a report in I/B/E/S issued by an analyst for a firm, divided by the stock price 50 days before the report date following Huang, Zang, and Zheng (2014).

Earnings forecast	Earnings forecast in a report in I/B/E/S issued by an analyst for a firm, divided by the stock price 50 days before the report date following Huang, Zang, and Zheng (2014).
Severity	The severity score in RepRisk, taking the value of one, two, or three, with three indicating the most severe (harsh) incident.
Novelty	The novelty score in RepRisk, taking the value of one or two, with two indicating that a firm is linked to a particular E&S issue for the first time.
Reach	The reach score in RepRisk, taking the value of one, two, or three, corresponding to the source of an E&S news event from local, national/regional, or international media, respectively.
Saliency	An indicator variable that takes the value of one if there is at least one salient E&S news event within the 90 days prior to an analyst's stock recommendation date, and zero otherwise. A salient news event is defined as an RepRisk E&S news event with a severity score over one, a novelty score over one, or a reach score over one.
Cognition	The fraction of words that reflects different ways people think or refer to their thinking. Cognition serves as the overarching dimension that includes the subcategories of all-or-none thinking, cognitive processes, and memory. The measure is constructed using the LIWC-22 software (Boyd et al. 2022).
All-or-none	The fraction of all-or-none or absolutist language (e.g., all, no, never, always). All-or-none thinking is formally known as "dichotomous thinking," refers to a style of thinking that tends to be over-generalized and more extreme. The measure is constructed using the LIWC-22 software (Boyd et al. 2022).
Cognitive processes	The fraction of words related to cognitive processes (e.g., but, not, if, or, know). Cognitive processes span six sub-categories: insight, causation, discrepancy, tentative, certitude, and differentiation. The measure is constructed using the LIWC-22 software (Boyd et al. 2022).
Insight	The fraction of words related to insight (e.g., know, how, think, feel). The measure is constructed using the LIWC-22 software (Boyd et al. 2022).
Causation	The fraction of causal words (e.g., how, because, make, why). The measure is constructed using the LIWC-22 software (Boyd et al. 2022).
Gunning-Fog index	The text readability score based on the average sentence length and the proportion of complex words (three or more syllables). A lower score indicates greater readability. A score of 12 corresponds to the reading level of a U.S. high school senior.
Flesch reading ease score	The text readability score based on the average sentence length and the average syllables per word. A higher score indicates greater readability. Scores range from 0 (very difficult) to 100 (very easy), with 60–70 capturing plain English.
Readability	The negative value of the estimated school grade level required to understand the text, based on eight readability measures, including the Gunning-Fog index and the Flesch reading ease score. A higher score indicates greater readability. The measure is constructed using textstat package: https://github.com/textstat/textstat .
CAR[-1,+1]	Cumulative three-day abnormal return (in percentage points) centered around the report date (day 0) based on a market model in which the market portfolio is the CRSP value-weighted market index.
Negative E&S tone	The negative value of the average tone of E&S-related sentences in an analyst report. Negative E (S) tone is defined analogously.
Negative non-E&S tone	The negative value of the average tone of non-E&S-related sentences in an analyst report.
Female	An indicator variable that takes the value of one if the lead analyst on an analyst report is a female, and zero otherwise.
Report length	Natural logarithm of the number of sentences in a report.
Recommendation revision	Recommendation in a report minus the last recommendation in I/B/E/S issued by the same analyst for the same firm following Huang, Zang, and Zheng (2014).

Target price revision	Target price in a report minus the last target price in I/B/E/S issued by the same analyst for the same firm, divided by the stock price 50 days before the report date following Huang, Zang, and Zheng (2014).
Earnings forecast revision	Earnings forecast in a report minus the last earnings forecast in I/B/E/S issued by the same analyst for the same firm, divided by the stock price 50 days before the report date following Huang, Zang, and Zheng (2014).
Prior CAR	Cumulative ten-day abnormal return (in percentage points) ending two trading days before the report date based on a market model in which the market portfolio is the CRSP value-weighted market index.

Analyst-call level

Having E&S questions	An indicator variable that takes the value of one if an analyst raises at least one E&S-related question during a firm's earnings conference call, and zero otherwise.
Having E questions	An indicator variable that takes the value of one if an analyst raises at least one environmental-related question during a firm's earnings conference call, and zero otherwise.
Having S questions	An indicator variable that takes the value of one if an analyst raises at least one social-related question during a firm's earnings conference call, and zero otherwise.
$\ln(1 + N_E\&S \text{ questions})$	Natural logarithm of one plus the number of E&S-related questions by an analyst during a firm's earnings conference call.
$\ln(1 + N_E \text{ questions})$	Natural logarithm of one plus the number of environmental-related questions by an analyst during a firm's earnings conference call.
$\ln(1 + N_S \text{ questions})$	Natural logarithm of one plus the number of social-related questions by an analyst during a firm's earnings conference call.
$N_questions$	The number of questions by an analyst during a firm's earnings conference call.
Female	An indicator variable that takes the value of one if an analyst who raises at least one question during a firm's earnings conference call is a female, and zero otherwise.

Analyst-year level

Education	An analyst's highest degree where 1 represents a bachelor's degree, 2 represents a master's degree, 3 represents a PhD, and 0 represents degrees below a bachelor's degree (such as a high school diploma).
CFA	An indicator variable that takes the value of one if an analyst is a Chartered Financial Analyst (CFA) charter-holder, and zero otherwise.
Star analyst	An indicator variable that takes the value of one if an analyst is accredited to All-Star status, and zero otherwise.
# firms followed	Number of firms for which an analyst makes at least one forecast in a given year.
# industries followed	Number of two-digit SIC industries for which an analyst makes at least one forecast in a given year.
General experience	Number of years since an analyst first appears in the I/B/E/S Detail History file following Bradley, Gokkaya, and Liu (2017).

Analyst-firm-year level

Forecast accuracy	The negative value of the average of the absolute forecast error made by an analyst in a given year demeaned by the average absolute forecast error of all analysts covering the same firm in the same year. The absolute forecast error is the absolute value of the difference between an analyst's annual EPS forecast and the actual EPS using the I/B/E/S Unadjusted Detail file.
Forecast frequency	Number of forecasts that an analyst makes of a firm in a given year.
Forecast horizon	The average of forecast horizons (in terms of the number of days between the forecast date and the earnings announcement date) that an analyst employs when making forecasts in a firm-year.

$\text{Ln}(1 + N_{\text{E\&S sentences}})$	Natural logarithm of one plus the average number of E&S-related sentences among the reports written by an analyst covering a firm in a given year.
$\text{Ln}(1 + N_{\text{E sentences}})$	Natural logarithm of one plus the average number of environmental-related sentences among the reports written by an analyst covering a firm in a given year.
$\text{Ln}(1 + N_{\text{S sentences}})$	Natural logarithm of one plus the average number of social-related sentences among the reports written by an analyst covering a firm in a given year.
$\text{Ln}(1 + N_{\text{E\&S questions}})$	Natural logarithm of one plus the average number of E&S-related questions raised by an analyst during a firm's conference calls in a given year.
$\text{Ln}(1 + N_{\text{E questions}})$	Natural logarithm of one plus the average number of environmental-related questions raised by an analyst during a firm's conference calls in a given year.
$\text{Ln}(1 + N_{\text{S questions}})$	Natural logarithm of one plus the average number of social-related questions raised by an analyst during a firm's conference calls in a given year.

Figure IA1
Active learning flowchart

This figure presents a flowchart of the active learning process.

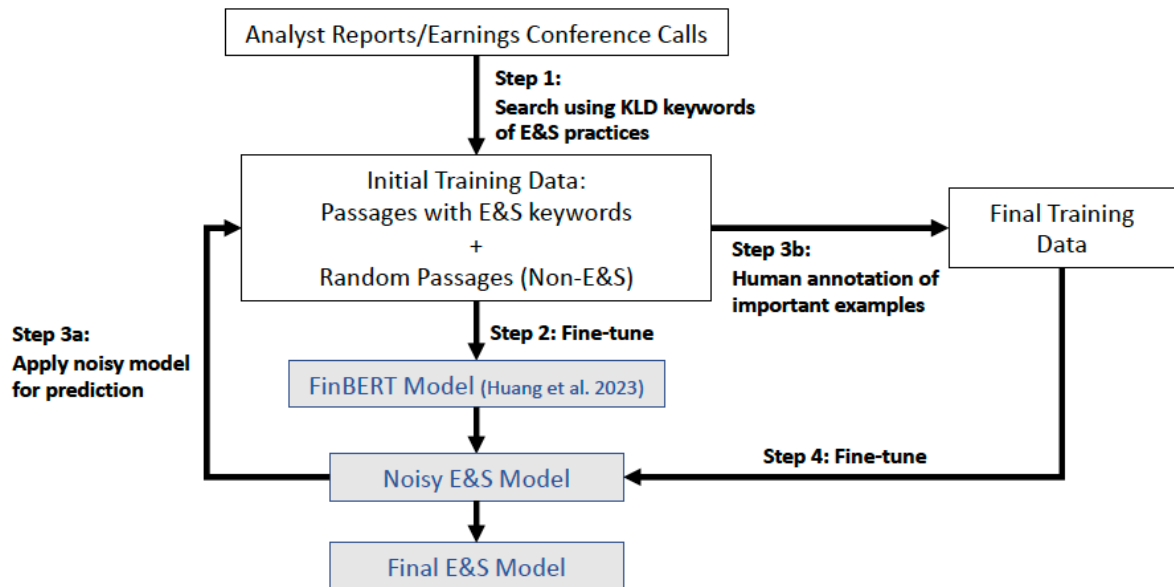


Figure IA2

Temporal trends in E&S-related discussions in analyst reports and during calls

This figure plots the temporal trend in E&S-related discussions in analyst reports and E&S-related questions during earnings conference calls. We obtain analyst reports over the period 2004–2020 from Thomson One’s Investtext database, and earnings call transcripts over the period 2007–2020 from Capital IQ. We capture discussions of E&S issues (E&S-related questions) using the fine-tuned FinBERT model to automatically classify E&S-related sentences (questions). We plot the yearly averages of the percentage of reports with E- or S-related sentences and the percentage of E- or S-related questions in calls.

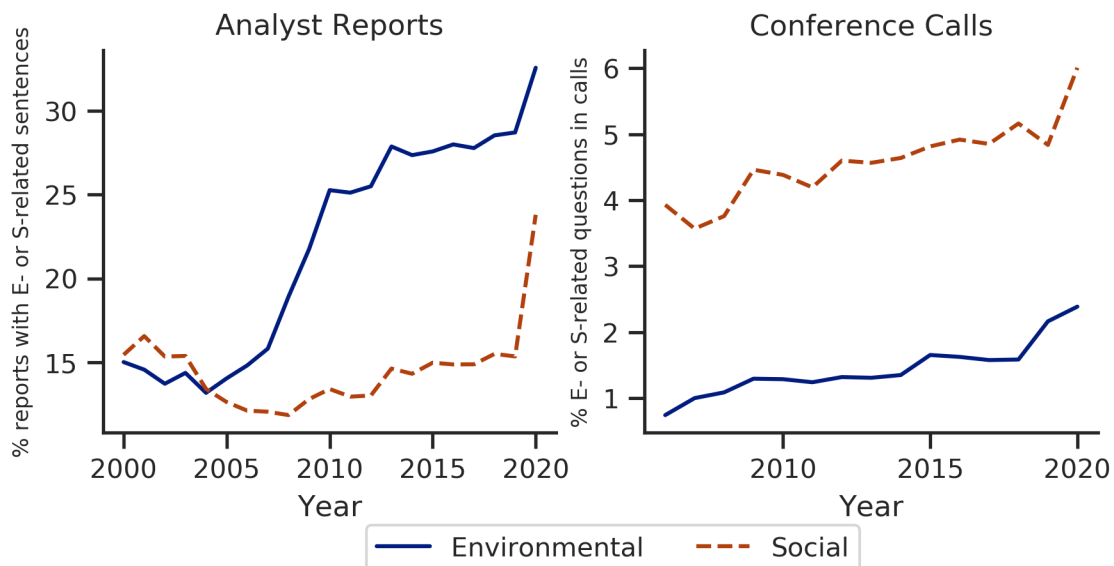


Figure IA3
Distribution of E&S-related discussions in analyst reports and during calls across Fama-French 12 industries

This figure plots the distribution of E&S-related discussions in analyst reports and E&S-related questions during earnings conference calls across Fama-French 12 industries. We obtain analyst reports over the period 2004–2020 from Thomson One’s Investext database, and earnings call transcripts over the period 2007–2020 from Capital IQ. We capture discussions of E&S issues (E&S-related questions) using the fine-tuned FinBERT model to automatically classify E&S-related sentences (questions). We plot the industry averages of the percentage of reports with E- or S-related sentences and the percentage of E- or S-related questions in calls.

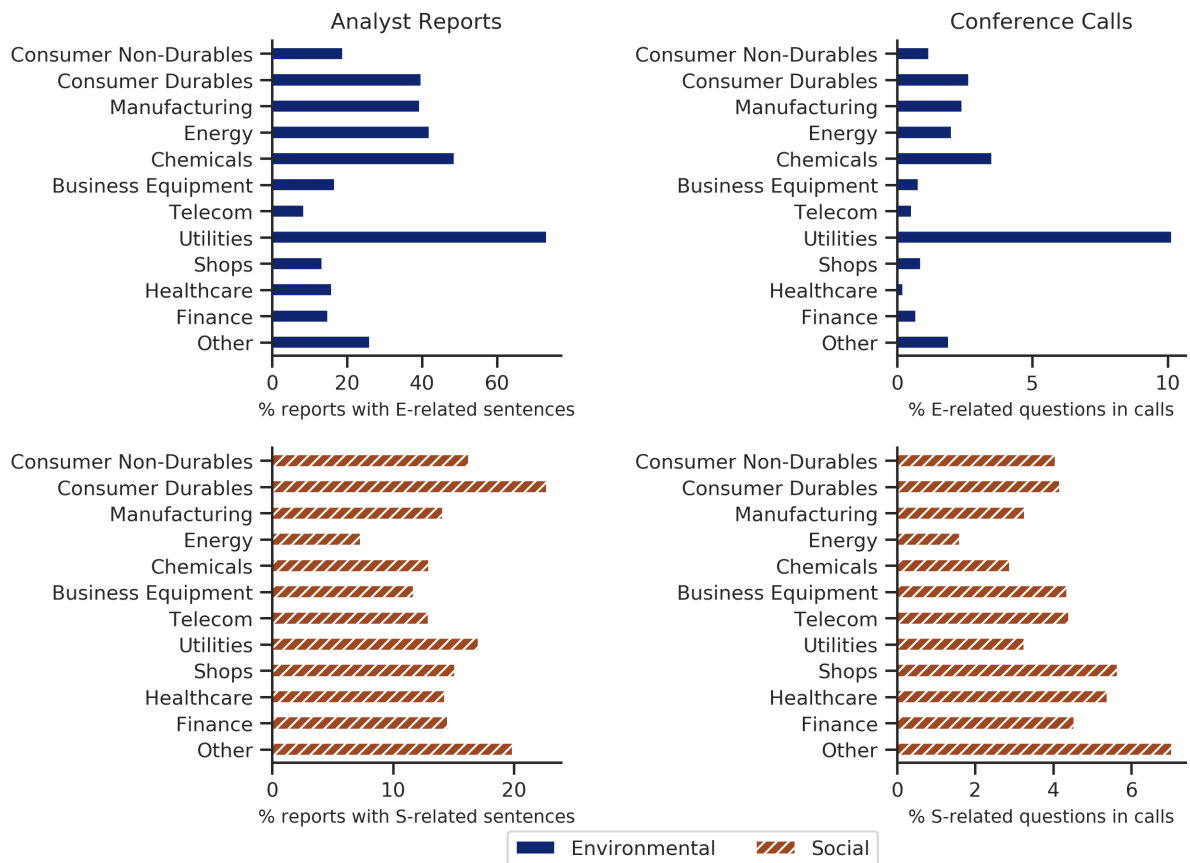
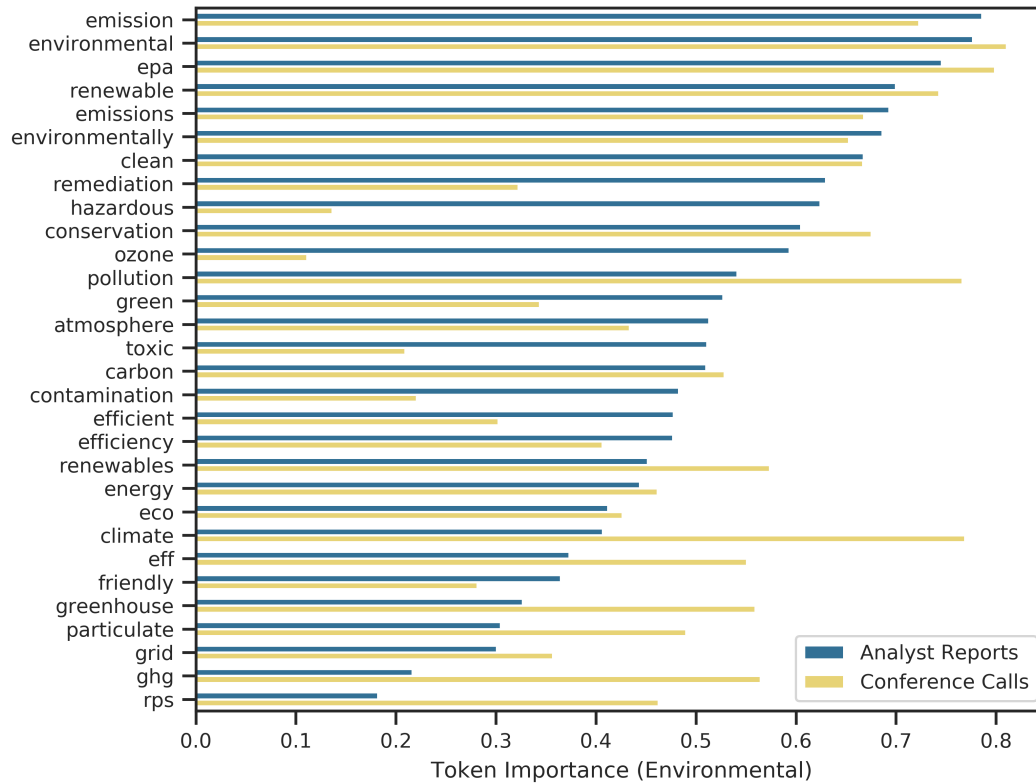


Figure IA4
Most important tokens in fine-tuned FinBERT models

This figure lists the most important tokens from our fine-tuned FinBERT models. We obtain analyst reports over the period 2004–2020 from Thomson One’s Investext database, and earnings call transcripts over the period 2007–2020 from Capital IQ. We capture discussions of E&S issues (E&S-related questions) using the fine-tuned FinBERT model to automatically classify E&S-related sentences (questions). We use the integrated gradients method (Sundararajan, Taly, and Yan 2017) to compute the token importance for each corpus. The integrated gradients method is a technique for explaining the prediction of a machine learning model by attributing the importance of each token to the model’s output. The FinBERT model, like many other transformer-based models, uses subword tokenization to break up words into smaller pieces (e.g., *resident* is tokenized to *re* and *##sident*).

Panel A: Most important tokens on environmental issues



Panel B: Most important tokens on social issues

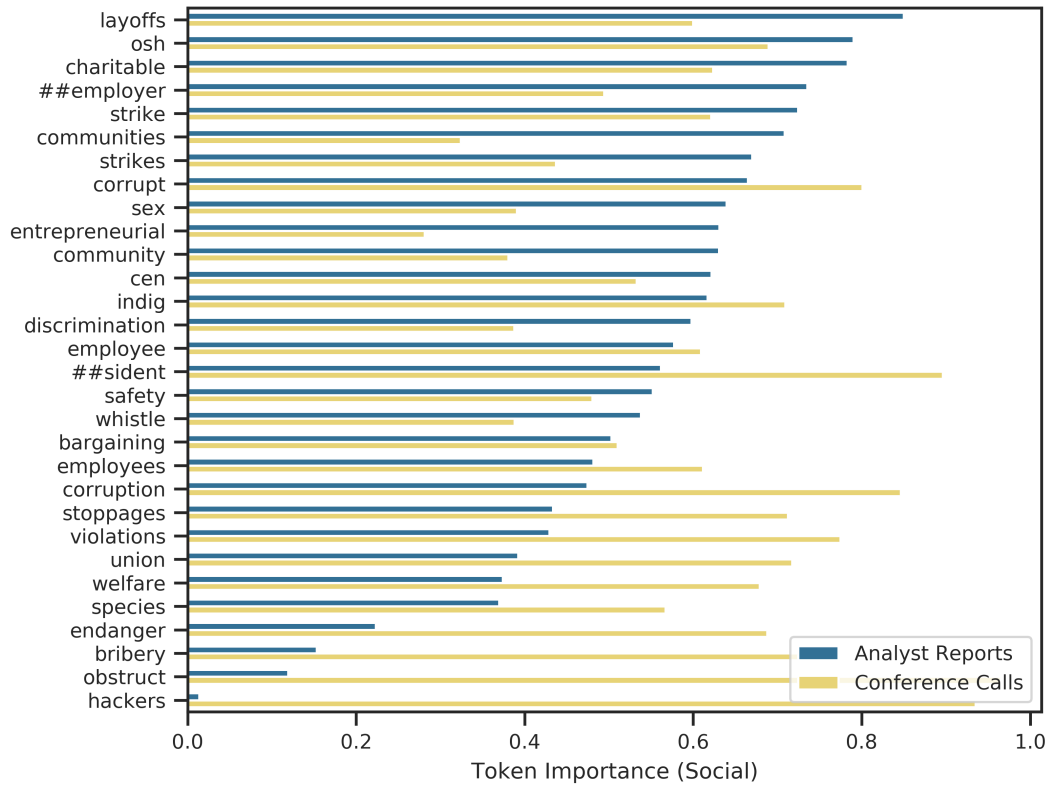


Table IA1
Analyst coverage over time and across Fama-French 12 industries

This table provides an overview of female analysts and analyst coverage over time and across Fama-French 12 industries over the period 2004–2020. The sample is the I/B/E/S-Capital IQ merged sample with hand-collected gender information, as described in Section 3.2. *N_female* is the number of unique female analysts present during a given time period or covering a specific industry, *N_analyst* is the number of unique analysts present during a given time period or covering a specific industry, and *Female analyst ratio* is the ratio of the number of female analysts to the total number of analysts.

Panel A: Temporal trend in female analysts and analyst coverage

	N_female	N_analyst	Female analyst ratio
2004-2007	764	4,893	0.14
2008-2011	704	4,919	0.13
2012-2015	607	4,605	0.12
2016-2020	634	4,435	0.13

Panel B: Industry distribution of female analysts and analyst coverage

	N_female	N_analyst	Female analyst ratio
Consumer Non-Durables	164	713	19%
Consumer Durables	70	564	11%
Manufacturing	168	1,394	11%
Energy	63	794	7%
Chemicals	65	401	14%
Business Equipment	230	2,214	9%
Telecom	40	434	8%
Utilities	16	90	15%
Shops	250	1,359	16%
Healthcare	162	949	15%
Finance	132	1,245	10%
Other	196	1,950	9%

Table IA2
Correlation matrices for the firm-year sample

This table presents the correlation matrices of the firm-year sample in Table 2. Definitions of the variables are provided in Appendix IA.B. Superscripts a, b, and c indicate significance at the 1, 5, and 10 percent levels, respectively.

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
1. E&S score	1																			
2. E score	0.957 ^a	1																		
3. S score	0.951 ^a	0.821 ^a	1																	
4. Carbon emissions	0.574 ^a	0.557 ^a	0.533 ^a	1																
5. Environmental-related penalties	0.176 ^a	0.175 ^a	0.156 ^a	0.463 ^a	1															
6. Environmental-related cases	0.187 ^a	0.186 ^a	0.167 ^a	0.187 ^a	0.917 ^a	1														
7. Workplace safety-related penalties	0.113 ^a	0.139 ^a	0.070 ^a	0.287 ^a	0.021 ^a	0.068 ^a	1													
8. Workplace safety-related cases	0.138 ^a	0.153 ^a	0.106 ^a	0.397 ^a	0.129 ^a	0.173 ^a	0.849 ^a	1												
9. N_female	0.270 ^a	0.238 ^a	0.278 ^a	0.281 ^a	-0.050 ^a	-0.055 ^a	-0.010	0.006	1											
10. N_male	0.316 ^a	0.284 ^a	0.321 ^a	0.411 ^a	0.090 ^a	0.100 ^a	-0.103 ^a	-0.015	0.461 ^a	1										
11. Firm size	0.436 ^a	0.390 ^a	0.443 ^a	0.447 ^a	0.048 ^a	0.083 ^a	-0.095 ^a	-0.01	0.524 ^a	0.981 ^a	1									
12. Tobin's Q	0.332 ^a	0.325 ^a	0.308 ^a	0.268 ^a	0.159 ^a	0.201 ^a	-0.098 ^a	-0.037 ^a	0.155 ^a	0.258 ^a	0.365 ^a	1								
13. ROA	0.559 ^a	0.514 ^a	0.555 ^a	0.528 ^a	-0.097 ^a	-0.115 ^a	-0.060 ^a	0.042 ^a	0.335 ^a	0.474 ^a	0.587 ^a	0.583 ^a	1							
14. Leverage	-0.072 ^a	-0.073 ^a	-0.065 ^a	-0.097 ^a	-0.030 ^b	-0.034 ^b	-0.005	-0.042 ^a	0.107 ^a	0.124 ^a	0.116 ^a	-0.126 ^a	-0.321 ^a	1						
15. SG&A	0.245 ^a	0.214 ^a	0.255 ^a	0.324 ^a	0.058 ^a	0.048 ^a	0.072 ^a	0.038 ^a	0.187 ^a	0.189 ^a	0.224 ^a	0.025 ^a	0.330 ^a	-0.052 ^a	1					
16. Cash holdings	0.094 ^a	0.108 ^a	0.070 ^a	0.236 ^a	-0.138 ^a	-0.168	0.067 ^a	0.063 ^a	0.036 ^a	0.016 ^b	-0.028 ^a	-0.022 ^a	0.151 ^a	-0.099 ^a	0.067 ^a	1				
17. Tangibility	-0.145 ^a	-0.134 ^a	-0.143 ^a	-0.131 ^a	-0.157 ^a	-0.160	0.027 ^c	-0.014	0.090 ^a	0.014 ^b	-0.051 ^a	-0.157 ^a	-0.457 ^a	0.551 ^a	-0.351 ^a	-0.162 ^a	1			
18. Board independence	-0.177 ^a	-0.160 ^a	-0.179 ^a	-0.228 ^a	0.308 ^a	0.327	-0.161 ^a	-0.155 ^a	-0.024 ^a	0.026 ^a	-0.001	-0.080 ^a	-0.363 ^a	0.506 ^a	-0.444 ^a	-0.251 ^a	0.591 ^a	1		
19. CEO duality	0.077 ^a	0.080 ^a	0.066 ^a	0.453 ^a	-0.057 ^a	-0.064	0.156 ^a	0.228 ^a	-0.025 ^a	-0.023 ^a	0.169 ^a	-0.076 ^a	0.008	-0.073 ^a	0.077 ^a	0.129 ^a	-0.136 ^a	-0.135 ^a	1	
20. Institutional ownership	-0.205 ^a	-0.186 ^a	-0.206 ^a	-0.348 ^a	0.056 ^a	0.058 ^a	0.072 ^a	0.022	-0.308 ^a	-0.433 ^a	-0.460 ^a	-0.091 ^a	-0.278 ^a	-0.154 ^a	-0.202 ^a	0.021 ^a	-0.049 ^a	-0.029 ^a	-0.005	

Table IA3
Queries of E&S issues

This table lists queries developed from the KLD User Guide (2010, the last version before RiskMetrics KLD was acquired by MSCI) to capture E&S issues. We apply these queries to analyst reports and earnings call transcripts to identify analysts' discussion of E&S issues. For example, a query using "communit* & involvement" means that only a passage includes both a word starting with "communit" (such as "community" and "communities") and the "involvement" is considered to be relevant discussions of E&S issues regarding community.

E&S practices	Queries
Community	communit* && charitable, "community involvement", "community reinvestment act" NOT "community reinvestment act funds", "disadvantaged people", "disadvantaged groups", communit* && donation, "underserved communities", "indigenous people", "local community" NOT banks NOT bank NOT "local community papers" NOT local , community pharmacy" NOT "local community hospital" NOT "merger" NOT "local community storefronts", communit* && ngo, communit* && "non-profit organizations", "socially responsible investing", "local communities" && support, "local communities" && sponsor, communit* && volunteer, communit* && youth && training
Diversity	diversity && bisexual, "black owned" NOT "national association of black owned broadcasters" NOT "orange is the new black", diversity && csr NOT market, diversity && esg, "female ceo", "female executives", diversity && gay, diversity && gender, glbt*, lgbt*, "ethnic diversity", diversity && inclusive NOT geographic, "sexual orientation", transgendered, "veteran owned", "female owned", diversity && inclusion, "work-life balance", diversity && workforce
Employee relations	"defined benefit" && underfunded, "health and safety" && employees, "no-layoff", osha NOT joe NOT "j. osha" NOT "joseph osha" NOT stericycle NOT "steri safe", employee* && "profit sharing", strike* && employee* NOT "strike me" NOT "strikes me" NOT "strikes you" NOT "strike , rice", "wrongful termination" NOT "please call us", "union relations" NOT "the company specific risks to our investment thesis include", health && safety && employee*, "significant layoffs", "significant workforce reduction", "major layoff"
Environment	environment* && "circular economy", "clean air act", "clean energy" NOT "clean energy ventures" NOT "allete clean energy" NOT "lu'an clean energy company" NOT "china sunergy co" NOT "clean energy group" NOT "s&p global clean energy index" NOT "okeechobee clean energy center" NOT "con edison clean energy businesses" NOT "clean energy fuels" NOT "global clean energy holdings", "clean water act", "climate change", "eco-friendly", "ecological restoration", "emission reduction", "energy efficient", "energy efficiency", environmental* && lawsuits, "environmental protection agency", epa && regulation*, environmental* && remediation, "environmental sustainability", "global reporting initiative", "green building", "green transport", "greenhouse gas", "gri guidelines", liabilities && hazardous, "iso 14001", "iso 50001", environment* && carbon && emission*, environment* && co2 && emission*, "carbon footprint", environment* && ozone , environment* && pollut* NOT "competitive environment", environment* && "renewable energy", environmental* && sourcing, "safe drinking water act", environmentally && sustainable, environment* && toxic
Human rights	"child labor", "forced labor", "free association", "free speech" NOT grounds, censorship, "human rights", "human trafficking", "labor rights", "prison labor"
Product	antitrust && violation, "product safety" && issues, cpsia, "food safety" && violations

Table IA4
Important training examples in active learning

This table lists examples from the two corpora (reports and calls) that are identified as important for human annotation in active learning, using Continuous Active Learning (CAL) and Simple Active Learning (SAL) protocols. CAL focuses on examples the model is most certain about, while SAL focuses on examples the model is uncertain about. For CAL-selected examples, the *Noisy E&S model's* predicted class probabilities, along with the human labels, are provided. For SAL-selected examples, the *Noisy E&S model's* predicted labels and the corresponding human labels are provided.

Panel A: Important examples from analyst reports

SAL (High uncertainty examples)
<p>Example 1: Unlike other equity analysts and market commentators, our focus is on leveraging overriding themes -an approach called Thematic Investing that looks to identify emerging economic, political, regulatory and social structural changes around the globe, and then seeks to determine which companies will be impacted by it -both those that stand to benefit from the tidal wave, and those that will be drowned out by it.</p> <p><i>Noisy E&S Model Class Probability:</i> P(E) = 0.53, P(S) = 0.4, P(N) = 0.4 <i>Human Label:</i> [N]</p>
<p>Example 2: However, in May, the Fourth Circuit vacated ACP's "incidental take" permit issued by the US Fish and Wildlife Service, meant to help protect endangered species at water crossings along the route.</p> <p><i>Noisy E&S Model Class Probability:</i> P(E) = 0.81, P(S) = 0.17, P(N) = 0.02 <i>Human Label:</i> [E]</p>
<p>Example 3: Below we discuss the business segment results: Revenues in its Consumer and Commercial Services division increased 15% to \$1 billion, reflecting strong revenue contribution from Terminix, American Home Shield (AHS), ARS/Rescue Rooter (ARS), ServiceMaster Clean and Merry Maids.</p> <p><i>Noisy E&S Model Class Probability:</i> P(E) = 0.29, P(S) = 0.56, P(N) = 0.15 <i>Human Label:</i> [N]</p>
CAL (High predicted probability examples)
<p>Example 1: Importantly, US Ecology management reiterated its commitment to the current dividend (\$0.72/share annually).</p> <p><i>Noisy E&S Model Label:</i> [E] <i>Human Label:</i> [N]</p>
<p>Example 2: Citi Holdings net income came in lower than expected with a net loss of \$802 million, driven by larger than expected losses in the Special Asset Pool segment ("SAP") as there were no recorded securities gains and lower loan balances in Local Community Lending ("LCL") which led to lower net interest revenue.</p> <p><i>Noisy E&S Model Label:</i> [S] <i>Human Label:</i> [N]</p>
<p>Example 3: When asked about the recent performance of its Southern California division, Safeways CEO stated only that it has been consistent with other post-strike periods, though cost savings have been realized a bit earlier than expected due to a larger number of employees not returning to work upon conclusion of the strike.</p>

Noisy E&S Model Label: [S]
Human Label: [S]

Panel B: Important examples from earnings conference calls

SAL (High uncertainty examples)

Example 1: In the United States, a kind of a two-part question. With a modest reflation, kind of a grinding economic expansion, some tax revenues to the municipal levels going up, what are you seeing in demand from U.S. fire departments be either volunteer rural or the full-service cities, with the demand from fire suits, and then also too after 9/11. I think the Bush administration set up a number of depots for hazmat suits, first responders. You've said, Chris, over time that those, the glue in those different suits tends to erode. How do you see from kind of a national security or a homeland security, any follow up to replace some of that aging 9/11 suit stuff?

Noisy E&S Model Class Probability: $P(E) = 0.19$, $P(S) = 0.57$, $P(N) = 0.24$
Human Label: [S]

Example 2: If somebody answered this, I can just go back and read the transcript later. But I'm just wondering if you talked at all about sort of your expectation for soy crush margins, as the Argentine farmer begins to release all these pent-up beans. Are you anticipating any sort of degradation in the crush outlook?

Noisy E&S Model Class Probability: $P(E) = 0.42$, $P(S) = 0.14$, $P(N) = 0.44$
Human Label: [N]

Example 3: Robert, I mean, and Emanuele, clearly, this is a bold move by the company. I mean, it looks well timed. I mean today, BP just came out and announced they're going to cut oil and gas CapEx by 40% and shift that into renewables. As you thought about this pivot into offshore wind, I guess, a little -- could you give us a little bit of color around when you started thinking about it? I'm assuming you had a lot of conversations with a lot of potential customers. Kind of just could you give us some of the genesis around the decision to make this move?

Noisy E&S Model Class Probability: $P(E) = 0.46$, $P(S) = 0.0$, $P(N) = 0.54$
Human Label: [E]

CAL (High predicted probability examples)

Example 1: Okay. Last housekeeping on the mediation with the Teamsters in Auburn. Has the union also agreed to mediation?

Noisy E&S Model Label: [S]
Human Label: [S]

Example 2: Okay. And also, going back to the market, there has been a great deal of publicity regarding cutbacks in spending by state and local governments due to lower tax revenues. Have you seen an impact upon your monthly run rate of orders this year so far from domestic law enforcement agencies because of it?

Noisy E&S Model Label: [S]
Human Label: [N]

Example 3: So the increase in nonperforming loans, it's interesting that, that is not at all energy-related or at least a majority of it is not energy-related and that is your nonenergy C&I. Can you just expand a little bit on what specifically drove that increase in 1Q?

Noisy E&S Model Label: [E]
Human Label: [N]

Table IA5
Examples of E&S-related sentences in analyst reports

This table lists examples of E&S-related sentences in analyst reports used in Table 4. At the report level, we capture discussions of E&S issues using the fine-tuned FinBERT model to automatically classify E&S-related sentences. Panel A lists examples of environmental-related sentences. Panel B lists examples of social-related sentences.

Panel A: Environmental-related sentences

Example 1: This report was written by Hayley Beth Wolff (Female) from Rochdale Securities LLC for Polaris Inc. released on 7/27/2009.

More significantly, we believe that eco-friendly engine may satisfy the growing demand from many government agencies such as the US Forest Service and US military, looking for more environmentally friendly solutions to gas-powered vehicles.

Example 2: This report was written by David Begleiter (Male) from Deutsche Bank for Eastman Chemical Company released on 3/7/2011.

Going forward Eastman has established a 10-year environmental target for 25% reduction in energy intensity, 20% reduction in greenhouse gas intensity, and 20% NO₂ & 40% SO₂ reductions.

Example 3: This report was written by Vishal Shah (Male) from Deutsche Bank for First Solar released on 9/15/2011.

However, we anticipate a paradigm shift going forward, with clean electricity generation, particularly solar, gaining traction in several end-markets, supported by favorable government policies and improving cost structures.

Example 4: This report was written by Ann Kohler (Female) from Imperial Capital for Valero Energy Corp released on 1/31/2013.

Although there are limited government mandated regulatory capital requirement for refiners in the near term, the federal government continues to seek to reduce refinery emissions, including greenhouse gases through increased regulations by the Environmental Protection Agency (EPA).

Example 5: This report was written by Ryan Brinkman (Male) from J.P. Morgan for Tenneco released on 8/22/2017.

Tenneco management stated in our meetings that looking just at the US Tier 3 regulation alone for light vehicles, the Environmental Protection Agency (EPA) has referenced a +\$72 content per vehicle cost to manufacturers in order to comply with these stricter regulations (with much of this representing revenue opportunity for Tenneco -some of the incremental content will be on the engine side, but much of it will be on the tailpipe end).

Panel B: Social-related sentences

Example 1: This report was written by Stacey Widlitz (Female) from Fulcrum Global Partners for Tiffany & CO. released on 7/1/2005.

To be sure, some human rights organizations have made accusations that De Beers Group mining resulted in the relocation of bushmen in Botswana.

Example 2: This report was written by Ann Duignan (Female) from J.P. Morgan for Eaton Corp released on 9/22/2011.

What was interesting about this facility was the strong sense of community -employees come in early to participate in team activities, some volunteer at the on-campus school for local, underprivileged children, and many participate in on-campus sports activities to represent "team Eaton" vs. other local companies.

Example 3: This report was written by Devina Mehra (Female) from First Global Stockbroking for Philip Morris International released on 12/23/2012.

With tobacco, the main constituent of cigarettes, being the single greatest cause of preventable death globally and highly addictive, PMI's operations (as well as of its competitors) are highly controversial and are increasingly the subject of litigation and restrictive legislation from governments concerned about the health impacts of tobacco products.

Example 4: This report was written by Joseph Bonner (Male) from Argus Research for Alphabet Inc. released on 12/9/2019.

Messrs. Page and Brin are leaving executive management just as Alphabet faces a daunting range of challenges: multiple antitrust investigations, both in the U.S. and abroad; intense competition for internet advertising from Facebook and Amazon; issues surrounding user privacy; YouTube's potential liabilities for endangering the welfare of children; and an increasingly restive workforce.

Example 5: This report was written by Jonathan Ho (Male) from William Blair & Company for Axon Enterprise released on 11/5/2020.

Regarding gender and racial/ethnic diversity, the company's board of directors is one-third female.

Table IA6
Examples of E&S-related questions during calls

This table lists examples of E&S-related questions during earnings conference calls used in Table 5. At the analyst-call level, we capture E&S-related questions using the fine-tuned FinBERT model to automatically classify E&S-related questions. Panel A lists examples of environmental-related questions. Panel B lists examples of social-related questions.

Panel A: Environmental-related questions

Example 1: The question was asked by Marc de Croisset (Male) from FBR Capital Markets & Co. on the FQ2 2011 earnings conference call of The Southern Company held on 07/27/2011.

If I may, I'd love to ask a quick question on your thoughts on the Cross State Air Pollution Rule. One of the arguments that I think the EPA has made is that SO2 compliance could be achieved by having utilities use existing scrubbers more effectively, and as a result, that would be one of the means to reduce -- to achieve SO2 compliance. And I'd be very interested in your reaction to this argument. And have you seen any indication in the industry that -- or in your region, that scrubbers, over the last several years, may not have been utilized as often or as effectively as they could be?

Example 2: The question was asked by Ryan J. Brinkman (Male) from KeyBanc Capital Markets Inc. on the FQ3 2015 earnings conference call of Tesla, Inc. held on 03/11/2015.

Just maybe going back to the Dieselgate issue again, but from a bigger picture perspective. I'm curious what impact you see to the electric vehicle market from these revelations at VW. Could it increase the demand for electric vehicles to your benefit? Does it maybe make nonelectric vehicles more expensive to produce to truly comply with the emission regulations? Does that help the Model 3 be more cost- competitive? I'm just curious what impact you see overall to the industry, and then to Tesla specifically.

Example 3: The question was asked by Noelle Christine Dilts (Female) from Stifel, Nicolaus & Company, Incorporated on the FQ3 2018 earnings conference call of Myr Group Inc. held on 11/01/2018.

Okay. And then in terms of your commentary on renewable energy and some of those projects seeing support. How are you thinking about the -- kind of the knock on the factor, what that does to transmission project demand? I mean, do you see that as driving some of the larger kind of highway projects that would move renewable energy from point A to point B? Or are you thinking about that as driving kind of more of the small to medium-sized intertie type of work? Just curious kind of how you're thinking about that.

Example 4: The question was asked by Angie Storzynski (Female) from Macquarie Research on the FQ1 2019 earnings conference call of Entergy Corporation held on 05/01/2019.

I'm sorry. I was just wondering about your regulated renewable power CapEx. You mentioned that some of your jurisdictions might consider more renewable spending going forward once renewables become more economic, but given that there is some sort of some of the tax subsidies, would you -- wouldn't you consider actually potentially accelerating this CapEx?

Example 5: The question was asked by Theresa Chen (Female) from Barclays Bank PLC. On the FQ3 2020 earnings conference call of Valero Energy Corporation held on 10/22/2020.

I guess a follow-up question on the renewable diesel front. Clearly, the energy transition is a big theme along with ESG investing and happy to see the additional disclosures consistent with the SASB framework. Can you talk about how you view your renewable diesel position as far as the defensibility of your projected returns? How many of these projects that have been recently announced are you factoring in as ones that could come to fruition? And also on the LCFS prices as well, do you see any risk there?

Panel B: Social-related questions

Example 1: The question was asked by Kelly A. Flynn (Female) from Credit Suisse AG on the FQ4 2012 earnings conference call of Robert Half International Inc. held on 01/29/2013.

Okay, great. So I still one -- a last one on the health care, ObamaCare issue. What portion of the temporary employees do you currently pay benefits to?

Example 2: The question was asked by Irene Oiyin Haas (Female) from Wunderlich Securities Inc. on the FQ4 2014 earnings conference call of PDC Energy, Inc. held on 02/19/2015.

Okay. In terms of community relations, how are things shaping out these days?

Example 3: The question was asked by Jeffrey Ted Kessler (Male) from Imperial Capital, LLC on the FQ4 2018 earnings conference call of ShotSpotter, Inc. held on 02/19/2019.

I recently was at a safe city, secure city's conference. And one of the things that they talked about in terms of funding various programs was a catalyst, something to kind of tie the various services together, around which the public/private partnerships could actually agree on funding. And my question to you is, do you think that you -- in your relationships with companies like Verizon, are you able to get that mind share in which these -- well, let's just say these groups will be able to get mind share around you, too? Taking on, essentially, you being the brand name that they use to go out to the community and try to get funds for, not just ShotSpotter but you serving as a catalyst for other types of safety and public safety measures. In other words, that builds up your value proposition as well.

Example 4: The question was asked by Laurie Katherine Havener Hunsicker (Female) from Stifel, Nicolaus & Company on the FQ1 2011 earnings conference call of Washington Trust Bancorp, Inc. held on 04/25/2011.

Okay. And then we're still going to see the charitable foundation this year? Charitable foundation contribution, is that likely?

Example 5: The question was asked by Richard Tobie Safran (Male) from The Buckingham Research Group Incorporated on the FQ1 2012 earnings conference call of Lockheed Martin Corporation held on 04/26/2012.

Yes, thanks, Bruce. Bob, at the risk of this being a somewhat sensitive topic, I want to know if I can get a comment from you on negotiations with the Machinists Union of Fort Worth. I want to know if you could talk about the impact of a protracted disagreement. Is this a situation that's serious enough where, for example, you think you have the potential to lay off personnel? And I'm only asking this because the news reports I'm looking at seem to indicate that the Union is making statements like they're ready for a long strike, that kind of thing.

Table IA7
Robustness checks: Using alternative ESG data

This table conducts robustness checks on our main findings in Table 2 Panel A using three alternative ESG data sets. Columns (1)-(3) present the results using the E&S scores from Thomson Reuters' ASSET4 over the period 2005–2018 when it was replaced by Refinitiv's ESG database used in our main analysis. Columns (4)-(6) present the results using the E&S scores from MSCI's KLD Stats over the period 2005–2018 when it was discontinued thereafter. Columns (7)-(9) present the results using the E&S scores from Morningstar's Sustainalytics over the period 2009–2018 when the legacy Sustainalytics database, which measures ESG preparedness and performance, was discontinued in 2019. Other control variables are the same as those in Table 2 and are omitted for brevity. Industry fixed effects are based on Fama-French 48-industry classifications. Definitions of the variables are provided in Appendix IA.B. Standard errors (in parentheses) are clustered at the firm level. ***, **, * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Variable	ASSET4			KLD			Sustainalytics		
	E&S score (1)	E score (2)	S score (3)	E&S score (4)	E score (5)	S score (6)	E&S score (7)	E score (8)	S score (9)
N_female	1.453*** (0.474)	0.736*** (0.278)	0.714*** (0.249)	0.006*** (0.001)	0.004*** (0.001)	0.006*** (0.001)	0.881*** (0.302)	0.500*** (0.180)	0.376** (0.162)
N_male	-0.129 (0.106)	-0.072 (0.063)	-0.059 (0.051)	0.001** (0.000)	0.000 (0.000)	0.001*** (0.000)	-0.200*** (0.065)	-0.095** (0.041)	-0.101*** (0.032)
Other controls	YES	YES	YES	YES	YES	YES	YES	YES	YES
Industry × Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Adjusted R ²	0.446	0.392	0.386	0.332	0.336	0.328	0.326	0.332	0.301
No. of observations	13,939	13,939	13,939	10,520	10,520	10,520	8,042	8,042	8,042

Table IA8
Analyst coverage and corporate E&S performance

This table examines the relation between analyst coverage ($N_{analyst}$) and firms' E&S performance. Panel A examines the relation between analyst coverage and firms' E&S performance: $E\&S$ score, E score, and S score. Panel B examines the relation between analyst coverage and real E&S outcomes: *Carbon emissions*, *Environment-related penalties*, *Environment-related cases*, *Workplace safety-related penalties*, and *Workplace safety-related cases*. Industry fixed effects are based on Fama-French 48-industry classifications. Definitions of the variables are provided in Appendix IA.B. Standard errors (in parentheses) are clustered at the firm level. ***, **, * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Analyst coverage and corporate E&S performance

Variable	E&S score (1)	E score (2)	S score (3)
$N_{analyst}/100$	-0.013 (0.058)	-0.005 (0.064)	-0.021 (0.058)
Firm size	0.125*** (0.003)	0.128*** (0.003)	0.122*** (0.003)
Tobin's Q	0.011*** (0.002)	0.011*** (0.002)	0.012*** (0.002)
ROA	0.059*** (0.018)	0.021 (0.019)	0.098*** (0.019)
Leverage	-0.073*** (0.017)	-0.070*** (0.019)	-0.076*** (0.018)
SG&A	0.130*** (0.019)	0.130*** (0.020)	0.131*** (0.019)
Cash holdings	-0.062*** (0.012)	-0.050*** (0.013)	-0.075*** (0.012)
Tangibility	-0.010 (0.015)	0.007 (0.016)	-0.026 (0.016)
Board independence	0.013 (0.032)	-0.003 (0.036)	0.028 (0.032)
CEO duality	-0.013** (0.006)	-0.012* (0.007)	-0.015** (0.006)
Institutional ownership	-0.031*** (0.012)	-0.048*** (0.014)	-0.015 (0.012)
Industry \times Year FE	YES	YES	YES
Adjusted R ²	0.558	0.521	0.513
No. of observations	20,402	20,402	20,402

Panel B: Analyst coverage and real E&S outcomes

Variable	Carbon emissions (1)	Environment-related penalties (2)	Environment-related cases (3)	Workplace safety-related penalties (4)	Workplace safety-related cases (5)
N_analyst	-0.007* (0.004)	-0.029* (0.016)	-0.007*** (0.002)	-0.059*** (0.016)	-0.008*** (0.002)
Firm size	0.984*** (0.023)	1.129*** (0.119)	0.147*** (0.019)	0.702*** (0.110)	0.142*** (0.020)
Tobin's Q	-0.034** (0.014)	0.219* (0.118)	0.022* (0.012)	-0.107 (0.111)	-0.017 (0.016)
ROA	1.669*** (0.159)	-2.293* (1.231)	-0.054 (0.143)	3.491*** (1.164)	0.338* (0.185)
Leverage	0.228* (0.129)	0.030 (0.591)	-0.040 (0.067)	-1.413** (0.582)	-0.156* (0.094)
SG&A	0.970*** (0.134)	0.269 (0.860)	0.069 (0.084)	0.956 (0.796)	0.234* (0.122)
Cash holdings	-0.502*** (0.097)	-1.539** (0.739)	-0.117 (0.077)	-2.569*** (0.748)	-0.299*** (0.109)
Tangibility	0.990*** (0.137)	3.216*** (0.568)	0.306*** (0.066)	0.835 (0.557)	0.211** (0.090)
Board independence	0.365* (0.187)	3.164*** (1.021)	0.344*** (0.120)	0.688 (0.997)	0.058 (0.163)
CEO duality	0.045 (0.036)	-0.148 (0.190)	-0.016 (0.023)	0.471*** (0.182)	0.049 (0.031)
Institutional ownership	-0.011 (0.087)	-0.702 (0.465)	-0.084 (0.058)	-0.401 (0.438)	-0.091 (0.084)
Industry × Year FE	YES	YES	YES	YES	YES
Adjusted R ²	0.873	0.328	0.374	0.270	0.333
No. of observations	14,651	5,167	5,167	5,167	5,167

Table IA9
Robustness checks

This table conducts robustness checks on our main findings in Table 2 Panel A. Panel A includes firm and year fixed effects (instead of industry \times year fixed effects). Panel B examines the relation between firms' female analyst ratio and firms' E&S performance. *Female analyst ratio* is the ratio of the number of female analysts to the total number of analysts covering a firm in a given year. Panel C examines the relation between having female analyst coverage and firms' E&S performance. *Having female analyst* is an indicator variable that takes the value of one if there is at least one female analyst who covers a firm in a given year, and zero otherwise. *Having male analyst* is defined analogously. Panel D examines the relation between female analyst coverage and firms' E&S performance separating by their brokerage size. *N_female_Top10* is the number of female analysts, from one of the top 10 brokers, who cover a firm in a given year. We determine whether a broker is one of the top 10 brokers based on its size. *N_female_non-Top10* is the number of female analysts, not from one of the top 10 brokers, who cover a firm in a given year. Panel E controls for SRI fund ownership. *SRI fund ownership* is the fraction of shares outstanding held by socially responsible investment (SRI) funds. The SRI fund ownership data is from Heath et al. (2023). The sample consists of 15,027 firm-year observations (representing 3,174 firms) with data on SRI fund ownership over the period 2010–2020. Panel F controls for analyst political ideology. *Female analyst ideology* is the average ideology score (liberal relative to conservative) of female analysts covering a firm in a given year. We use the ideology score of the state in which an analyst's office is located as a proxy for her ideology. The state-year ideology scores are from Berry et al. (1998). *Male analyst ideology* is defined analogously. The sample consists of 6,154 firm-year observations (representing 1,333 firms) with data on the ideology score over the period 2004–2016. Other control variables are the same as those in Table 2 Panel A and are omitted for brevity. Industry fixed effects are based on Fama-French 48-industry classifications. Definitions of the variables are provided in Appendix IA.B. Standard errors (in parentheses) are clustered at the firm level. ***, **, * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Including firm and year fixed effects

Variable	E&S score (1)	E score (2)	S score (3)
N_female	0.007*** (0.002)	0.008** (0.002)	0.006*** (0.002)
N_male	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Other controls	YES	YES	YES
Firm FE	YES	YES	YES
Year FE	YES	YES	YES
Adjusted R ²	0.833	0.806	0.794
No. of observations	19,990	19,990	19,990

Panel B: Female analyst ratio and corporate E&S performance

Variable	E&S score (1)	E score (2)	S score (3)
Female analyst ratio	0.053** (0.025)	0.050* (0.028)	0.055** (0.026)
Other controls	YES	YES	YES
Industry \times Year FE	YES	YES	YES
Adjusted R ²	0.563	0.525	0.515
No. of observations	16,717	16,717	16,717

Panel C: Having female analyst coverage and corporate E&S performance

Variable	E&S score (1)	E score (2)	S score (3)
Having female analyst	0.017*** (0.006)	0.019*** (0.007)	0.015** (0.007)
Having male analyst	-0.062*** (0.011)	-0.062*** (0.012)	-0.062*** (0.011)
Other controls	YES	YES	YES
Industry × Year FE	YES	YES	YES
Adjusted R ²	0.562	0.524	0.517
No. of observations	20,402	20,402	20,402

Panel D: Separating female analysts by their brokerage size

Variable	E&S score (1)	E score (2)	S score (3)
N_female_Top10	0.015*** (0.005)	0.013** (0.005)	0.017*** (0.005)
N_female_non-Top10	0.008** (0.004)	0.009** (0.004)	0.007* (0.004)
t-test [N_female_Top10 = N_female_non-Top10] p-value	0.243	0.541	0.115
Other controls	YES	YES	YES
Constant	YES	YES	YES
Industry × Year FE	YES	YES	YES
Adjusted R ²	0.560	0.523	0.516
No. of observations	20,402	20,402	20,402

Panel E: Controlling for SRI fund ownership

Variable	E&S score (1)	E score (2)	S score (3)
N_female	0.014*** (0.003)	0.013*** (0.003)	0.015*** (0.003)
N_male	-0.001 (0.001)	-0.000 (0.001)	-0.001 (0.001)
SRI fund ownership	1.543*** (0.501)	1.504*** (0.563)	1.579*** (0.529)
Other controls	YES	YES	YES
Industry × Year FE	YES	YES	YES
Adjusted R ²	0.586	0.546	0.545
No. of observations	15,027	15,027	15,027

Panel F: Controlling for analyst political ideology

Variable	E&S score (1)	E score (2)	S score (3)
N_female	0.009** (0.005)	0.005 (0.005)	0.014*** (0.005)
N_male	0.001 (0.001)	0.002 (0.001)	0.001 (0.001)
Female analyst ideology	0.076 (0.064)	0.064 (0.074)	0.087 (0.068)
Male analyst ideology	0.243** (0.110)	0.327*** (0.125)	0.158 (0.114)
Other controls	YES	YES	YES
Industry × Year FE	YES	YES	YES
Adjusted R ²	0.543	0.512	0.484
No. of observations	6,154	6,154	6,154

Table IA10
DID sample formation and broker closures over time

This table describes sample formation steps for the DID firm-year sample and lists the 13 broker closure events over the period 2005–2020 used in Table 3. Panel A reports the impact of various data matching steps and data filters on sample formation. Panel B lists the broker closure events used in Table 3. The number of treated firms is the number of firms previously covered by a female analyst from an exited broker, and has experienced a drop in female analyst coverage from the year prior to the broker closure event to the year after. The number of control firms is the number of firms previously covered by a male analyst from an exited broker, and has experienced a drop in male analyst coverage from the year prior to the broker closure event to the year after. The number of female (male) analysts is the number of female (male) analysts who were affected by a broker closure event, i.e., they were covering a (potential) treated (control) firm in the year before the closure event and not covering the same firm after.

Panel A: DID sample formation

	# firm-year obs.	# firms removed	# firms remaining (# broker closures)
I/B/E/S database, 2005–2021			11,432
Keep firms that were affected by at least one of the 133 broker closure events	45,212	7,585	3,847 (99 closures)
Keep firms with at least a two-year gap between two consecutive broker closure events that affected them	23,538	1,619	2,228 (94 closures)
Remove firms that 1) miss analyst coverage information either before or after a closure event; 2) lost both the female and male analyst coverage in the same year due to the same closure event; and 3) did not experience a drop in (female/male) analyst coverage between the year before a closure event and the year after. Only keep firm-year observations in the year before a closure event and the year after a closure event	1,082	1,687	541 (69 closures)
Merge I/B/E/S firm-year observations using the ticker-gvkey link file with Compustat by gvkey, as the unique firm identifier in I/B/E/S is ticker and the unique firm identifier in Compustat is gvkey	801	116	425 (62 closures)
Merge with Refinitiv to get E&S scores and Compustat to get firm characteristics.	126	362	63 (17 closures)
Use propensity score matching, keep only five control firms for each treated firm. The DID sample comprises five treated firms and 25 control firms	60		30 (13 closures)
Final sample	60		30

Panel B: Broker closure events over time

Broker	Last appeared in I/B/E/S stop files	Event type	Acquirer	Treated	Control	# female analysts	# male analysts	Source
Bear Stearns Companies Inc.	2008/6	M&A	JPMorgan Chase & Co.	0	1	52	423	https://www.sec.gov/Archives/edgar/data/19617/000089882208000286/pressrelease.htm
Lehman Brothers Holdings Inc.	2008/9	Closure		4	3	38	228	https://www.npr.org/2023/09/15/1199321274/lehman-brothers-collapse-2008-mortgages#:~:text=Fifteen%20years%20ago%2C%20the%20world,recessionary%20economy%20into%20a%20tailspin.https://www.federalreserve.gov/newsevents/testimony/bernanke20090625a.htm#:~:text=The%20agreement%20in%20principle%20on,pleased%20to%20take%20your%20questions.
Banc of America Securities LLC	2008/12	M&A	Merrill Lynch & Co Inc.	0	1	82	355	https://www.federalreserve.gov/newsevents/testimony/bernanke20090625a.htm#:~:text=The%20agreement%20in%20principle%20on,pleased%20to%20take%20your%20questions.https://www.capitaliq.com/CIQDotNet/company.aspx?companyId=61712978
HFP Capital Markets LLC	2010/8	Closure		0	1	5	5	https://www.capitaliq.com/CIQDotNet/company.aspx?companyId=61712978
Morgan Keegan & Company, Inc.	2012/4	M&A	Raymond James Financial	0	2	0	198	https://www.raymondjames.ca/waynemeneill/-/media/75715133b34645c5bb2fb17a0673800b.ashx
Murphy & Durieu, L.P.	2012/6	Closure		0	1	0	14	https://www.capitaliq.com/CIQDotNet/company.aspx?companyId=61376824&fromSearchProfiles=Truehttps://www.credit-agricole.com/en/press-releases/citic-securities-company-limited-and-credit-agricole-corporate-and-investment-bank-are-pleased-to-jointly-announce-the-completion-of-the-sale-and-phttps://www.globenewswire.com/news-release/2015/06/05/1109593/0/en/Stifel-Completes-Acquisition-of-Sterne-Agee.html
CLSA Americas, LLC	2013/7	M&A	CITIC Securities Company	0	3	19	82	https://www.credit-agricole.com/en/press-releases/citic-securities-company-limited-and-credit-agricole-corporate-and-investment-bank-are-pleased-to-jointly-announce-the-completion-of-the-sale-and-phttps://www.globenewswire.com/news-release/2015/06/05/1109593/0/en/Stifel-Completes-Acquisition-of-Sterne-Agee.html
Sterne Agee & Leach Inc.	2015/11	M&A	Stifel Financial Corp	0	3	0	106	https://www.globenewswire.com/news-release/2015/06/05/1109593/0/en/Stifel-Completes-Acquisition-of-Sterne-Agee.htmlhttps://brokercheck.finra.org/firm/summary/154133
Topeka Capital Markets Inc.	2016/6	Closure		0	1	2	66	https://brokercheck.finra.org/firm/summary/154133https://ir.brileyfin.com/2017-07-05-B-Riley-Financial-Closes-Acquisition-of-Wunderlich-Securitieshttps://ir.brileyfin.com/2017-06-01-B-Riley-Financial-Closes-Acquisition-of-FBR-Co#:~:text=Riley%20Financial%20and%20FBR%2C%20including,other%20filings%20with%20the%20SEC.https://www.prnewswire.com/news-releases/hilliard-lyons-agrees-to-join-baird-300756362.html
Wunderlich Securities Inc.	2017/7	M&A	B. Riley Financial	0	3	24	98	https://ir.brileyfin.com/2017-07-05-B-Riley-Financial-Closes-Acquisition-of-Wunderlich-Securitieshttps://ir.brileyfin.com/2017-06-01-B-Riley-Financial-Closes-Acquisition-of-FBR-Co#:~:text=Riley%20Financial%20and%20FBR%2C%20including,other%20filings%20with%20the%20SEC.https://www.prnewswire.com/news-releases/hilliard-lyons-agrees-to-join-baird-300756362.html
B. Riley Securities, Inc.	2017/11	M&A	FBR & Co.	1	4	8	74	https://ir.brileyfin.com/2017-06-01-B-Riley-Financial-Closes-Acquisition-of-FBR-Co#:~:text=Riley%20Financial%20and%20FBR%2C%20including,other%20filings%20with%20the%20SEC.https://www.prnewswire.com/news-releases/hilliard-lyons-agrees-to-join-baird-300756362.html
Hilliard Lyons	2018/5	M&A	Robert W Baird & Co Inc.	0	1	0	92	https://www.prnewswire.com/news-releases/hilliard-lyons-agrees-to-join-baird-300756362.htmlhttps://brokercheck.finra.org/firm/summary/40375
IFS Securities, Inc.	2019/8	Closure		0	1	0	29	https://brokercheck.finra.org/firm/summary/40375

Table IA11
Analyst gender and tones in discussing E&S issues

This table examines the relation between analyst gender and their tones in discussing E&S issues. Panel A presents the report-level regressions examining the relation between analyst gender and their tones in discussing E&S issues in analyst reports. The sample consists of 160,172 reports covering 13,409 firm-year observations (representing 1,572 unique firms). Panel B presents the call-analyst level regressions examining the relation between analyst gender and their tones in discussing E&S issues during earnings calls. The sample consists of 27,648 calls covering 10,439 firm-year observations (representing 1,319 unique firms). Definitions of the variables are provided in Appendix IA.B. Standard errors (in parentheses) are clustered at the analyst time year level. ***, **, * correspond to statistical significance at the 1, 5, and 10 percent levels, respectively.

Panel A: Analyst gender and their tones in discussing E&S issues in reports

Variable	Negative E&S tone (1)	Negative E tone (2)	Negative S tone (3)
Female	-1.608 (0.991)	0.012 (0.008)	0.006 (0.007)
Negative non-E&S tone	0.741*** (0.014)	-0.005*** (0.000)	-0.003*** (0.000)
Report length	-1.266*** (0.386)	0.010*** (0.003)	0.015*** (0.003)
Star analyst	-2.165* (1.279)	0.011 (0.011)	0.001 (0.009)
Education	0.415 (0.498)	-0.008* (0.004)	0.005 (0.004)
CFA	-0.882 (0.630)	-0.002 (0.005)	0.011** (0.005)
Forecast frequency	-0.216* (0.111)	0.001 (0.001)	0.002*** (0.001)
Forecast horizon	0.015*** (0.003)	-0.000*** (0.000)	-0.000*** (0.000)
# firms followed	0.048 (0.054)	0.000 (0.000)	-0.000 (0.000)
# industries followed	-0.516** (0.217)	0.001 (0.002)	0.003 (0.002)
General experience	0.025 (0.080)	0.000 (0.001)	-0.001 (0.001)
Firm × Year FE	YES	YES	YES
Broker × Year FE	YES	YES	YES
Adjusted R ²	0.278	0.265	0.219
No. of observations	160,172	160,172	160,172

Panel B: Analyst gender and their tones in raising E&S-related questions during calls

Variable	Negative E&S tone (1)	Negative E tone (2)	Negative S tone (3)
Female	-0.009 (0.011)	-0.007 (0.006)	-0.002 (0.010)
Negative non-E&S tone	0.049*** (0.010)	0.020*** (0.006)	0.028*** (0.009)
N_questions	-0.007 (0.009)	0.002 (0.006)	-0.011 (0.008)
Star analyst	-0.023 (0.015)	-0.003 (0.009)	-0.021* (0.012)
Education	-0.004 (0.005)	-0.003 (0.003)	-0.000 (0.004)
CFA	0.004 (0.007)	0.003 (0.004)	-0.001 (0.006)
Forecast frequency	0.002 (0.002)	0.001 (0.001)	0.001 (0.001)

Forecast horizon	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
# firms followed	0.000 (0.001)	-0.000 (0.000)	0.000 (0.001)
# industries followed	-0.004* (0.002)	-0.000 (0.001)	-0.003* (0.002)
General experience	0.002** (0.001)	0.000 (0.000)	0.002*** (0.001)
Firm × Year FE	YES	YES	YES
Broker × Year FE	YES	YES	YES
Adjusted R ²	0.008	-0.008	0.019
No. of observations	27,648	27,648	27,648
