Do Managers Walk the Talk on Environmental and Social Issues?_NZFM2021 *

Sudheer Chava† Wendi Du ‡ Baridhi Malakar§

Abstract

We train a deep-learning based Natural Language Processing (NLP) model on various corporate sustainability frameworks in order to construct a comprehensive Environmental and Social (E&S) dictionary that incorporates materiality. We analyze the earnings conference calls of U.S. public firms during 2007-2019 using this dictionary. We find that the discussion of environmental topics is associated with higher pollution abatement and more future green patents. Firms reduced their air pollution even after the U.S. announced its withdrawal from the Paris Agreement. Similarly, the discussion of social topics is positively associated with improved employee ratings. Overall, our results provide some evidence that firms do walk their talk on E&S issues.

Keywords: ESG, Deep Learning, Earnings Conference Calls, NLP
JEL Classification: G30, M14, G14, Q52, Q53, Q55

*We thank participants at 2020 Journal of Investment Management (JOIM) Conference and Georgia Institute of Technology, for comments that helped to improve the paper.

†Scheller College of Business, Georgia Tech, *Email:* `sudheer.chava@scheller.gatech.edu`

‡Scheller College of Business, Georgia Tech, *Email:* `wendi.du@scheller.gatech.edu`.

§Scheller College of Business, Georgia Tech, *Email:* `baridhi.malakar@scheller.gatech.edu`

1 Introduction

Over the past decade, there has been a tremendous increase in investor interest in the Environmental and Social (E&S) impact of corporations. In parallel, there is a significant increase in corporate communications about E&S issues. Around 90% of S&P 500 firms published a sustainability report in 2020, up from 20% in 2011. In fact, the Business Round Table, an association of Chief Executive Officers of America's leading companies, has redefined their longstanding emphasis on the purpose of corporations from shareholder primacy to benefiting all its stakeholders. At the same time, there are concerns about greenwashing and how much of the managerial talk actually translates into improved firm performance on E&S issues[1]. In this paper, we study whether managers walk their talk on the E&S issues by analyzing the E&S discussion in earnings calls and the subsequent performance of these firms on the E&S dimensions.

It is not clear whether managers have incentives to walk their talk on the E&S issues. Sustainability disclosures are not mandated corporate disclosures and there is ambiguity on how to measure firms' performance on the E&S issues. Unsurprisingly, there is significant divergence among various E&S ratings (see Berg, Koelbel, and Rigobon (2019)). In addition, asset managers do not always vote in support of the E&S issues (see de Groot, de Koning, and van Winkel (2021)). As a result, managerial talk may be cheap, while commensurate E&S-related corporate actions may involve significant costs.

In this paper, we develop a dictionary of environmental and social phrases using a deep learning–based model trained on nine leading sustainability standards and frameworks. This allows us to directly evaluate whether managers' talk on environmental and social aspects from earnings conference calls translates into real firm outcomes, without relying on the E&S ratings. We find that firms whose managers discuss more environmental issues also undertake more pollution abatement activities. Moreover, firms which discuss more environmental topics on their earnings conference calls have a higher number of green patents granted in the future.

[1]Some of the concerns about *greenwashing*, defined as making people believe that the company is doing more to protect the environment than it really is, are highlighted in the 2021 EU Sustainable Finance Disclosure Regulation (SFDR) that aims to partly address it.

Equivalently, employee ratings are higher for firms with more discussion of social issues in the earnings conference calls. Overall, our results suggest that the E&S discussion in earnings conference calls is credible and associated with improved performance on the E&S dimensions.

Identifying E&S discussion in earnings conference calls is challenging for multiple reasons. The first challenge is that there is no widely accepted definition on the environmental and social topics (Chatterji, Durand, Levine, and Touboul, 2016). Second, even if there is a consensus on the E&S topics, they may not necessarily be used verbatim by firms' executives or the business press. They may instead rely on short-hand synonyms or "nicknames" to refer to the same topics. Third, some of the E&S topics may be material for some industries but not for others. For example, air quality is important for the transportation industry, but not for financial services.

We mitigate these challenges by deploying a newly-developed deep learning model called Robustly optimized BERT approach (RoBERTa) (Liu, Ott, Goyal, Du, Joshi, Chen, Levy, Lewis, Zettlemoyer, and Stoyanov, 2019), to construct a novel dictionary. We identify the environmental and social phrases separately and distinguish whether they are material to the specific industry under that E&S topic. Importantly, this dictionary contains both the terminologies associated with E&S topics as well as all their synonyms used in common discourse.

We begin by treating the task of constructing a dictionary of E&S phrases as a text classification problem, which aims to assign "environmental" or "social" label to a large number of phrases. First, we construct a high-quality training dataset using the documents from nine leading sustainability standards and frameworks[2]. These documents describe the conceptual frameworks and definitions of the E&S topics viewed as important to investors, as well as the metrics used for each topic. These documents together provide a high-quality self-labeled textual data for us to train the deep learning model. We thus label all the textual sequences from the documents as "environmental", "social" or others.

The second step is to train the state-of-the-art deep learning model, RoBERTa, to solve the text classification problem. RoBERTa improves upon the Bidirectional Encoder Repre-

[2]Sustainability Accounting Standards Board (SASB), Global Reporting Initiative (GRI), Task Force on Climate-related Financial Disclosures (TCFD), Asset4, Dow Jones Sustainability Indices (DJSI), FTSE4Good, Innovest, MSCI KLD, and Sustainalytics.

sentations from Transformers (BERT) (Devlin, Chang, Lee, and Toutanova, 2018) model by expanding the training data and optimizing the training procedure. RoBERTa outperforms the BERT-Large model and the recently-developed deep learning model, XLNet (Yang, Dai, Yang, Carbonell, Salakhutdinov, and Le, 2019), in a variety of downstream tasks like question-answering and sentiment analysis. In this paper, we train the RoBERTa model using the labeled E&S training data developed in the previous step. During the training process, RoBERTa learns the inherent association between the textual sequence and its label to make better predictions.

Next, we use the trained RoBERTa model with the learned knowledge to make predictions on a large number of candidate phrases extracted from the sustainability documents and Wikipedia. We keep the phrases predicted as "environmental" or "social" by RoBERTa model with probability larger than or equal to 80%.

To validate whether the E&S phrases selected by the RoBERTa model are actually used in the business world, we generate a corpus of words from 415 corporate social responsibility (CSR) reports [3] and filter out the phrases which never occur in the CSR corpus. We manually check 14,037 phrases occurring more than five times over three rounds, and get a dictionary of 614 "environmental" phrases and 697 "social" phrases. Since materiality is important for sustainability analysis (Khan, Serafeim, and Yoon, 2016), we further train RoBERTa to classify the resulting phrases into "material" and "immaterial" for each industry under the E&S topics.

The use of the deep–learning model to construct the E&S dictionary has several advantages. First, RoBERTa utilizes *transfer learning*, using the knowledge gained through pre-training process, to adapt to the downstream tasks through *fine-tuning* on a domain-specific labeled data. The pre-trained deep learning model absorbs the general semantic and syntactic knowledge of the language from a large corpus of English words. Deep learning models with transfer learning technique represent a significant upgrade over the earlier deep learning models, such as Word2Vec (Mikolov, Sutskever, Chen, Corrado, and Dean, 2013) and FastText (Joulin, Grave, Bojanowski, and Mikolov, 2016), involving building models from scratch. Second, the large architecture of the deep learning model allows for a cascading of non-linear transformations

[3] We collect the CSR reports from SASB website: `https://www.sasb.org/company-use/sasb-reporters/`. Our downloaded reports are from firms who voluntarily post their CSR reports on SASB website during 2017 to 2020.

to define the hyperspace of textual information. Specifically, RoBERTa's 12-layer architecture makes it more powerful than Word2Vec's 2-layer architecture in analyzing textual information. Finally, the modern deep learning model allows for a more holistic classification of textual data (receiving the text sequence as a whole), as opposed to a piecemeal characterization of features in the traditional machine learning methods. By absorbing contextual information bidirectionally, the attention mechanism in RoBERTa model learns which part of the sequence needs more attention to improve the accuracy of its prediction.

Our dictionary reveals that managers refer to environmental and social topics about four times on average, during an earnings conference call. These discussions are dominated by environmental references during 2007–2019 with a gradual increase over the years. For environmental-related discussion, there is heterogeneity at the industry level. Our sample suggests that over 70% of firms relying on traditional fossil fuels consistently discuss environmental topics in their earnings calls. Meanwhile, the fraction of telecommunication firms involving discussion of environmental topics has increased from around 40% in 2007 to 68% in 2019. The proportion of earnings calls referring to social issues has remained fairly constant between 40–50% during the period. Overall, more than 70% of earnings calls in our sample discuss E&S topics.

Using this dictionary, we analyze whether the firm's discussion of E&S-related issues in the earnings calls is followed by corresponding actions along these dimensions. First, we consider the abatement of firm-level pollution. We use pollution prevention (P2) data from the Environmental Protection Agency (EPA). Our analysis focuses on the two most common forms of abatement: changes to operating practices and production improvements. Our baseline regression specification controls for firm fixed effects and time-varying heterogeneity at the industry × year levels. We find that the discussion of environmental topics is positively associated with the investment in pollution abatement. Firms appear to increase operations-related pollution abatement initiatives with more environmental content in their earnings calls. The effect is stronger when the environmental discussion is material to the industry. We find that a 1% change in environmental discussion in the earnings call is associated with a 9.47% increase in operations-related abatement activities. We do not find significant improvements in production-related processes on aggregate; however, there is a weak positive association when the environmental discussion is material to that industry.

We also relate environmental discussion in the earnings calls to the firm's pollution emissions. After controlling for unobserved heterogeneity at the firm and industry × year levels, we find that a 10% increase in discussion of E&S topics in the earnings call is associated with a 0.99% decrease in the firm's total pollution. For the average firm in our sample, this amounts to 25,720 pounds. The corresponding effect on air pollution is a reduction of 8,809 pounds (∼ 1% of the average value), which is the primary component driving our results. We find similar results by using the fraction of environmental topics alone as the main explanatory variable. Our evaluation of future years of pollution reveals some evidence of this negative correlation with total pollution in the subsequent year. But the effect diminishes statistically and economically thereafter. We also find evidence that managers seem to discuss environmental issues more prominently after the Paris Agreement. Firms seem to be committed to reducing emissions when managers talk about the environment, even after the US withdrawal from the international treaty on climate change.

Next, we find a positive association between the managers' talk about environmental topics and the number of green patents granted to their firms. This relationship holds over the one- to three-year window after the earnings call date. We also test the market reaction to the E&S talk in the earnings calls, and find a positive stock price reaction to the discussion of environmental topics. Overall, we find that firms with more environmental discussion in earnings calls have more pollution abatement initiatives and more green patenting activity. Correspondingly, we find a negative association between the environmental discussion and the overall pollution.

For the social dimension, we rely on anonymous employee reviews obtained from Glassdoor. We aggregate the overall rating and its components from Glassdoor at the firm-year-quarter level by taking an average. We examine the association between employee ratings and the discussion of social topics in earnings conference calls after controlling for firm and industry × year-quarter fixed effects. We find that a 1% increase in the discussion of E&S topics is associated with a 0.24-notch improvement in the overall rating from employees, which represents a 7.62% increase on the average overall rating of 3.15 in our sample. We also find a similar magnitude (although statistically weaker) by using the fraction of social topics alone as the explanatory variable.

This relationship becomes weaker in magnitude during the subsequent quarter of the earnings conference call, and becomes statistically insignificant thereafter.

We also analyze the individual components that constitute the overall employee rating from Glassdoor. We find that a 1% increase in the discussion of social topics is associated with a 0.24-notch improvement in ratings corresponding to compensation and benefits in that quarter. We find similar associations in terms of work-life balance (0.26-notch) and organizational culture (0.31-notch). We argue that these components primarily drive the improvement in overall employee rating. As further evidence, we show that the positive associations in work-life balance and organizational culture ratings are more strongly associated with the discussion of social topics that are material to that industry. Overall, our findings suggest that managers' talk on social dimensions are reflected in the observed employee ratings of the firm.

Our paper contributes to the growing literature in finance and economics on environmental, social, and governance (ESG) considerations (for example, Hong and Kostovetsky (2012); Chava (2014); Krüger (2015); Lins, Servaes, and Tamayo (2017); Liang and Renneboog (2017); Gibson, Krueger, and Schmidt (2019); Dyck, Lins, Roth, and Wagner (2019); Chava, Kim, and Lee (2020); Engle, Giglio, Kelly, Lee, and Stroebel (2020)). The measurement of ESG performance and the assignment of ratings is a complex process. There is substantial disagreement across rating agencies (Berg, Koelbel, and Rigobon (2019). Christensen, Serafeim, and Sikochi (2019) propose that greater ESG disclosure leads to higher disagreement across ratings. Barber, Morse, and Yasuda (2021) analyze the impact and traditional venture capital funds and show that investors are willing to sacrifice returns for non-pecuniary benefit. We build on this literature by employing deep learning techniques to construct a comprehensive dictionary covering a broad scope of E&S issues and incorporating materiality that would be relevant to investors. Based on this E&S dictionary, we document the relationship between the firms' talk and actions on the E&S topics. To the best of our knowledge, we are the first to address this problem of credibility in E&S talk by comparing the actual firm-level actions with the E&S discussion by the firm's managers.

Our paper is related to the literature on generating and applying various dictionaries to financial documents. Loughran and McDonald (2011) create a word list to capture six tone-related characteristics of 10-K filings. Bodnaruk, Loughran, and McDonald (2015) construct measures of financial constraints by parsing 10-K filings using their list of words on financial constraints. Bochkay, Hales, and Chava (2020) construct a dictionary of linguistic extremity from earnings conference calls. We build on this literature by creating a dictionary of E&S phrases employing the deep learning model, RoBERTa, instead of the bag-of-words approach and other similar methods previously followed in this literature.

Finally, our paper also contributes to the emerging literature that uses machine learning techniques in finance. To our knowledge, we are the first to apply a RoBERTa approach to analyze financial documents in the finance literature. Buehlmaier and Whited (2018) predict the probability of a firm being financially constrained based on the textual analysis of firms' annual reports. A recent study by Jha, Liu, and Manela (2020) applies the BERT model to study the cross-cultural differences in sentiment towards finance across eight countries. In a separate study, Jha, Liu, and Manela (2021) analyze the response of this finance sentiment to natural disasters and its impact on economic outcomes. Chava, Du, and Paradkar (2020) use BERT to analyze whether the discussion of emerging technologies in earnings calls is just hype or whether it conveys credible information to investors. In contrast, we fine-tune the pre-trained RoBERTa model with sustainability documents, to first generate a dictionary of E&S-related phrases. We then apply this dictionary to parse out references to E&S topics in earnings conference calls and also incorporate materiality.

2 Dictionary Development

To capture the environmental and social discussion in firms' earnings conference calls, we first need to identify the phrases that are relevant to the E&S topics. After aggregating the documentation from a broad range of sustainability standards and frameworks, we deploy the state-of-the-art deep learning model, Robustly optimized BERT approach (RoBERTa). Using RoBERTa, we construct a dictionary which contains both the E&S terminologies and their synonyms used in common discourse. Following prior literature (see Krüger (2015); Dyck, Lins, Roth, and Wagner (2019)), we exclude the governance-related issues from our consideration.

In this paper, we treat the task of constructing a dictionary of the environmental and social phrases as a text classification problem, which aims to assign "environmental" and "social" labels to a large number of phrases. There are three main steps to construct our dictionary: (1) construct a high-quality training sample; (2) train the deep learning model for classification purposes; (3) use the trained model to make out-of-sample predictions, and validate the selected phrases. Figure 1 provides an overview of the steps involved in the construction of the E&S dictionary. Finally, we leverage our dictionary to identify E&S phrases that are material to a given industry. We provide additional details below.

2.1 E&S Labeled Training Data

We construct a high-quality training data with each sequence labeled as "environmental", "social", or other dimensions based on the sustainability documents collected from nine leading sustainability standards and frameworks, namely: Sustainability Accounting Standards Board (SASB), Global Reporting Initiative (GRI), Task Force on Climate-related Financial Disclosures (TCFD), Asset4, Dow Jones Sustainability Indices (DJSI), FTSE4Good, Innovest, MSCI KLD, and Sustainalytics. The sustainability documents contain the textual information of the topics' definitions and the corresponding metrics. Guided by the conceptual framework provided by each organization, we label the sequences in the documents as per the corresponding dimension, i.e. "environmental", "social" or other. For example, as shown in Figure 2, the "greenhouse gas emissions" category is classified into "environmental" dimension by the SASB framework. Thus, we label all the textual sequences under this category as "environmental". Based this method, we get the E&S training data with five labels as shown in appendix Table IA4.

2.2 Deep Learning Model – RoBERTa

For text classification, we use the deep learning model, RoBERTa, which shows the state-of-the-art performance in a variety of downstream tasks like question-answering and sentiment analysis (Liu, Ott, Goyal, Du, Joshi, Chen, Levy, Lewis, Zettlemoyer, and Stoyanov, 2019).

RoBERTa is the latest in the suite of deep learning techniques that are adept at capturing the nuances and complexities of the English language. According to the *Distributional Hypothesis* (Harris, 1954), distinct phrases used in similar contexts tend to have similar meanings (Firth, 1957). RoBERTa model learns the meaning of phrases from the context in which they are used during the training process, and then uses this knowledge to identify the E&S phrases in general discourse.

RoBERTa offers several advantages over existing approaches commonly used for natural language processing (NLP) tasks. First, RoBERTa utilizes the most up-to-date method of *transfer learning*, where a deep learning model pre-trained with a large corpus can be employed again in downstream NLP applications through second training (or fine-tuning) process. RoBERTa is pre-trained with over 160GB of uncompressed text, consisting of BookCorpus + Wikipedia (16GB), CC-News (76GB), OpenWebText (38GB), and Stories (31GB), which enables it to absorb the general semantic and syntactic knowledge of the English language. Deep learning models with transfer learning techniques represent a significant upgrade over earlier deep learning techniques, such as Word2Vec and FastText, which build models from scratch. Second, RoBERTa is able to receive the entire sequences of various lengths as inputs to absorb the word order information (Loughran and McDonald, 2016). This is in contrast to splitting the sequence into fixed-length features in isolation as used in the bag-of-words (BOW) approach and the traditional machine learning (ML) techniques in finance and economics research (Hassan, Hollander, van Lent, and Tahoun, 2019; Bloom, Hassan, Kalyani, Lerner, and Tahoun, 2020). Moreover, the feature design in traditional ML techniques requires domain knowledge, thus limiting the model's ability to be generalized to other downstream tasks (Minaee, Kalchbrenner, Cambria, Nikzad, Chenaghlu, and Gao, 2021). Whereas, the attention mechanism in RoBERTa model allows it to learn which part of the input sequence is important to make accurate predictions. The third advantage of RoBERTa is its ability to extract a word's contextual information bidirectionally, which is an upgrade over the unidirectional deep learning models like OpenAI GPT (Radford, Narasimhan, Salimans, and Sutskever, 2018), which process words sequentially. Based on the contextual information, RoBERTa is able to understand multiple connotations of the *same* phrase depending on how the phrase is used, compared to the models such as Word2Vec and Glove (Pennington, Socher, and Manning, 2014), which are only capable of understanding one meaning for each unique phrase.[4]

During the training process with the E&S labeled data, RoBERTa learns the inherent association between the input sequences and their E&S labels. We then use this RoBERTa model with the learned knowledge to make predictions on a large number of candidate phrases, and focus on the ones which are labeled as "environmental" and "social".

2.3 Model Prediction

To make out-of-sample predictions, we extract 45,595 key phrases from the sustainability documents, and augment them by collecting 3,798,530 topics from Wikipedia under the categories related to sustainability. The trained RoBERTa model not only labels each phrase into a dimension (e.g., "environmental") but also reports the probability of this prediction.

We keep the 998,089 phrases predicted as "environmental" or "social" with a probability higher than or equal to 80%. In order to collect an exhaustive list of keyword phrases, we extract all unigrams (one-), bigrams (two-), and trigrams (three-word combinations) from each phrase. For instance, *LGBT rights in the Cook Islands* would yield *LGBT*, *LGBT rights*, *the Cook Islands*, etc. as the shorter phrases. This process generates over 2 million such phrases in our sample. We then supply these shorter phrases into RoBERTa, and retain those with the same threshold as above. Next, we get 792,453 unique phrases after cleaning those with stop words at the head or tail.

Having thus identified the phrases, it remains unclear whether they are actually used by the business society to discuss issues on sustainability. To address this concern, we generate a corpus with 415 CSR reports spanning 10,151,937 words. We drop the phrases which never occur in the CSR corpus. Thereafter, we manually check 14,037 phrases with a frequency larger than five in the CSR corpus. After three rounds of manual checking, we arrive at our dictionary

[4]For example, the word "Apple" represents a brand and a fruit respectively in the following two sentences: "I went to the Apple Store to buy an iPhone." and "Apple trees are planted in the backyard." Word2Vec produces a single vector to represent both meanings of the word "Apple." In contrast, RoBERTa is able to encode contextual information like "iPhone" and "planted in the backyard," thereby generating two different vectors for each meaning of the word "Apple" in the two sentences.

of E&S topics with 614 "environmental" phrases and 697 "social" phrases.

2.4 Material Environmental and Social Phrases

Identifying sustainability issues that are material to the industry is relevant for firms' disclosures and for guiding investors' asset allocation (Khan, Serafeim, and Yoon, 2016; Amel-Zadeh and Serafeim, 2018). Some E&S topics are important for some industries, but not for others. For example, "climate change" is material for the firms in oil and gas industry, but immaterial for those in the financial services. SASB provides the industry-specific guidance on material sustainability issues. Khan, Serafeim, and Yoon (2016) deploy this to distinguish the material and immaterial ratings, and find that firms with good ratings on material sustainability issues significantly outperform those with poor ratings on these issues. In earnings calls, some firms may focus on the sustainability issues which are material for their business, while others may prefer to discuss the immaterial issues. To analyze this aspect of E&S discussion, we follow the guidelines of SASB and identify the E&S phrases referring to the sustainability issues material to each industry.

For the E&S dictionary thus constructed, we use RoBERTa model to further classify each phrase as "material" or "immaterial" for each industry. In each industry, we generate the training data by labeling the sequences under the SASB definition as topics "material" to this industry, and otherwise. We then train the RoBERTa model to make predictions on the E&S phrases. Based on the E&S phrases which are material for each industry, we can further split the E&S discussions into material and immaterial ones.

3 Methodology and Identification Challenges

3.1 Empirical Specification

We use a multivariate panel regression approach to measure the association between firm-level outcomes and discussion of environmental and social issues in earnings conference call transcripts. Using the dictionary of E&S phrases, we quantify the fraction of environmental and social topics in a given call transcript as detailed in Section 4.1. Our baseline empirical specification is given below:

$$Y_{i,f,t} = \alpha + \beta_0 * Level(E)_{i,f,t} + \beta_1 * Level(S)_{i,f,t} + Controls_{i,t} + \delta_{f,t} + \phi_i + \epsilon_{i,f,t} \qquad (1)$$

where $Y_{i,f,t}$ represents the outcome variable corresponding to firm i (operating in industry f) in time t. The key independent variables, $Level(E)_{i,f,t}$ and $Level(S)_{i,f,t}$, represent the fraction of words devoted to the discussion of environmental and social topics respectively, in firm i's earnings conference calls during time t, belonging to industry f. The construction of these variables is discussed in more detail in Section 4.1. $Controls_{i,t}$ represents a vector of firm-level time-varying observable characteristics which might be correlated with the outcome variables. The baseline specification includes firm fixed effects (ϕ_i) and industry × time fixed effects ($\delta_{f,t}$). We double cluster standard errors at the firm and time levels. We use firm-level outcome variables from multiple sources that are aggregated/observed at an annual or quarterly frequency. We provide descriptions of key variables used in the analysis in Table A1.

The inclusion of firm fixed effects in the specification helps account for any firm-specific, time-invariant unobserved characteristics, such as a firm's general proclivity to always or never include any E&S discussion in its conference calls. Moreover, controlling for the industry × time fixed effects helps us account for any time-varying trends within industries that are potentially correlated with the E&S concerns relevant to specific industries. We suitably modify the equation when we apply it to outcome variables that are available at a quarterly or annual frequency.

3.2 Identification Challenges

The first major challenge we observe in our empirical research is the validity of the dictionary. We follow a layered approach in developing the dictionary, starting with training the RoBERTa model and subsequent validation using the corpus of words obtained from corporate social responsibility (CSR) reports filed by corporations. The CSR reports come from 2017-2020, while our sample of earnings conference calls dates back to 2007. We assume that the set of "environmental" and "social" phrases/words do not change substantially over this period. In any case, we further substantiate our validation by manually checking the list of words in our dictionary.

To further mitigate the concern over our dictionary construction, we also analyze the words in our dictionary for material significance to a given industry. This helps us to narrow down phrases and references that are particularly relevant to "environmental" or "social" issues in that industry. To this end, we rely on SASB materiality map at the industry level. Broadly, this enables us to identify the discussion of the E&S issues which are important for that industry. See Section 2.4 for more details.

A second major threat to our identification comes from an omitted variable problem. Firms with higher profitability, which perform well financially, could be the ones which also discuss more environmental and social issues. Because they have the financial resources, they may also be involved in actively reducing emissions or treating employees better. We cannot completely rule out this possibility. In light of this limitation, we restrict our inference to signify association between the firm action/outcome and the discussion of relevant environmental or social topics in their conference calls.

4 Data

In addition to the construction of our dictionary detailed in Section 2, we use data on four major aspects, namely: i) earnings conference calls, ii) environmental pollution, iii) employee ratings, and iv) green patents. We provide details on each one of them in this Section below. We obtain firm-level financial information from Compustat, IBES, and CRSP.

4.1 Earnings Conference Calls

Earnings conference calls provide a good setting to analyze the discussion of E&S topics by firms' management for two reasons. First, existing literature has documented that earnings conference calls convey critical corporate information to the market (Bowen, Davis, and Matsumoto, 2002;

Brown, Hillegeist, and Lo, 2004). Second, managers speaking at earnings conference calls are less constrained than in the regulatory filings to the Securities and Exchange Commission (SEC) (e.g. 10-K used in Hoberg and Maksimovic (2015)) because they are able to interact with participants during the Q&A session in a more conversational format.

We collect data on 159,138 earnings conference call transcripts of U.S. public firms during 2007 to 2019 from SeekingAlpha.[5] Using the firm name, the stock ticker, and the earnings conference call date, we find that 143,473 transcripts in our data can be merged to CRSP data. Next, we follow the approach in Bochkay, Hales, and Chava (2020) and only retain transcripts with more than one thousand words [6], and find 132,295 transcripts that meet these criteria.

Based on the dictionary constructed (see Section 2) to identify the level of E&S discussion, we aggregate the frequency of E&S phrases mentioned in a given firm's earnings calls during a period of time. We scale this by the total number of words in the corresponding earnings call transcripts. For example, while studying the association between a firm's E&S discussion and its annual pollution performance (see Section 5.1), we sum up the frequency of E&S phrases appearing in this firm's earnings call transcripts for each year, scaled by the total number of words in these transcripts during the year.

$$Level(E)_{i,t} = \frac{Environmental\ Frequency_{i,t}}{Number\ of\ Words\ in\ Transcript_{i,t}}$$

$$Level(S)_{i,t} = \frac{Social\ Frequency_{i,t}}{Number\ of\ Words\ in\ Transcript_{i,t}}$$

$$Level(E+S)_{i,t} = \frac{Environmental\ Frequency_{i,t} + Social\ Frequency_{i,t}}{Number\ of\ Words\ in\ Transcript_{i,t}}$$

We report the transcript-level summary statistics for the fraction of environmental and social words of the sample used in analyzing the market reaction in Panel A of Table 1. The distribution on the corresponding frequency of words is provided in Panel A of Table IA1. On average, the E&S phrases are mentioned around 4 times during each earnings conference call.

[5] We thank James Moon for graciously sharing his earnings conference calls data.

[6] We retain the longest earnings call transcript for a given firm-earnings call date, if there are multiple filings for that date. Some of the shorter files correspond to the previous version of earnings call transcript.

The average frequency of environmental phrases is more than twice as large as that for the social phrases. There are five environmental and social words per 10,000 words in the transcripts of this sample. When we split these into their constituent groups, we find an average of three environmental words and two social words for every 10,000 words in the conference call transcripts. In Figure 3, we present a yearly variation of the proportion of transcripts mentioning E&S topics during the earnings calls. Over 40% of the transcripts consistently include some discussion on social topics throughout the sample period. On the other hand, there is a gradual increase in the discussion of environmental topics during 2007-2011. Thereafter, this metric remains fairly uniform between 55-60%. Further, our dictionary also allows us to classify the E&S phrases into those that are material to the specific industry of the firm versus those that are not. Figure 4 shows the proportion of earnings calls mentioning E&S phrases across years and Fama-French 12 industries. We observe heterogeneity across industries in this regard. For example, companies in the utilities and energy industries are more likely to discuss E&S topics in their earnings conference calls than the financial services firms.

4.2 Environmental Pollution

We use plant-level data on toxic chemical releases reported under the Toxic Release Inventory (TRI) program from the Environmental Protection Agency (EPA) during 2007-2019. While the EPA has collected this information since 1987, our sample starting year is restricted by the availability of earnings conference call transcripts from Seeking Alpha (Section 4.1). The EPA reports chemical-level emissions data in TRI for plants with a facility of over 10 full-time employees, operating in one of the 400 industries defined at the six-digit NAICS level, and using one of nearly 600 chemicals. The TRI data are self-reported by firms and the EPA conducts audits to investigate anomalies. Misreporting could lead to penalties of a civil or criminal nature (Xu and Kim (2020)). Bui and Mayer (2003) find limited evidence in favor of over– or underestimation of toxic releases in the EPA-TRI data. For each firm in a given year of our sample, we aggregate the number of pounds released into the ground, air, and water from all its plants. We require each firm to be present throughout all the years of our sample period.

Our data on abatement activities comes from EPA's Pollution Prevention (P2) database. Under this program, plants included in the EPA-TRI database are required to document source reduction activities at the chemical level to restrict the amount of hazardous substances released into the environment. We use the firm names from various datasets by the EPA to create a crosswalk between EPA and firm names in the earnings conference calls data. We implement a combination of such matches across various files in the EPA to capture permutations of facility spellings and firm names used in the EPA. Unfortunately, these names are not absolutely consistent over the years and we overcome this limitation through iterations of developing the crosswalk. To account for possible differences in names, we standardize the strings by removing common suffixes like "Corp.", "LLC", "Inc." etc. Our matched dataset contains 1,067 unique firms for which we observe E&S discussions in the earnings call transcripts and EPA-TRI pollution. We provide the coverage of unique firms over the years in Figure IA1. This remains broadly uniform during the sample period with a slight decline toward 2019. The associated coverage of the plant-level facilities from EPA-TRI is shown in Figure IA2.

The average firm in our sample emits total emissions of about 2.58 million pounds, with ground emissions forming the largest component amounting to 1.50 million pounds. Figure 5 shows the time series of aggregate emissions for the three categories of pollution over the sample period. The overall trend is consistent with previous findings that report a decline in emissions, primarily driven by air pollution (Akey and Appel, 2021; Shapiro and Walker, 2018). On average, there are 24% of firms in the sample which adopt some measure of pollution abatement based on operating procedures. The corresponding fraction for process improvement methods is 23%. We provide details on the distribution of pollution in Panel B of Table 1.

4.3 Green Patents

We study firms' patenting activities to analyze their performance on green innovation. The data on patents granted are provided by Kogan, Papanikolaou, Seru, and Stoffman (2017).

The primary challenge is to identify whether a patent is related to the environmental-related technology or not. The Organization for Economic Co-operation and Development (OECD) reports a list of Cooperative Patent Classification (CPC) codes (Haščič and Migotto, 2015) to identify patents which contribute to solving environmental problems, like climate change, air pollution, water scarcity, etc.[7] Cohen, Gurun, and Nguyen (2020) make use of this framework to analyze green innovation and ESG scores across industries. Based on the CPC data obtained from USPTO, we identify the set of green patents granted to U.S. public firms.

Our sample construction in this analysis follows the literature. Specifically, we require that the transcripts are associated with firms with positive total assets, with non-missing earnings announcement dates in COMPUSTAT, with stock prices greater than $1 at the end of the fiscal quarter, with market values of equity larger than $5 million at the end of the fiscal quarter, and with non-missing dependent and control variables. Following the approach in Chava, Oettl, Subramanian, and Subramanian (2013), we further restrict the sample to only include potentially innovative firms that have filed at least a single patent during 2000–2019. We use Poisson regression for our analysis of the number of green patents. After accommodating these considerations, we are left with 32,074 earnings call transcripts during the one-year window analysis. We provide the summary statistics for this sample in Panel C of Table 1. The average firm in the sample has 2 green patents granted. Of these, nearly 1.85 are classified as related to climate change mitigation by the OECD. In addition, we report the distribution of the two-year and three-year samples in Panel C of Table IA1.

4.4 Employee Ratings

Our employee ratings come from Glassdoor reviews for the corresponding firm during 2008-2019[8]. The Glassdoor website was launched in 2008 to host anonymously written voluntary reviews from current and former employees. We use the overall firm rating and its components

[7] `https://www.oecd.org/environment/consumption-innovation/ENV-tech%20search%20strategies,%20version%20for%20OECDstat%20(2016).pdf`

[8] We are thankful to Reza Ferhadi for sharing their data with us.

provided by employees on a scale of one to five. Glassdoor requires email verification from an active email address or a valid social networking account to prevent self-promotion by the company. Further, the site also moderates content using their two-step approach consisting of algorithmic detection and human checking to eliminate fraudulent reviews.

We consider the average employee ratings in a given calendar year-quarter for each firm in the sample. Our final sample consists of 2,125 unique firms matched to their earnings conference call transcripts. We report the summary statistics for the distribution of overall employee reviews and its components in Panel D of Table 1. The mean value of the overall ratings is 3.15 which is similar to the average rating for compensation and benefits. The standard deviation across all the ratings is over 0.70 notches. Given that this distribution is not very skewed, we do not perform log transformation of the raw values thus obtained. Our distribution of the ratings looks similar to other papers in the literature (Green, Huang, Wen, and Zhou, 2019; Chemmanur, Rajaiya, and Sheng, 2019).

5 Results

We begin our analysis by showing results for the association of firm-level pollution abatement with the discussion of environmental words in Section 5.1. In this regard, we also evaluate the firm-level emissions. Next, we examine firms' innovation activities using green patents (Section 5.2). In Section 5.3, we provide evidence from employee ratings and discussion of social topics in earnings conference calls. Finally, we consider the stock market's immediate reaction to the discussion of environmental and social topics in earnings conference calls (Section 5.4).

5.1 Firm Pollution and Environmental Topics

Firms that are concerned about the environment may likely invest in pollution abatement. Such investments may reduce overall firm emissions provided firms do not increase production in order to capitalize on the emission reduction technologies. We measure abatement activities at the firm level using EPA's P2 database. Specifically, we focus on the two most common abatement categories: modifications to operating practices and process improvements. While good operating practices are related to improving maintenance or quality control, process improvements may be associated with improving chemical reactions or enhancing process controls. We follow Akey and Appel (2021) to classify abatement activities into these two types.

In Table 2, we test whether discussion on environmental and social issues is associated with changes in annual pollution abatement by firms using Equation (1). We double cluster standard errors at the firm and year levels. The dependent variable in Columns (1)-(3) is an indicator for abatement related to operating practices, and the dependent variable in Columns (4)-(6) is an indicator for abatement related to process improvements. We count the total number of abatement activities of each type across all plants for a given firm in a year. We denote the abatement indicator as one if the firm undertakes at least one such activity for the year. Our results suggest that firms increase abatement related to operating practices when they discuss environmental issues. The association with process improvements is relatively weaker.

Column (1) shows that when the level of environmental and social discussion increases by 1%, the associated abatement in operations increases by 8.49%. Similarly, in Column (2) we find evidence for a 9.47% increase in operations associated with a 1% change in environmental discussions alone. In Column (3), we provide results by splitting the environmental topics into those that are material or immaterial to the firm's industry. We find that the impact is mostly driven by words material to the industry. The economic magnitude of this association is sizable given that the average likelihood of operating process abatement is 24%. For abatement related to the production process, we find similar evidence in Column (6) for association with environmental topics material to the industry. However, the statistical significance is weaker.

Based on the analysis above, we expect that firms would reduce their overall emission of pollution in association with managers' discussion on environmental topics. We study the pollution emissions aggregated up to the firm level using Equation (1) at the firm-year level. To this end, we consider the air, water, and ground emissions from the Toxic Release Inventory database provided by the EPA. Table 3 shows the results with standard errors double clustered at the firm and year levels. We normalize emissions by the firm-level cost of goods sold in our baseline approach. Our dependent variable is the natural logarithm of one plus the normalized toxic chemical release, $\log(1 + \frac{Release_t}{COGS_{t-1}})$. We show robustness to our choice of scaling the pollution emissions in Table IA2.

In Column (1) of Panel A, we show the unconditional correlation of total pollution with the proportion of environmental and social words used in the earnings conference call discussions with firm and year fixed effects. In Column (2), we show results by introducing firm-level controls consisting of logged revenue, leverage, logged assets and capital intensity. Our preferred specification is presented in Column (3) where we control for unobserved heterogeneity at the firm and industry × year levels. We use Fama-French 30 industry classification. The coefficient of -54.23 suggests that a 10% increase in discussion of E&S topics is associated with 0.995% decrease in total pollution. For the average firm in our sample, this amounts to 25,720 pounds. This effect is statistically significant and economically meaningful, primarily driven by the reduction in air emissions (Column (4)). Our results suggest that a 10% increase in E&S discussion is associated with a decrease of 8,809 pounds of air emissions (~1% of the average). We find similar results by using only the level of environmental discussion in Panel B.

We also evaluate the association of future pollution and environmental discussion. In Panel C, we show our results by changing the dependent variable to total pollution in the future. Column (1) suggests a similar negative association implying 0.994% reduction in total pollution one year after a 10% increase in discussion of environmental topics. This effect reduces substantially (economically and statistically) in the longer horizon of two and three years ahead, as shown in Columns (2) and (3), respectively. Further, in Panel D we replicate our analysis using forward years of air pollution alone which seems to drive the results in Panel A. As before, Column (1) suggests that a 10% increase in the discussion of environmental topics is associated with 0.997% reduction in air pollution one year after. This effect reduces to 0.98% two years after (Column (2)) but is only marginally significant. Overall, our results suggest a negative correlation with pollution in the year of E&S discussion reported in the earnings conference calls but the effect reduces in the future years.

Next, we exploit the novelty of our dictionary data to understand whether our results are driven by the topics that are material to a given industry or not[9]. Section 2.4 provides details

[9]While it is appealing to consider whether the impact is stronger for consumer-facing industries, we are restricted by our sample. Only a small fraction of the firm-year observations belong to the retail industry as per two-digit SIC classification.

on how we classify environmental and social topics into those that are material to the firm's industry or otherwise. Based on this classification, we report our results from the baseline Equation (1) in Table 4. Column (1) suggests that most of the correlation in total pollution and environmental discussion is driven by topics material to the industry. A 10% increase in those discussions is associated with almost 1% reduction in total pollution. This nearly accounts for the overall effect reported earlier in Table 3. However, we lose some statistical power given the high requirement imposed on the data due to the industry × year fixed effects. We do not find similar evidence by using air pollution as the outcome variable in Column (2). We argue that this may be due to the lower specificity of industry-relevant words given the prominence acquired by air pollution during the sample period. For instance, Bi (2017) finds that coal-fired plants respond to media-specific environmental regulation by shifting their toxic air releases to waterways, land and offsite recycling transfers.

5.1.1 The Paris Agreement and Environmental Discussion

The Paris Agreement is a legally binding international treaty on climate change which was adopted on December 12, 2015. Its primary goal is to restrict the rise in global temperatures to 1.5 degrees Celsius during the twenty-first century. The Agreement calls upon signatory nations to submit their plans aimed at reducing greenhouse gas emissions. In light of the national plans, firms would likely have to comply with more stringent environmental regulations to reduce pollution and mitigate climate change. We assess the implication for our measure of environmental discussion in the earnings conference calls based on the Paris Agreement.

Using a modified version of our baseline specification in Equation (1), we regress the *Level(E)* (in %) on annual dummies corresponding to calendar years while controlling for firm and industry fixed effects. We double cluster standard errors by firm and year-quarter. Panel A in Figure 6 shows the estimates of the regression coefficients along with the 95% confidence intervals. We use the years prior to 2015 in our sample as the omitted benchmark. Our results suggest that there is an increase in environmental discussion in earnings calls after the Paris Agreement, with the effect nearly doubling in 2019. The effect is statistically significant and distinguishable from zero starting with the first full year after the Agreement in 2016.

The tone of the environmental discussion reveals the views of the speaker regarding the environmental issues. In this regard, we explore whether the tone in which environmental topics are discussed in earnings calls changes over time. For each earnings call transcript in our data, we conduct sentiment analysis on sentences containing environmental phrases. We use the RoBERTa model trained on the Stanford Sentiment Treebank data (SST-2, Socher, Perelygin, Wu, Chuang, Manning, Ng, and Potts (2013)) to classify sentences containing environmental phrases into those with "positive" or "negative" sentiment. For each call transcript, we separately calculate the frequency of environmental phrases with positive sentiment and with negative sentiment. We depict these dynamic trends in Panel B and Panel C of Figure 6 using a similar regression approach as before, benchmarked to years prior to 2015. Our results show that the positive sentiment has increased since 2016 onward. Whereas, there was a dip in the negative sentiment around 2018. We also calculate the net sentiment of the environmental discussion in each earnings call as the difference between the levels of positive and negative environmental discussions. As shown in Panel D of Figure 6, the net sentiment of environmental discussion in earnings conference calls has increased substantially over the recent years.

Our results above suggest that managers seem to discuss environmental issues more prominently after the Paris Agreement. They may be motivated to acknowledge the importance of environmental and climate-related issues due to greater emphasis in the news media and related coverage. Alternatively, firms may indeed be adopting measures to align their carbon footprint in line with the expected regulatory changes. In order to understand how firms respond to the Agreement, we use a modified version of our baseline Equation (1). Since the United States had announced to cease all participation in the Paris Agreement on June 1, 2017, we consider three periods around the Agreement. We interact the equation with dummies corresponding to before, during and after withdrawal from the Paris Agreement. We additionally control for group-year fixed effects in this analysis.

We report our regression results in Table 5. First, we find that the negative association between total pollution and discussion of environmental topics becomes stronger after the Paris Agreement. Second, our results suggest that this effect is magnified after the US announced its withdrawal from the climate accord. This is not surprising given that there is some anecdotal evidence to suggest that CEOs and business leaders had been supportive[10] of the US maintaining its commitment to the Paris Agreement. Across the different media of pollution, we find that this result is primarily driven by the effect on air pollution. This is consistent with our earlier results which suggest a greater reduction in air emissions associated with the discussion of environmental topics. In particular, the inverse association between pollution and *Level(E)* is statistically more pronounced during the withdrawal period than during the Agreement, both for total pollution and air pollution. Overall, our results suggest that firms seem to be committed to reducing emissions when their managers discuss environmental topics even after the US withdrawal from the international treaty on climate change.

5.2 Future Green Innovation and Environmental Topics

In this section, we examine whether firms follow up the discussion of E&S phrases in their earnings calls with the measurable environmental-related innovation after the earnings calls. We measure innovation outputs in terms of the green patenting activities of firms following their earnings calls. For each earnings call, we identify the green patents granted to the firm in the three-year window after the earnings call date.

In Table 6, we examine how the level of E&S discussion in an earnings conference call is related to the number of green patents granted ex post. We use a modified version of Equation (1), by omitting the firm fixed effects while ensuring comparison across firms in the same industry over time. The specifications are estimated using a Poisson regression model with fixed effects based on industry × year-quarter level and standard errors double clustered at the firm and year-quarter levels. In Columns (1) and (2) of Panel A, the dependent variable is the number of the green patents granted in the one-year horizon after the earnings call. We find that firms with a higher level of E&S discussion in their earnings calls have a higher number of green patents granted in the one-year window after the earnings calls. In addition, this effect is mainly driven by the level of environmental discussion in the earnings calls. For the average firm in

[10] https://www.reutersevents.com/sustainability/75-ceos-call-us-stay-paris-agreement-emissions-continue-rise
https://hbr.org/2017/05/u-s-business-leaders-want-to-stay-in-the-paris-climate-accord

the sample, one standard deviation increase (0.08%) in the level of environmental discussion is associated with around 16.77% ($e^{193.8*0.08\%}-1$) increase in the number of green patents granted in the next one year (Column (2)). This inference remains unchanged when we examine the two-year window (Panel B) and three-year window (Panel C) after the earnings calls.

Beyond the evidence on the overall green innovation provided above, we also separately examine the green patents of different classifications after the earnings calls. Based on OECD's guidelines, we further consider the green patents which are related to environmental management and water-related adaptation technologies (***Env&Water***, Columns (3) and (4)), and those which are related to climate change mitigation (***Climate***, Columns (5) and (6)). These groups are not mutually exclusive. As before, we find that the level of E&S discussion is positively associated with the number of green patents of each group. Moreover, this association is stronger if we focus on the discussion of environmental topics. Overall, our results in this section suggest that there is a positive relationship between the discussion of E&S topics with the firms' green patenting activity.

5.3 Employee Ratings and Social Topics

In this Section, we present our analysis for the correlation between firm outcomes that might reflect the discussion on social issues in earnings conference calls. Specifically, we evaluate the association of average firm-level employee ratings aggregated up to the quarterly frequency from Glassdoor.

First, in Panel A of Table 7, we report the results from our baseline Equation (1) with average employee overall ratings from Glassdoor as the outcome variable. We double cluster standard errors at the firm and year-quarter levels. Our firm level controls consist of logged value of assets and market-to-book ratio, leverage, return on assets and stock returns. Column (1) shows the correlation with just the firm and year-quarter fixed effects. We find a statistically strong positive correlation amounting to 25.26 units. In Column (2), we show our results by adding firm-quarter level controls. Our preferred specification comes from Column (3) where we introduce industry × year-quarter fixed effects to further absorb unobserved variation across industries over time. The coefficient of 24.60 is statistically significant and suggests that a 1% increase in the discussion of environmental and social topics is associated with a 0.24-notch improvement in overall rating from employees in that quarter[11]. This represents an improvement of 7.62% on the average overall rating of 3.15 in our sample. Further, in Column (4), we report our results by changing the dependent variable to the overall rating observed in the subsequent quarter. The impact is economically weaker, although still statistically significant. Columns (2) and (3) show results based on the overall ratings observed during two and three quarters ahead. We find that the statistical relationship does not hold in these cases.

We follow a similar analysis in Panel B using only the discussion of social topics. As before, our preferred baseline specification comes from Column (3) where we control for firm and industry-quarter fixed effects, in addition to firm-quarter controls. Our results show a magnitude of 29.73 although the statistical significance is weaker. The result suggests that a 1% increase in discussion of social topics is associated with a 0.29-notch improvement in overall ratings from employees in that quarter. When we project this relationship onto the future quarters in Columns (4)-(6), we do not find this relationship to hold. Overall, we find that there is a strong positive association between overall ratings from employees and the discussion of environmental and social topics. The relationship seems to become weaker gradually over the subsequent quarters.

Next, we examine the different components of the Glassdoor employee ratings that constitute the overall rating. We show our results in Panel C of Table 7 using the fraction of social topics as the key explanatory variable. We start with our baseline specification in Column (1) with the overall rating as the dependent variable. In the subsequent columns, we show the association corresponding to each component of the overall rating as the dependent variable. Column (3) suggests that a 1% increase in discussion of social topics is associated with a 0.24-notch improvement in ratings corresponding to compensation and benefits in that quarter. Likewise, we find that a higher discussion on social topics is also correlated with better

[11] It is not surprising to find that employees seem to care about environmental topics as well, when submitting employer reviews. For example, survey evidence suggests that 72% of respondents among UK office workers were concerned about environmental ethics. A focus on sustainability may also help firms in retaining their employees and in engaging with them more beneficially.

ratings on work-life balance (Column (5)) and culture (Column (6)). These associations are statistically significant and the economic magnitudes are comparable to the impact on overall rating discussed before. In summary, the improvement in overall ratings seems to be driven by ratings on compensation and benefit, work-life balance, and culture.

In order to understand more about what drives these results, we follow a similar approach as in Section 5.1. We divide the fraction of social topics into those that are material to the given firm's industry versus otherwise. We report our results in Panel D using the preferred baseline specification. In Column (1), we find that there is no statistically significant correlation between overall rating and the fraction of industry-material social topics. However, Columns (3) and (5) suggest that impact on ratings for compensation and benefits, and work-life balance respectively, is indeed driven by social topics material to that industry. The coefficients suggest an improvement of 0.34-notch and 0.39-notch in compensation and benefits and work-life balance ratings, respectively. These correspond to 10.76% and 12.04% of the average values of the corresponding ratings in the sample.

Taken together, our results in this section provide some evidence toward marginal improvement in average employee ratings, especially following the discussion of industry-specific material words. The effect is driven by ratings corresponding to compensation and benefits, and work-life balance. We do not find evidence for future improvements in ratings.

5.4 Stock price response to E&S discussion

We also explore the immediate price response to the use of E&S phrases in earnings conference calls. The dependent variable in the analysis is the seven-day cumulative abnormal return (CAR[−3,+3]), calculated using the market model. Our independent variables of interest are the level of environmental and social discussion ($Level(E+S)$). We also consider them separately ($Level(E)$, $Level(S)$) for our analysis. The construction of these variables is described in Section 4. Our findings are presented in Table 8.[12]

In Column (1), we study the impact of the overall E&S discussion on the immediate stock

[12] The sample used in this analysis is similar to that of the green innovation analysis, except for removing the innovative–firm restriction and not requiring to have next three-year patent records.

price response to earnings conference calls. We control for time-varying firm characteristics at the time of the earnings conference call, the firm fixed effects, and the industry × year-quarter fixed effects. We find that the coefficient estimate on $Level(E + S)$ is positive and significant at the 10% level, which suggests that earnings calls with a greater discussion of E&S issues generate a positive immediate price reaction. In Column (4), we separately examine the impact of environmental and social discussion on the immediate stock price response to the earnings conference calls. We find that the market only positively react to the discussion of environmental topics in the earnings call, while no significant stock price response is found to the discussion of social topics.

6 Conclusion

There has been an increasing focus on firms' performance on environmental and social dimensions. However, recent evidence suggests that there is much disagreement and confusion among ESG ratings provided by different agencies and there are significant concerns about greenwashing. We construct a novel dictionary on environmental and social topics, using a deep learning–based NLP model trained on comprehensive sustainability frameworks. Our approach also provides the crucial advantage of incorporating industry-specific materiality of topics. We deploy this dictionary into parsing out phrases within earnings conference calls of U.S. public firms. We show that firms talking about the environment also invest in pollution abatement initiatives, especially when such topics are material to that industry. Consistent with these results, we find a negative association in overall pollution with respect to earnings calls that discuss the environment. Our findings also suggest that there is a positive association between the green patenting activities and managers' talk about the environmental topics. On the social dimension, we find a positive relationship between the discussion of environmental and social topics, and the overall employee rating, suggesting that employees care about the E&S profile of the firm. Our evidence suggests that managers do walk some of their talk on the E&S issues.

References

Akey, P. and I. Appel (2021). The limits of limited liability: Evidence from industrial pollution. *The Journal of Finance 76*(1), 5–55.

Amel-Zadeh, A. and G. Serafeim (2018). Why and how investors use esg information: Evidence from a global survey. *Financial Analysts Journal 74*(3), 87–103.

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

Berg, F., J. F. Koelbel, and R. Rigobon (2019). *Aggregate confusion: The divergence of ESG ratings.* MIT Sloan School of Management.

Bi, X. (2017). "cleansing the air at the expense of waterways?" empirical evidence from the toxic releases of coal-fired power plants in the united states. *Journal of Regulatory Economics 51*(1), 18–40.

Bloom, N., T. A. Hassan, A. Kalyani, J. Lerner, and A. Tahoun (2020). The geography of new technologies. *Institute for New Economic Thinking Working Paper Series* (126).

Bochkay, K., J. Hales, and S. Chava (2020). Hyperbole or reality? investor response to extreme language in earnings conference calls. *The Accounting Review 95*(2), 31–60.

Bodnaruk, A., T. Loughran, and B. McDonald (2015). Using 10-k text to gauge financial constraints. *Journal of Financial and Quantitative Analysis*, 623–646.

Bowen, R. M., A. K. Davis, and D. A. Matsumoto (2002). Do conference calls affect analysts' forecasts? *The Accounting Review 77*(2), 285–316.

Brown, S., S. A. Hillegeist, and K. Lo (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics 37*(3), 343–366.

Buehlmaier, M. M. and T. M. Whited (2018). Are financial constraints priced? evidence from textual analysis. *The Review of Financial Studies 31*(7), 2693–2728.

Bui, L. T. and C. J. Mayer (2003). Regulation and capitalization of environmental amenities: evidence from the toxic release inventory in massachusetts. *Review of Economics and statistics 85*(3), 693–708.

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

Chava, S. (2014). Environmental externalities and cost of capital. *Management Science 60*(9), 2223–2247.

Chava, S., W. Du, and N. Paradkar (2020). *Buzzwords?* Georgia Institute of Technology.

Chava, S., J. Kim, and J. Lee (2020). *Doing Well by Doing Good? Risk, Return, and Environmental and Social Ratings.* Georgia Institute of Technology.

Chava, S., A. Oettl, A. Subramanian, and K. V. Subramanian (2013). Banking deregulation and innovation. *Journal of Financial economics 109*(3), 759–774.

Chemmanur, T. J., H. Rajaiya, and J. Sheng (2019). How does online employee ratings affect external firm financing? evidence from glassdoor. *Evidence from Glassdoor (December 16, 2019).*

Christensen, D., G. Serafeim, and A. Sikochi (2019). Why is corporate virtue in the eye of the beholder? the case of esg ratings. Technical report, Working paper.

Cohen, L., U. G. Gurun, and Q. H. Nguyen (2020). The esg-innovation disconnect: Evidence from green patenting. Technical report, National Bureau of Economic Research.

de Groot, W., J. de Koning, and S. van Winkel (2021). Sustainable voting behavior of asset managers: Do they walk the walk? *Available at SSRN 3783454*.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

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

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

Firth, J. R. (1957). Applications of general linguistics. *Transactions of the Philological Society 56*(1), 1–14.

Gibson, R., P. Krueger, and P. S. Schmidt (2019). Esg rating disagreement and stock returns. *Swiss Finance Institute Research Paper* (19-67).

Green, T. C., R. Huang, Q. Wen, and D. Zhou (2019). Crowdsourced employer reviews and stock returns. *Journal of Financial Economics 134*(1), 236–251.

Harris, Z. S. (1954). Distributional structure. *Word 10*(2-3), 146–162.

Haščič, I. and M. Migotto (2015). Measuring environmental innovation using patent data.

Hassan, T. A., S. Hollander, L. van Lent, and A. Tahoun (2019). Firm-level political risk: Measurement and effects. *The Quarterly Journal of Economics 134*(4), 2135–2202.

Hoberg, G. and V. Maksimovic (2015). Redefining financial constraints: A text-based analysis. *The Review of Financial Studies 28*(5), 1312–1352.

Hong, H. and L. Kostovetsky (2012). Red and blue investing: Values and finance. *Journal of Financial Economics 103*(1), 1–19.

Jha, M., H. Liu, and A. Manela (2020). Does finance benefit society? a language embedding approach. *A Language Embedding Approach (July 18, 2020)*.

Jha, M., H. Liu, and A. Manela (2021). Natural disaster effects on popular sentiment toward finance. *Available at SSRN 3833110*.

Joulin, A., E. Grave, P. Bojanowski, and T. Mikolov (2016). Bag of tricks for efficient text classification. *arXiv preprint arXiv:1607.01759*.

Khan, M., G. Serafeim, and A. Yoon (2016). Corporate sustainability: First evidence on materiality. *The accounting review 91*(6), 1697–1724.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics 132*(2), 665–712.

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of financial economics 115*(2), 304–329.

Liang, H. and L. Renneboog (2017). On the foundations of corporate social responsibility. *The Journal of Finance 72*(2), 853–910.

Lins, K. V., H. Servaes, and A. Tamayo (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *the Journal of Finance 72*(4), 1785–1824.

Liu, Y., M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of finance 66*(1), 35–65.

Loughran, T. and B. McDonald (2016). Textual analysis in accounting and finance: A survey. *Journal of Accounting Research 54*(4), 1187–1230.

Mikolov, T., I. Sutskever, K. Chen, G. S. Corrado, and J. Dean (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pp. 3111–3119.

Minaee, S., N. Kalchbrenner, E. Cambria, N. Nikzad, M. Chenaghlu, and J. Gao (2021). Deep learning–based text classification: A comprehensive review. *ACM Computing Surveys (CSUR) 54*(3), 1–40.

Pennington, J., R. Socher, and C. D. Manning (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pp. 1532–1543.

Radford, A., K. Narasimhan, T. Salimans, and I. Sutskever (2018). Improving language understanding with unsupervised learning. *Technical report, OpenAI*.

Shapiro, J. S. and R. Walker (2018). Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review 108*(12), 3814–54.

Socher, R., A. Perelygin, J. Wu, J. Chuang, C. D. Manning, A. Y. Ng, and C. Potts (2013). Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, pp. 1631–1642.

Xu, Q. and T. Kim (2020). Financial constraints and corporate environmental policies. *Available at SSRN 3028768*.

Yang, Z., Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le (2019). Xlnet: Generalized autoregressive pretraining for language understanding. In *Advances in neural information processing systems*, pp. 5753–5763.

Table A1: Description of Key Variables

This table reports variable definitions. Data sources include the Environmental Protection Agency's Toxic Releases Inventory (EPA-TRI), Environmental Protection Agency's Pollution Prevent (EPA-P2) database, and Glassdoor reviews (GD), Centre for Research in Security Prices (CRSP), Transcripts: Seeking Alpha (TSA), Kogan, Papanikolaou, Seru, and Stoffman (2017) (KPSS).

Variable	Description	Source
Dependent variables		
$\mathbb{1}$*(Abatement-Operating)*	Dummy variable indicating one if the firm undertakes operations-related abatement activities across any of its plants in the year, and zero otherwise	EPA-P2
$\mathbb{1}$*(Abatement-Process)*	Dummy variable indicating one if the firm undertakes process-related abatement activities across any of its plants in the year, and zero otherwise	EPA-P2
Pollution	Firm-level pollution obtained by aggregating emissions corresponding to various chemicals across all plants of a firm in a given year, for a given medium (air, water, ground or total)	EPA-TRI
#Patents$[t,t+x]$	The number of green patents granted to the firm in the $[t,t+x]$ window (measured in years) after the earnings call occurring on date t. Using information of the Cooperative Patent Classification from the USPTO, we identify the green patents following the guidelines from OECD.	KPSS
Overall Rating	The average employee overall ratings in a given calendar year-quarter.	GD
CAR[−3,+3]	Seven-day cumulative abnormal return centered on the earnings conference call date, calculated using the market model.	CRSP
Key independent variables		
Level(E+S)	The total frequency of environmental and social phrases in the earnings conference call scaled by the total number of words in the earnings call.	TSA
Level(E)	The total frequency of environmental phrases in the earnings conference call scaled by the total number of words in the earnings call.	TSA

Variable	Description	Source
Level(S)	The total frequency of social phrases in the earnings conference call scaled by the total number of words in the earnings call.	TSA
Level(E_Material)	The total frequency of environmental phrases which are material for the firm's industry in its earnings conference call scaled by the total number of words in the earnings call.	TSA
Level(E_Immaterial)	The total frequency of environmental phrases which are immaterial for the firm's industry in its earnings conference call scaled by the total number of words in the earnings call.	TSA
Level(S_Material)	The total frequency of social phrases which are material for the firm's industry in its earnings conference call scaled by the total number of words in the earnings call.	TSA
Level(S_Immaterial)	The total frequency of social phrases which are immaterial for the firm's industry in its earnings conference call scaled by the total number of words in the earnings call.	TSA
Firm-level control variables		
Earnings surprise	Actual earnings per share (EPS) from IBES minus the consensus (median) of EPS forecasts issued or reviewed in 90 days before the earnings announcement date. The difference is scaled by stock price at the end of quarter.	IBES
HighUE	Dummy variable which equals to 1 if the earnings surprise is in the highest decile in a given quarter, 0 otherwise.	IBES
LowUE	Dummy variable which equals to 1 if the earnings surprise is in the lowest decile in a given quarter, 0 otherwise.	IBES

Variable	Description	Source
Negative EPS	Dummy variable which equals 1 if announced earnings per share (EPS) is negative, 0 otherwise.	IBES
Pre-event return	Average stock return in window [−71,−11] in terms of trading days relative to the earnings conference call date.	CRSP
Pre-event volume	Average trading volume in window [−71,−11] in terms of trading days relative to the earnings conference call date.	CRSP
ROA	Earnings before extraordinary items divided by total assets.	COMPUSTAT
Accruals	Earnings minus cash flows from operations scaled by total assets.	COMPUSTAT
Size	Natural logarithm of the market cap at the end of quarter.	COMPUSTAT
MTB	Market cap plus book value of liabilities scaled by total assets at the end of quarter.	COMPUSTAT
Leverage	Long-term debt divided by total assets.	COMPUSTAT
NetPPEA	Net property, plant and equipment scaled by total assets	COMPUSTAT
Earnings volatility	Standard deviation of earnings-to-assets ratio in the past 20 quarters, with a minimum of 8 quarters required.	COMPUSTAT
Return volatility	Standard deviation of monthly stock returns in the last 12 months, with a minimum of 6 months required.	CRSP
# Analysts (log)	Natural logarithm of the number of analysts whose forecasts were issued or reviewed in the 90 days before the earnings announcement date.	IBES
Firm age (log)	Natural logarithm of the number of years since the stock's first date in CRSP.	CRSP

Variable	Description	Source
PreROA	Earnings before extraordinary items change of current quarter (q) over one year ago ($q-4$), scaled by total assets of one year ago ($q-4$).	COMPUSTAT
PreSale	Sales change of current quarter (q) over one year ago ($q-4$), scaled by total assets of one year ago ($q-4$).	COMPUSTAT
Call Transcript Controls		
Uncertainty	Percentage of uncertain words in earnings call transcript based on Loughran and McDonald (2011) dictionary and the code from Bill McDonald's website.	TSA
Sentiment	Percentage of positive words minus percentage of negative words in earnings call transcript based on Loughran and McDonald (2011) dictionary and the code from Bill McDonald's website.	TSA

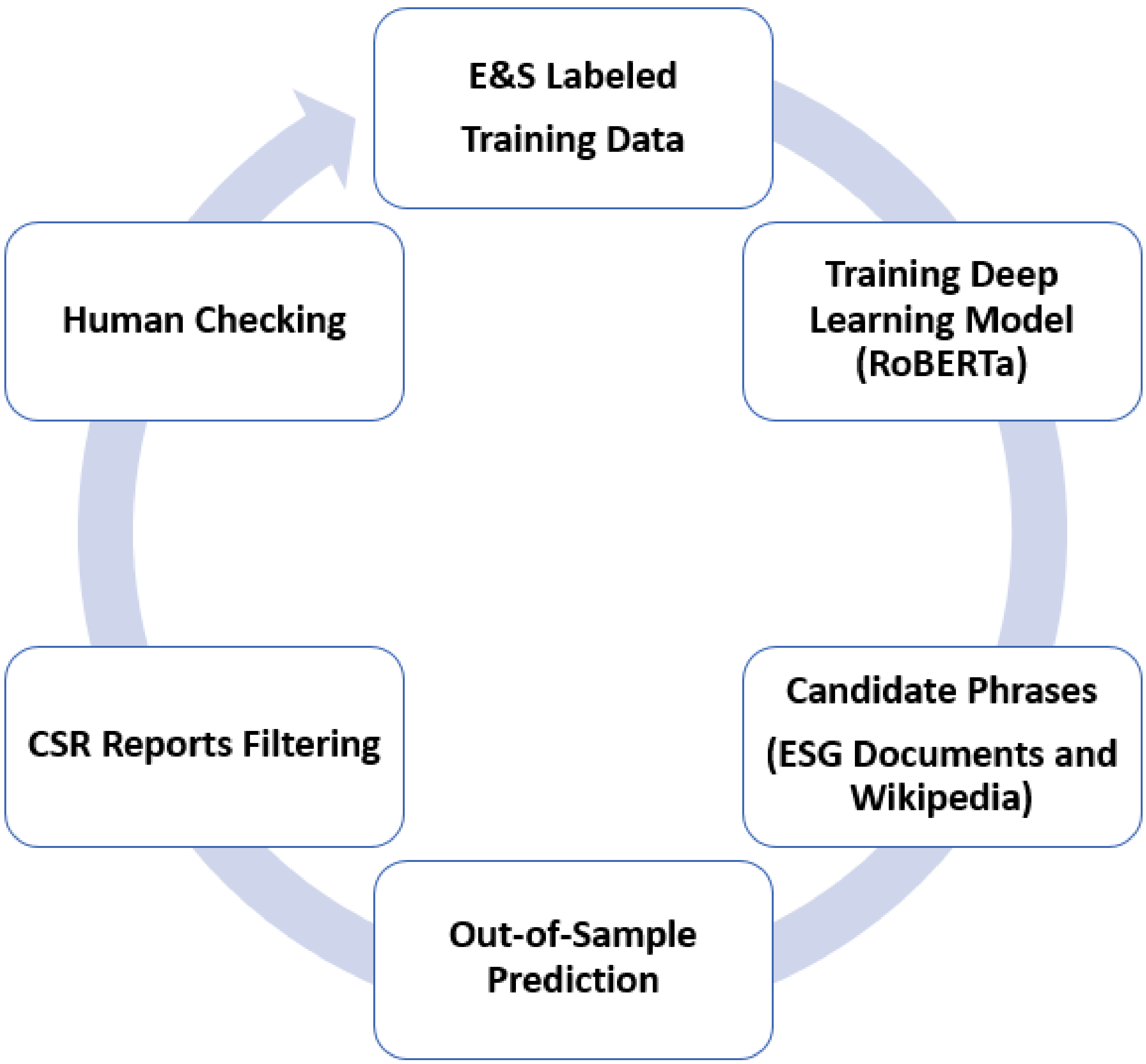

Figure 1: Work Flow of Dictionary Construction: This figure provides a high-level overview of the steps we take to construct our dictionary of environmental and social (E&S) topics. We provide detailed descriptions of these aspects in Section 2.

		Consumer Goods	Extractives & Minerals Processing	Financials
Dimension	General Issue Category	Click to expand	Click to expand	Click to expand
Environment	GHG Emissions			
	Air Quality			
	Energy Management			
	Water & Wastewater Management			
	Waste & Hazardous Materials Management			
	Ecological Impacts			
Social Capital	Human Rights & Community Relations			
	Customer Privacy			
	Data Security			
	Access & Affordability			
	Product Quality & Safety			
	Customer Welfare			
	Selling Practices & Product Labeling			

Figure 2: Sustainability Framework Example – SASB: This figure shows an example from SASB, one of the sustainability frameworks used in the construction of our dictionary. For each E&S category, the institution provides the detailed description and its corresponding dimension. We construct the E&S labeled data from the sustainability documents to train our deep learning model.

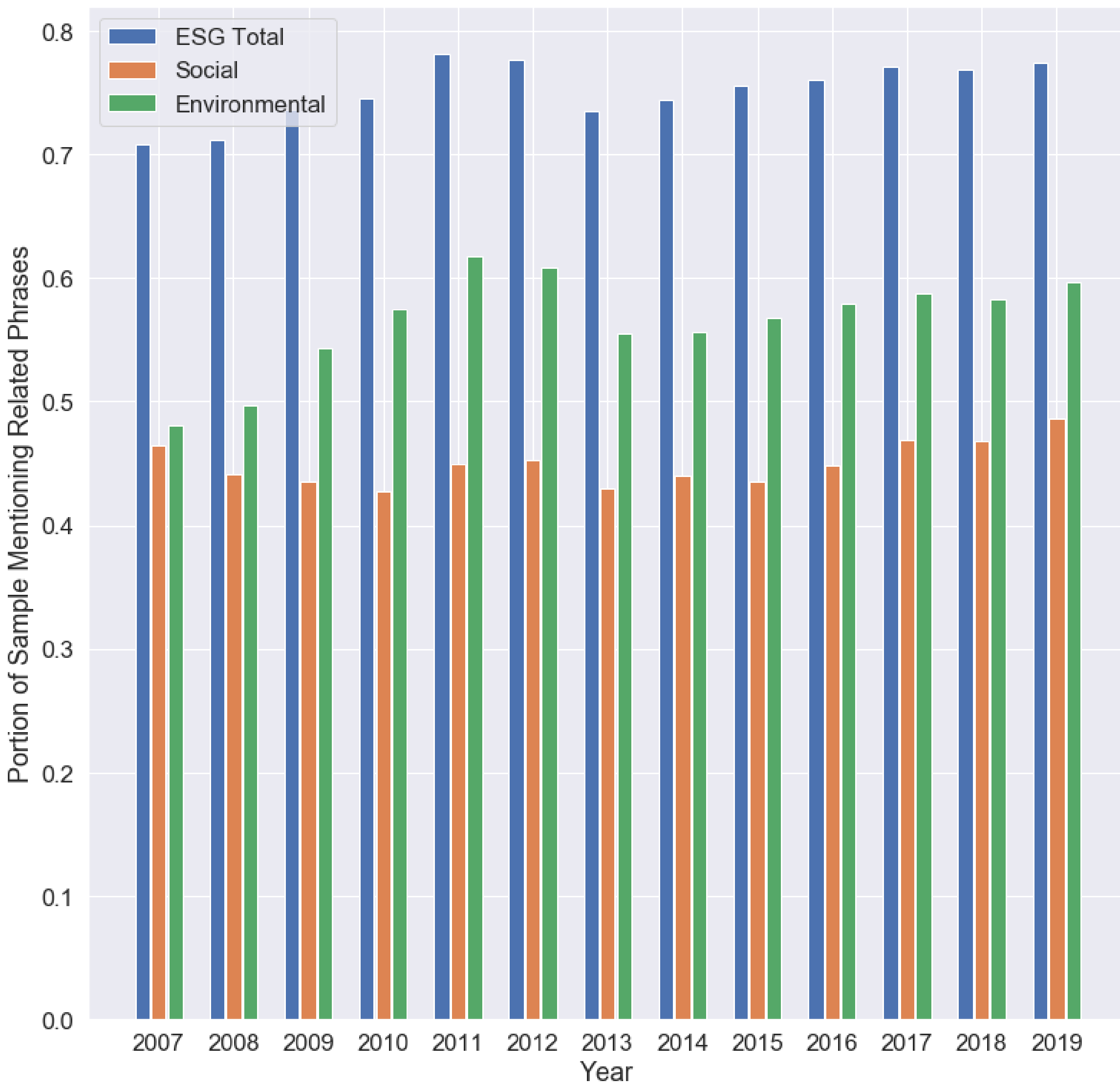

Figure 3: Distribution of Transcripts with E&S Phrases by Year: This figure shows the annual percentage of earnings calls mentioning E&S phrases during 2007-2019. The figure also separately documents the annual percentage of earnings calls mentioning environmental and social phrases separately.

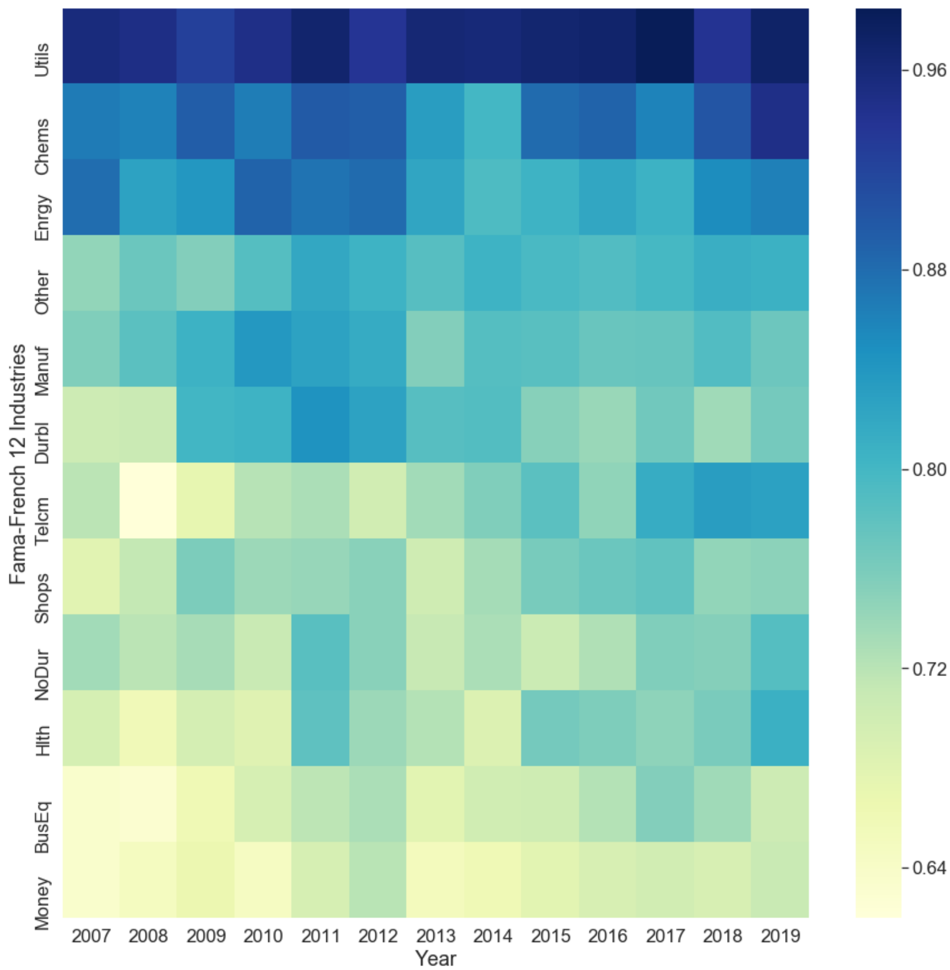

Figure 4: Distribution of Transcripts with E&S Phrases by Industry and Year: This figure shows the percentage of earnings calls mentioning E&S phrases across industries and years. The y-axis shows the abbreviations of Fama-French 12 industries sorted by the cumulative percentage of sample mentioning E&S phrases. We show the industry distribution as per Fama-French 12 industries classification due to space constraints. The x-axis represents different years of the sample. The density bar on the right of the graph displays the percentage of earnings calls mentioning E&S phrases for each industry between 2007 and 2019.

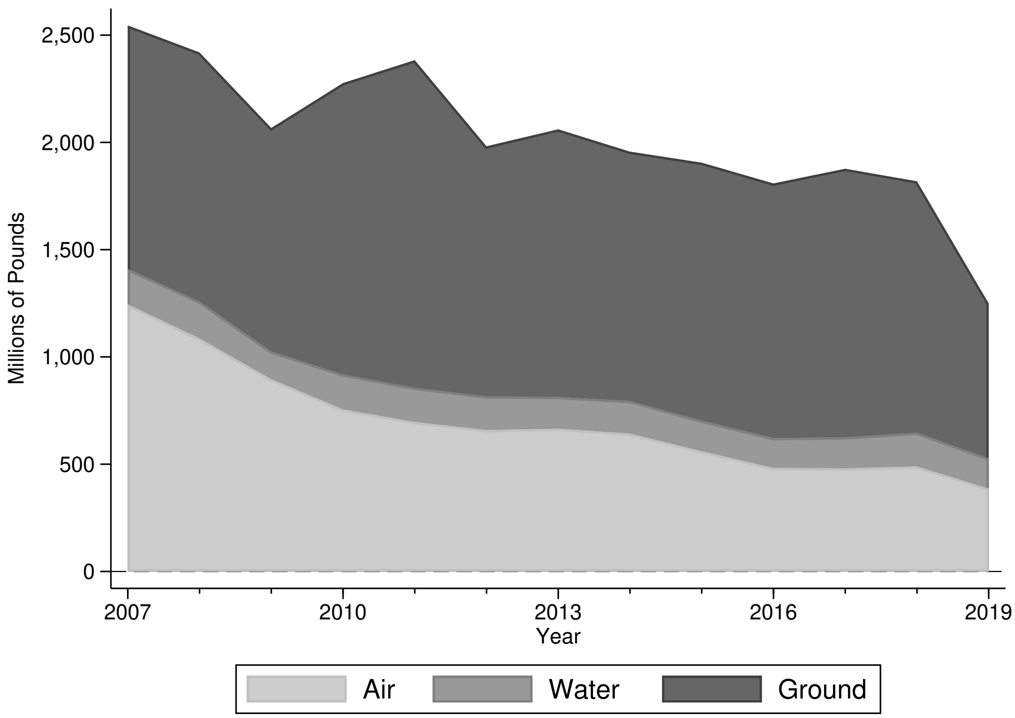

Figure 5: Total Toxic Releases: This figure shows the aggregate pollution (in millions of pounds) across different media of the environment for the firms in our sample during 2007-2019, as reported to the EPA's Toxic Release Inventory. The sample excludes firms that were not required to report emissions during the entire period.

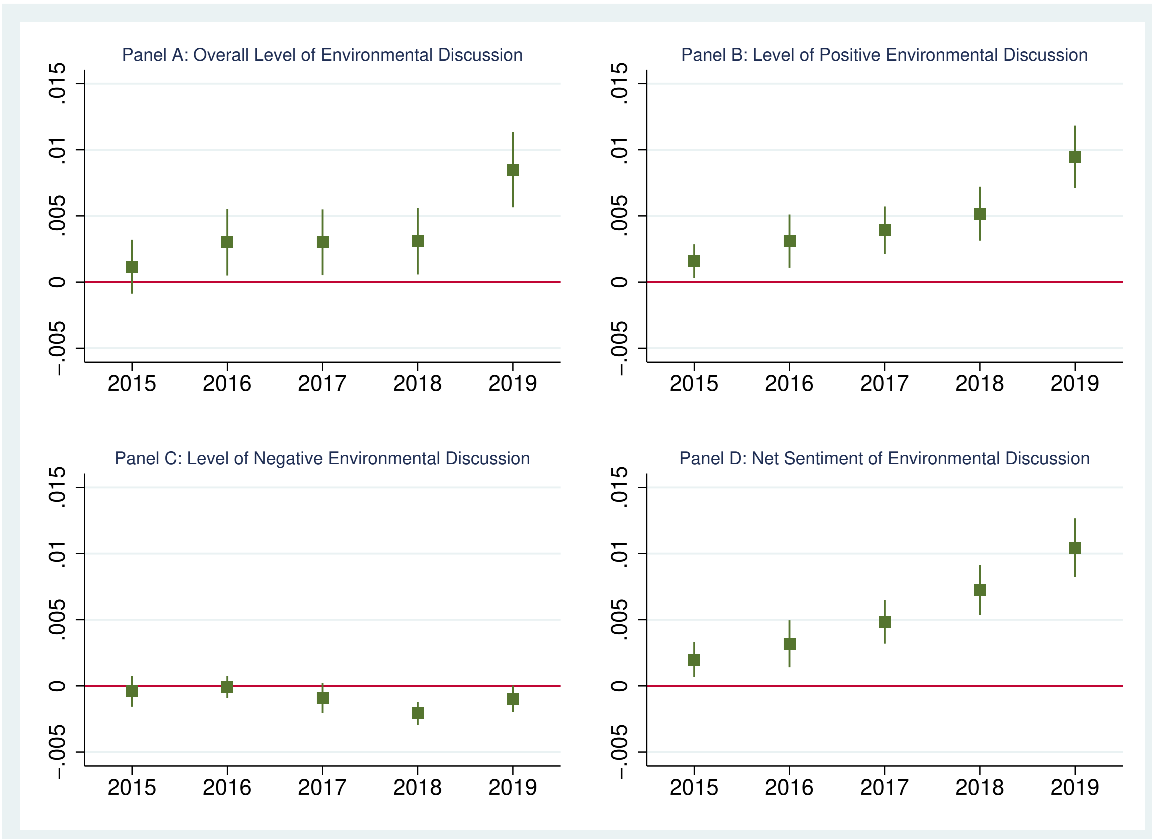

Figure 6: Paris Agreement and Environmental Discussion: This figure shows the dynamic trends of the level of environmental discussion in Panel A after the Paris Agreement on December 12, 2015. Panels B, C and D show the coefficients using different tones of the sentiment as the dependent variables. The y-axis represents the estimated coefficients obtained from regressing the level of various kinds of environmental discussion (%) in the earnings conference call transcripts over yearly dummies, after including firm and industry fixed effects. The benchmark period comprises the years before 2015 in our sample. Standard errors are double clustered by firm and year-quarter. The verticals lines represent 95% confidence intervals.

Table 1: Summary Statistics

This table reports the summary statistics for key variables in our analysis during the sample period. Panel A shows the transcript-level frequency of E&S phrases in earnings call transcripts in the sample of market reaction analysis during 2007-2019. In Panel B, we report the distribution of annual firm pollution from the Toxic Releases Inventory database of the EPA during 2007-2018. Panel C shows the number of green patents granted in the 1-year window after the earnings call date in the sample used in green innovation analysis. Panel D provides the statistics for overall employee rating, as obtained from Glassdoor reviews for the sample period of 2008-2019.

Panel A: Transcript-Level E&S Discussion

	Count	Mean	Median	Std. Dev.
Freq(E+S)	81,662	3.959	2.000	6.962
Freq(E)	81,662	2.662	1.000	6.184
Freq(S)	81,662	1.297	0.000	2.938
Freq(E-Material)	81,662	0.900	0.000	2.857
Freq(E-Immaterial)	81,662	1.762	0.000	4.743
Freq(S-Material)	81,662	0.613	0.000	1.947
Freq(S-Immaterial)	81,662	0.684	0.000	1.942

Panel B: Annual Firm Pollution (in pounds)

	Count	Mean	Median	Std. Dev.
Total Pollution	9,103	2,583,485	30,140	11,636,114
Air Pollution	9,103	883,111	23,242	3,889,917
Water Pollution	9,103	193,763	0	1,270,401
Ground Pollution	9,103	1,506,614	0	10,222,305
$\mathbb{1}$ *(Abatement-Operating)*	9,081	0.24	0	0.42
$\mathbb{1}$ *(Abatement-Process)*	9,081	0.23	0	0.42

Panel C: Number of Green Patents Granted

	Count	Mean	Median	Std. Dev.
Overall	32,074	2.236	0.000	8.597
Env&Water	32,074	0.465	0.000	2.159
Climate	32,074	1.850	0.000	7.073

Panel D: Quarterly Glassdoor Employee Reviews

	Count	Mean	Median	Std. Dev.
Overall Rating	41,387	3.15	3.17	0.75
Career Opp.	41,307	3.03	3.00	0.71
Comp.+Benefits	41,305	3.16	3.18	0.76
Senior Mgmt.	32,548	2.78	2.79	0.76
Work-life Balance	41,317	3.24	3.25	0.75
Culture	41,311	3.08	3.07	0.80

Table 2: Abatement of Pollution

This table reports the results from Equation (1) using EPA Pollution Prevention data on firm pollution abatement during 2007-2019. The dependent variables are indicators for whether a firm undertakes operations- or process-related abatement activities across any of its plants in the year. *Level(E+S)* corresponds to the fraction of words devoted to discussion of environmental and social topics in the earnings conference calls during the year. Whereas, *Level(E)* denotes the fraction of words devoted to discussion of environmental topics in the earnings conference calls during the year. *Level(E-Material)* denotes the fraction of words devoted to the discussion of environmental topics material to the industry in the earnings conference calls during the year. Similarly, *Level(E-Immaterial)* corresponds to the discussion of remaining environmental topics. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent Variable:	$\mathbb{1}$*(Abatement-Operating)*			$\mathbb{1}$*(Abatement-Process)*		
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	8.49**			8.15		
	[2.36]			[1.24]		
Level(E)		9.47**			7.50	
		[2.34]			[1.05]	
Level(E-Material)			31.51**			30.15*
			[2.39]			[1.92]
Level(E-Immaterial)			2.00			-0.17
			[0.55]			[-0.02]
Log(Total Revenue) $_{t-1}$	0.02	0.02	0.02	0.00	0.00	0.00
	[0.92]	[0.92]	[0.91]	[0.22]	[0.23]	[0.22]
Leverage $_{t-1}$	-0.01	-0.01	-0.00	-0.10**	-0.10**	-0.10**
	[-0.13]	[-0.13]	[-0.11]	[-2.70]	[-2.70]	[-2.69]
Log(Assets) $_{t-1}$	-0.02	-0.02	-0.02	0.01	0.01	0.01
	[-1.27]	[-1.27]	[-1.30]	[0.47]	[0.46]	[0.44]
NetPPEA $_{t-1}$	0.09	0.09	0.09	0.05	0.05	0.05
	[1.12]	[1.12]	[1.13]	[0.64]	[0.63]	[0.65]
Firm FE	✓	✓	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Adj.-R^2	0.452	0.452	0.452	0.387	0.387	0.387
Obs.	9,035	9,035	9,035	9,035	9,035	9,035

Table 3: Firm Pollution and Discussion on Environmental and Social Issues

This table reports the results from Equation (1) using yearly firm pollution from the EPA Toxic Release Inventory (TRI) as the dependent variable during 2007-2019. *Level(E+S)* corresponds to the fraction of words devoted to discussion of environmental and social topics in the earnings conference calls during the year. *Level(E)* denotes the fraction of words devoted to discussion of environmental topics in the earnings conference calls during the year. In Panel A, we show regression results using *Level(E+S)* as the explanatory variable. Panel B shows results using *Level(E)* as the independent variable. In Panel C, we show results using forward years of total pollution as the dependent variable. Panel D reports results using forward years of air pollution as the dependent variable. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A: Using Environmental & Social Discussion

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)					
	Total			Air	Water	Ground
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	-81.02***	-61.14***	-54.23***	-59.83**	-9.53	20.00
	[-3.63]	[-3.07]	[-3.10]	[-3.04]	[-0.80]	[0.47]
Log(Total Revenue) $_{t-1}$		-0.74***	-0.72***	-0.73***	-0.39***	-0.30***
		[-9.92]	[-9.31]	[-8.59]	[-3.38]	[-4.07]
Leverage $_{t-1}$		0.09	0.04	-0.05	0.00	-0.20*
		[0.61]	[0.32]	[-0.36]	[0.03]	[-1.96]
Log(Assets) $_{t-1}$		0.15	0.12	0.21*	0.15	0.15**
		[1.65]	[1.32]	[2.12]	[1.68]	[2.63]
NetPPEA $_{t-1}$		0.08	0.12	0.15	0.29	0.00
		[0.31]	[0.46]	[0.50]	[1.15]	[0.01]
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓				
Controls		✓	✓	✓	✓	✓
Industry-Year FE			✓	✓	✓	✓
Adj.-R^2	0.936	0.940	0.942	0.928	0.946	0.957
Obs.	9,001	9,001	8,967	8,967	8,967	8,967

Panel B: Using Only Environmental Discussion

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)					
	Total			Air	Water	Ground
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E)	-76.48***	-55.98**	-50.63**	-58.11**	-8.71	19.90
	[-3.11]	[-2.54]	[-2.57]	[-2.60]	[-0.69]	[0.42]
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓				
Controls		✓	✓	✓	✓	✓
Industry-Year FE			✓	✓	✓	✓
Adj.-R^2	0.936	0.940	0.942	0.928	0.946	0.957
Obs.	9,001	9,001	8,967	8,967	8,967	8,967

Panel C: Using Forward Years of Total Pollution

Dependent Variable:	Log (1+Poll.$_{t+1}$/COGS$_{t-1}$)	Log (1+Poll.$_{t+2}$/COGS$_{t-1}$)	Log (1+Poll.$_{t+3}$/COGS$_{t-1}$)
	(1)	(2)	(3)
Level(E)	-50.85**	-33.95	-9.55
	[-2.84]	[-1.78]	[-0.51]
Firm FE	✓	✓	✓
Industry-Year FE	✓	✓	✓
Controls	✓	✓	✓
Adj.-R^2	0.946	0.950	0.953
Obs.	8,343	7,696	7,035

Panel D: Using Forward Years of Air Pollution

Dependent Variable:	Log (1+Air$_{t+1}$/COGS$_{t-1}$)	Log (1+Air$_{t+2}$/COGS$_{t-1}$)	Log (1+Air$_{t+3}$/COGS$_{t-1}$)
	(1)	(2)	(3)
Level(E)	-57.93**	-40.61*	-17.13
	[-2.75]	[-2.17]	[-0.93]
Firm FE	✓	✓	✓
Year FE			
Industry-Year FE	✓	✓	✓
Controls	✓	✓	✓
Adj.-R^2	0.933	0.938	0.940
Obs.	8,343	7,696	7,035

Table 4: Pollution Based on Industry-Specific Material Discussion

This table reports the results from Equation (1) using yearly firm pollution from the EPA TRI as the dependent variable during 2007-2019. *Level(E-Material)* denotes the fraction of words devoted to the discussion of environmental topics material to the industry in the earnings conference calls during the year. Similarly, *Level(E-Immaterial)* corresponds to the discussion of remaining environmental topics. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)			
	Total	Air	Water	Ground
	(1)	(2)	(3)	(4)
Level(E-Material)	-80.60*	-36.57	-34.84	-37.12
	[-1.79]	[-0.84]	[-1.18]	[-0.72]
Level(E-Immaterial)	-40.68	-65.26**	-0.04	38.83
	[-1.66]	[-2.28]	[-0.00]	[0.60]
Firm FE	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Adj.-R^2	0.942	0.928	0.946	0.957
Obs.	8,967	8,967	8,967	8,967

Table 5: Pollution and Paris Agreement

This table reports the results based on a modified version of Equation (1) using yearly firm pollution from the EPA TRI as the dependent variable during 2007-2019. We interact the equation with dummies corresponding to periods around the inclusion of the US into the Paris Agreement. We additionally control for group-year fixed effects. *Level(E)* denotes the fraction of words devoted to the discussion of environmental topics. The dummy *Before Paris* corresponds to the years before the Paris Agreement up to 2015. *During Paris* refers to years during which the United States was part of the Paris Agreement in 2016 and 2017. Finally, *Paris Withdrawal* indicates years after the withdrawal from the Paris Agreement during 2018 and 2019. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)			
	Total (1)	Air (2)	Water (3)	Ground (4)
Before Paris × Level(E)	-32.88*	-36.36*	-1.90	48.16
	[-1.90]	[-1.87]	[-0.15]	[0.89]
During Paris × Level(E)	-63.81*	-85.95**	-7.59	1.94
	[-2.06]	[-2.25]	[-0.32]	[0.05]
Paris Withdrawal × Level(E)	-96.92**	-104.23**	-32.06	-56.50
	[-2.84]	[-2.86]	[-1.30]	[-1.51]
Diff(*Before- Withdrawal*)	64.05	67.87	30.16	104.66
P-value	0.06	0.05	0.24	0.03
Diff(*During- Withdrawal*)	33.11	18.28	24.47	58.44
P-value	0.05	0.03	0.24	0.02
Firm FE	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Group-Year FE	✓	✓	✓	✓
Adj.-R^2	0.942	0.928	0.946	0.957
Obs.	8,967	8,967	8,967	8,967

Table 6: Discussion of E&S Topics and Future Green Patenting Activity

This table presents results exploring whether the greater discussion of environmental and social topics in earnings calls is associated with the ex post granting of a higher number of green patents. The regressions are estimated using a Poisson model. The dependent variable in Columns (1) and (2) is the number of overall green patents granted after the earnings calls ($Overall$). The dependent variable in Columns (3) and (4) is the the number of green patents which are related to environmental management and water-related adaptation technologies ($Env\&Water$). The dependent variable in Columns (5) and (6) is the the number of green patents about climate change mitigation ($Climate$). In Panel A, we use the green patents which are granted during the one-year window after the earnings call date. Similarly, Panel B (Panel C) shows the number of green patents which are granted during the two-year (three-year) window after the earnings call date. The number of green patents and control variables are winsorized at 1% and 99%. All variables are defined in Table A1. The Z-statistics (in brackets) are based on double clustering standard errors at the firm and year-quarter levels, unless otherwise specified. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: 1-Year Window

Dependent Variable:	Overall		Env&Water		Climate	
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	168.0***		319.6***		143.4**	
	[2.82]		[3.89]		[2.45]	
Level(E)		193.8***		366.8***		169.0***
		[3.05]		[4.52]		[2.74]
Industry × Year-Qtr. FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Pseudo R^2	0.587	0.588	0.531	0.534	0.571	0.572
Obs.	32,074	32,074	29,310	29,310	31,884	31,884

Panel B: 2-Year Window

Dependent Variable:	Overall		Env&Water		Climate	
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	170.8***		319.2***		150.1**	
	[2.81]		[3.82]		[2.49]	
Level(E)		195.8***		367.6***		173.9***
		[3.04]		[4.47]		[2.76]
Industry × Year-Qtr. FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Pseudo R^2	0.619	0.619	0.565	0.568	0.609	0.610
Obs.	32,593	32,593	30,642	30,642	32,517	32,517

Panel C: 3-Year Window

Dependent Variable:	Overall		Env&Water		Climate	
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	169.3***		313.0***		151.5**	
	[2.71]		[3.63]		[2.43]	
Level(E)		194.6***		362.3***		175.1***
		[2.95]		[4.28]		[2.69]
Industry × Year-Qtr. FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Pseudo R^2	0.628	0.629	0.584	0.586	0.619	0.620
Obs.	32,686	32,686	31,505	31,505	32,644	32,644

Table 7: Employee Ratings and Discussion on Environmental and Social Topics

This table reports the results from Equation (1) using overall quarterly ratings from Glassdoor as the dependent variable during 2007-2019. *Level(E+S)* corresponds to the fraction of words devoted to discussion of environmental and social topics in the earnings conference calls during the year-quarter. Whereas, *Level(S)* denotes the fraction of words devoted to discussion of social topics in the earnings conference calls during the year-quarter. *Level(S-Material)* denotes the fraction of words devoted to the discussion of social topics material to the industry in the earnings conference calls during the year-quarter. Similarly, *Level(S-Immaterial)* corresponds to the discussion of remaining social topics. Rating_{t+i} denotes the overall rating in the i^{th} year-quarter ahead. In Panel A, we show regression results using *Level(E+S)* as the main explanatory variable. Panel B shows results using *Level(S)* as the main independent variable. In Panel C, we analyze different components of the overall rating. Panel D provides regression results based on the industry-specific material discussion across different components of overall rating. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year-quarter levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A: Using Environmental & Social Discussion

Dependent Variable:	Overall Rating					
	Rating_t			Rating_{t+1}	Rating_{t+2}	Rating_{t+3}
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	25.26***	24.76***	24.60***	17.66**	6.28	-3.11
	[2.84]	[2.80]	[2.88]	[2.05]	[0.82]	[-0.33]
$\text{Log(Assets}_{t-1})$		0.03	0.01	0.01	0.00	0.01
		[1.44]	[0.78]	[0.72]	[0.24]	[0.54]
Log(MTB Ratio)		0.10***	0.10***	0.10***	0.07***	0.06***
		[4.87]	[4.58]	[4.49]	[3.98]	[2.69]
Leverage		-0.05	-0.05	-0.09**	-0.07	-0.01
		[-1.41]	[-1.36]	[-2.12]	[-1.63]	[-0.38]
Return on Assets		0.47**	0.42	0.81***	0.46**	0.24
		[2.06]	[1.66]	[3.61]	[2.07]	[1.09]
Return_{t-1}		0.03	0.03	0.09***	0.04*	0.01
		[1.34]	[1.56]	[4.05]	[2.00]	[0.63]
Firm FE	✓	✓	✓	✓	✓	✓
Year-Qtr. FE	✓	✓				
Controls		✓	✓	✓	✓	✓
Industry × Year-Qtr. FE			✓	✓	✓	✓
Adj.-R^2	0.276	0.277	0.283	0.285	0.282	0.287
Obs.	41,267	41,267	41,199	41,139	40,884	40,401

Panel B: Using Only Social Discussion

Dependent Variable:	Overall Rating					
	$Rating_t$			$Rating_{t+1}$	$Rating_{t+2}$	$Rating_{t+3}$
	(1)	(2)	(3)	(4)	(5)	(6)
Level(S)	28.18*	27.87*	29.73*	15.79	9.89	-0.23
	[1.83]	[1.82]	[1.99]	[1.03]	[0.69]	[-0.02]
Firm FE	✓	✓	✓	✓	✓	✓
Year-Qtr. FE	✓	✓				
Controls		✓	✓	✓	✓	✓
Industry × Year-Qtr. FE			✓	✓	✓	✓
Adj.-R^2	0.276	0.277	0.283	0.285	0.282	0.287
Obs.	41,267	41,267	41,199	41,139	40,884	40,401

Panel C: Components of Rating

Dependent Variable:	Overall Rating	Career Opp.	Comp.+Benefits	Senior Mgmt.	Work-life Balance	Culture
	(1)	(2)	(3)	(4)	(5)	(6)
Level(S)	29.73*	22.43	24.56*	29.85	26.12**	31.67*
	[1.99]	[1.45]	[1.95]	[1.67]	[2.14]	[2.00]
Firm FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Industry × Year-Qtr. FE	✓	✓	✓	✓	✓	✓
Adj.-R^2	0.283	0.276	0.378	0.277	0.307	0.301
Obs.	41,199	41,118	41,116	32,403	41,128	41,122

Panel D: Based on Industry-Specific Material Discussion

Dependent Variable:	Overall Rating	Career Opp.	Comp.+Benefits	Senior Mgmt.	Work-life Balance	Culture
	(1)	(2)	(3)	(4)	(5)	(6)
Level(S-Material)	28.29	25.31	34.33**	37.85	39.29***	28.24
	[1.30]	[1.09]	[2.23]	[1.66]	[2.73]	[1.22]
Level(S-Immaterial)	31.61	18.68	11.84	18.80	8.96	36.13
	[1.41]	[0.90]	[0.59]	[0.72]	[0.45]	[1.38]
Firm FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Industry × Year-Qtr. FE	✓	✓	✓	✓	✓	✓
Adj.-R^2	0.283	0.276	0.378	0.277	0.307	0.301
Obs.	41,199	41,118	41,116	32,403	41,128	41,122

Table 8: Stock Price Response to E&S Discussion

This table presents results for the immediate stock price response to the E&S discussion in earnings calls. The dependent variable across all columns is CAR[−3,+3], calculated using the market model. The key independent variables are $Level(E+S)$, $Level(E)$ and $Level(S)$, which are computed as the frequency of environmental and social phrases, environmental phrases, and social phrases in an earnings call scaled by the total number of words in the earnings call. The market reaction variable and control variables are winsorized at 1% and 99%. All control variables are described in Table A1. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year-quarter levels, unless otherwise specified. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent Variable:	CAR[−3,+3]			
	(1)	(2)	(3)	(4)
Level(E+S)	1.10*			
	[1.72]			
Level(E)		1.56**		1.56**
		[2.35]		[2.37]
Level(S)			-0.27	-0.34
			[-0.21]	[-0.26]
Controls	✓	✓	✓	✓
Firm FE	✓	✓	✓	✓
Industry × Year-Qtr. FE	✓	✓	✓	✓
Adj.-R^2	0.143	0.143	0.143	0.143
Obs.	81,662	81,662	81,662	81,662

For Online Publication–Internet Appendix

Table IA1: Additional Summary Statistics

This table summarizes the descriptive statistics for independent variables used in the paper. Panel A corresponds to the fraction of E&S phrases (reported as %) and control variables related to the market reaction analysis. Panel B shows the distribution of the proportion of environmental and social phrases in our sample, aggregated up to the firm–year level. Panel C presents the number of green patents granted in the 2-year and 3-year windows after the earnings calls in the samples of the Poisson regression.

Panel A: Transcript-Level Variables

	Count	Mean	Median	Std. Dev.
Level(E+S)	81,662	0.051	0.024	0.089
Level(E)	81,662	0.035	0.011	0.081
Level(S)	81,662	0.016	0.000	0.036
Level(E-Material)	81,662	0.012	0.000	0.037
Level(E-Immaterial)	81,662	0.023	0.000	0.062
Level(S-Material)	81,662	0.008	0.000	0.024
Level(S-Immaterial)	81,662	0.009	0.000	0.024
Earnings surprise	81,662	0.000	0.001	0.012
Negative EPS	81,662	0.228	0.000	0.420
HighUE	81,662	0.098	0.000	0.298
LowUE	81,662	0.098	0.000	0.298
Pre-event return	81,662	0.000	0.001	0.003
Pre-event volume (Millions)	81,662	2.060	0.718	3.930
ROA	81,662	0.004	0.009	0.037
Size	81,662	7.765	7.750	1.766
Accruals	81,662	-0.014	-0.011	0.036
Earnings volatility	81,662	0.024	0.012	0.033
MTB	81,662	1.894	1.461	1.271
Leverage	81,662	0.230	0.199	0.205
Return volatility	81,662	0.104	0.088	0.061
# Analysts (log)	81,662	1.611	1.609	0.918
Firm age (log)	81,662	2.924	2.962	0.781
Uncertainty	81,662	1.018	0.996	0.248
Sentiment	81,662	0.580	0.584	0.598
Panel B: Firm-Year Aggregation				
Level(E+S)	39,388	0.056	0.029	0.093
Level(E)	39,388	0.037	0.013	0.085
Level(S)	39,388	0.018	0.008	0.038
Level(E-Material)	39,388	0.013	0.000	0.040
Level(E-Immaterial)	39,388	0.025	0.007	0.064
Level(S-Material)	39,388	0.008	0.000	0.024
Level(S-Immaterial)	39,388	0.010	0.003	0.026
Panel C: Number of Green Patents				
2-Year Window				
Overall	32,593	4.625	0.000	18.037
Env&Water	32,593	0.936	0.000	4.315
Climate	32,593	3.913	0.000	15.382
3-Year Window				
Overall	32,686	7.132	0.000	27.796
Env&Water	32,686	1.453	0.000	6.764
Climate	32,686	6.034	0.000	23.623

Table IA2: Robustness: Firm Pollution Using Alternative Scaling and E&S Discussion

This table reports the results from Equation (1) using yearly firm pollution as the dependent variable during 2007-2019. *Level(E+S)* corresponds to the fraction of environmental and social phrases discussed in earnings conference calls during the year. Similarly, *Level(E)* denotes the fraction of environmental phrases discussed in earnings conference calls during the year. Each column corresponds to a different variable used to scale the pollution emissions in the outcome variable. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A: Using Environmental & Social Discussion

Dependent Variable:	Log (1+Pollution/Scaling Variable$_{t-1}$)			
Scaling Variable:	COGS$_t$	Revenue$_{t-1}$	Sale$_{t-1}$	Total Assets$_{t-1}$
	(1)	(2)	(3)	(4)
Level(E+S)	-56.07***	-68.82***	-68.25***	-43.31***
	[-3.09]	[-3.85]	[-3.79]	[-3.21]
Firm FE	✓	✓	✓	✓
Industry-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Adj.-R^2	0.942	0.947	0.946	0.941
Obs.	8,971	9,046	8,968	9,054

Panel B: Using Only Environmental Discussion

Dependent Variable:	Log (1+Pollution/Scaling Variable$_{t-1}$)			
Scaling Variable:	COGS$_t$	Revenue$_{t-1}$	Sale$_{t-1}$	Total Assets$_{t-1}$
	(1)	(2)	(3)	(4)
Level(E)	-52.48**	-67.43***	-66.82***	-40.80**
	[-2.54]	[-3.29]	[-3.24]	[-2.76]
Firm FE	✓	✓	✓	✓
Year FE				
Industry-Year FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Adj.-R^2	0.942	0.946	0.946	0.941
Obs.	8,971	9,046	8,968	9,054

Table IA3: Robustness: Firm Pollution and E&S Discussion Using Matched Panel

This table reports the results from Equation (1) using yearly firm pollution from the EPA Toxic Release Inventory (TRI) as the dependent variable during 2007-2019 using a matched panel. We do not impute zero levels of environmental or social discussion to firms for which we do not find earnings conference call transcripts. *Level(E+S)* corresponds to the fraction of words devoted to discussion of environmental and social topics in the earnings conference calls during the year. *Level(E)* denotes the fraction of words devoted to discussion of environmental topics in the earnings conference calls during the year. In Panel A, we show regression results using *Level(E+S)* as the explanatory variable. Panel B shows results using *Level(E)* as the independent variable. In Panel C, we show results using *Level(E)* with at least one environmental word in the sample. Industry definition is based on Fama-French 30 classification. The T-statistics (in brackets) are based on double clustering standard errors at the firm and year levels, unless otherwise specified. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Panel A: Using Environmental & Social Discussion

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)					
	Total			Air	Water	Ground
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E+S)	-66.97**	-61.98**	-56.45**	-66.51**	-12.92	32.65
	[-2.76]	[-2.81]	[-2.63]	[-2.59]	[-0.86]	[0.52]
Log(Total Revenue) $_{t-1}$		-0.71***	-0.72***	-0.71***	-0.32***	-0.23**
		[-7.62]	[-7.57]	[-6.92]	[-3.87]	[-2.91]
Leverage $_{t-1}$		0.19	0.12	0.03	0.04	-0.10
		[1.10]	[0.69]	[0.18]	[0.54]	[-0.86]
Log(Assets) $_{t-1}$		0.16*	0.15*	0.26***	0.04	0.07
		[1.99]	[1.88]	[3.14]	[0.61]	[1.06]
NetPPEA $_{t-1}$		0.11	0.13	0.16	0.02	0.05
		[0.47]	[0.55]	[0.53]	[0.08]	[0.19]
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓				
Industry-Year FE			✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Adj.-R^2	0.942	0.945	0.947	0.933	0.953	0.959
Obs.	7,347	7,347	7,318	7,318	7,318	7,318

Panel B: Using Only Environmental Discussion

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)					
	Total			Air	Water	Ground
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E)	-59.37**	-54.09**	-51.49**	-64.87**	-13.42	33.05
	[-2.20]	[-2.23]	[-2.20]	[-2.36]	[-0.87]	[0.48]
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓				
Industry-Year FE			✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Adj.-R^2	0.942	0.945	0.947	0.933	0.953	0.959
Obs.	7,347	7,347	7,318	7,318	7,318	7,318

Panel C: At Least One Environmental Word

Dependent Variable:	Log (1+Pollution/COGS$_{t-1}$)					
	Total			Air	Water	Ground
	(1)	(2)	(3)	(4)	(5)	(6)
Level(E)	-60.00**	-56.16**	-53.09**	-66.73**	-14.24	32.37
	[-2.21]	[-2.33]	[-2.27]	[-2.40]	[-0.93]	[0.46]
Log(Total Revenue) $_{t-1}$		-0.73***	-0.76***	-0.75***	-0.36***	-0.27**
		[-8.08]	[-8.25]	[-7.30]	[-4.08]	[-2.98]
Leverage $_{t-1}$		0.13	0.02	-0.09	0.04	-0.14
		[0.82]	[0.15]	[-0.50]	[0.48]	[-1.18]
Log(Assets) $_{t-1}$		0.13*	0.15*	0.27***	0.04	0.08
		[1.80]	[1.98]	[3.28]	[0.57]	[1.04]
NetPPEA $_{t-1}$		0.11	0.08	0.17	-0.07	0.04
		[0.44]	[0.34]	[0.56]	[-0.26]	[0.12]
Firm FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓				
Industry-Year FE			✓	✓	✓	✓
Controls		✓	✓	✓	✓	✓
Adj.-R^2	0.945	0.948	0.949	0.935	0.953	0.960
Obs.	6,714	6,714	6,677	6,677	6,677	6,677

Table IA4: Labels in Training Data

This table shows the label names and the number of sequences under each label in the training data generated from the sustainability documents.

Label	Number of Sequences
Environmental	5,141
Social	6,921
Governance	2,291
Business Model Innovation	3,563
Economic	634

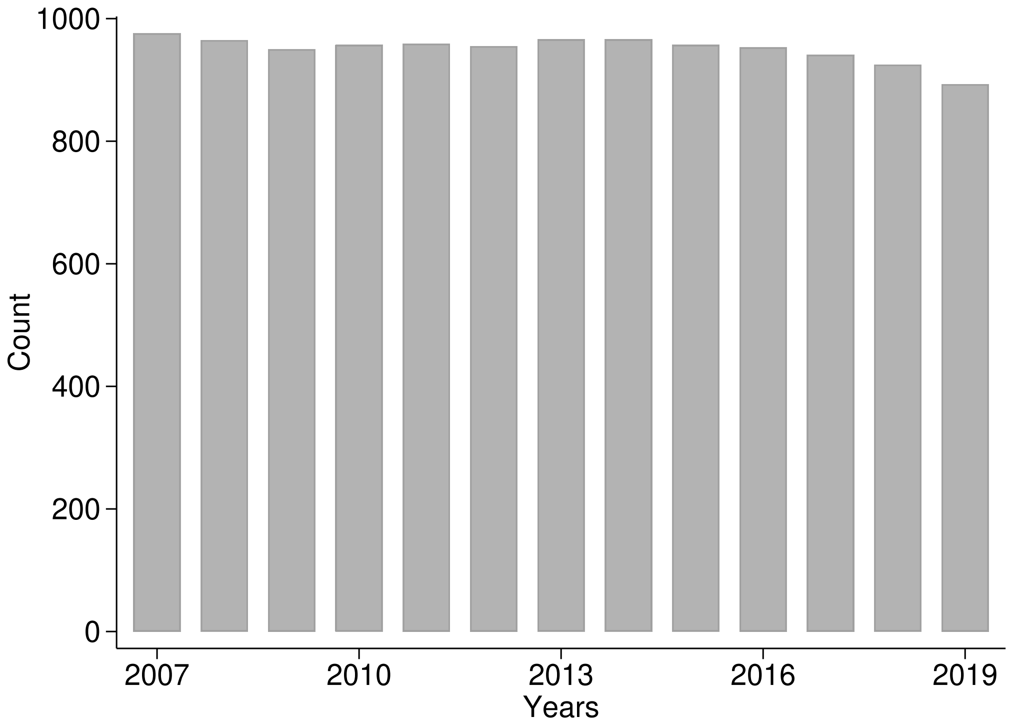

Figure IA1: Toxic Releases Inventory (TRI) - Number of Firms: The figure shows the the number of firms in our sample for which we obtain data on toxic releases from EPA's TRI.

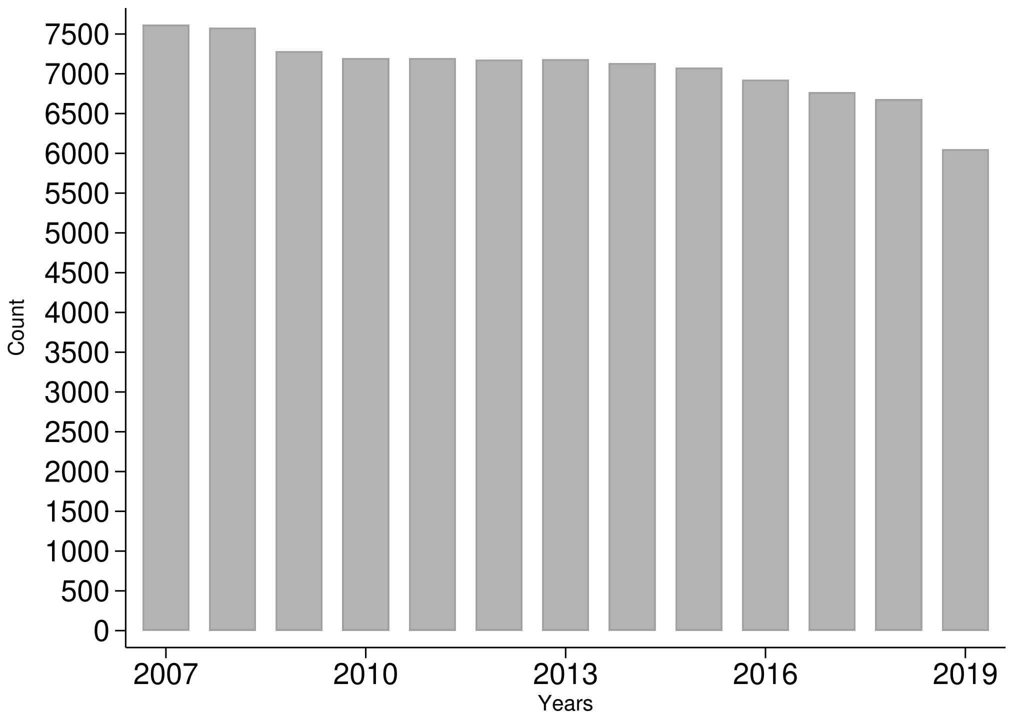

Figure IA2: Toxic Releases Inventory (TRI) - Number of Facilities: The figure shows the the number of facilities in our sample for which we obtain data on toxic releases from EPA's TRI.