Towards Objectivity and Interpretability: An Anatomy of Climate Change Exposure

Weiwei Hu, Kai Li, and Tingyu Yu *

August 31, 2023

Abstract

Traditional word-counting methods prove insufficient in evaluating intricate climate change exposures from texts. This paper leverages context-aware machine learning techniques to automatically detect and categorize climate-related sentences with minimal human bias, generating more objective and interpretable firm-level climate change exposure measures. We uncover novel mechanisms by distinguishing between risks and efforts within transition exposure, identifying diverse transition efforts in line with the emission mitigation hierarchy, and detecting physical exposures with different persistence. Firms reducing their own and customers' carbon emissions enjoy higher valuations and lower expected returns. Conversely, firms exposed to physical risks experience profit losses and increased equity costs.

JEL Codes: D83, G12, G14, Q54

Keywords: Climate change exposure, natural language processing, cross-section returns

First Draft: July 30, 2021

*Weiwei Hu (weiwei.hu@phbs.pku.edu.cn) is an assistant professor of economics at Peking University HSBC Business School. Kai Li (kaili825@gmail.com) is an associate professor of finance at Peking University HSBC Business School. Tingyu Yu (tingyu.yu@bf.uzh.ch) is a postdoc at University of Zurich. We are grateful for the helpful comments from Michal Dzielinski, Stefano Ramelli, Giorgio Valente, Chi-Yang Tsou, Markus Leippold, Zacharias Sautner, and other participants from AFA PhD Student Poster Session at the 2022 Annual Meeting, HKIMR Webinar, FutFinInfo Webinar of Stockholm Business School, The 19th Chinese Finance Annual Meeting 2022 and The Second China Finance Advancement Forum 2023. We acknowledge financial support from the General Research Fund of the Research Grants Council of Hong Kong (Project Number: 16506020) and Hong Kong Institute for Monetary and Financial Research. This paper represents the views of the author(s), which are not necessarily the views of the Hong Kong Monetary Authority, Hong Kong Institute for Monetary and Financial Research, or its Board of Directors or Council of Advisers. The above-mentioned entities except the author(s) take no responsibility for any inaccuracies or omissions contained in the paper.

1 Introduction

Assessing financial exposure to climate change is increasingly important but remains a complex task. Firms face multifaceted impacts from climate change, ranging from physical asset damage due to extreme weather events to substantial operational shifts necessitated by the transition to a low-carbon economy. The continually evolving responses of firms, shaped by the changing societal attitudes and technology developments, further complicate climate change influences on companies. Given these challenges, researchers are turning towards corporate disclosure texts as invaluable resources for a thorough exploration of climate change exposure measures at the firm level.

In this study, we utilize natural language processing techniques on earnings call transcripts to quantify firm-level exposure to climate change comprehensively. Current textual analysis research in finance largely depends on word-counting methodologies with predetermined topics; for example, two companion works, Sautner et al. (2023) and Li et al. (2023), investigate the same issue using the same dataset. Such methods prove insufficient in evaluating the intricate nature of climate change exposure due to the incorporation of human bias and the lack of contextual understanding. Our work distinguishes itself by employing advanced context-driven machine learning techniques (LDA, word2vec, and FinBERT) to formulate the anatomy of climate change exposure with two advantages – improved objectivity and enhanced economic interpretability. Our approach uncovers novel mechanisms associated with climate change that exhibit a stronger correlation with firm operations and asset pricing than prior studies. Table 1 summarizes the comparison between our method and those using keyword counting.

[Insert Table 1 here]

The first primary limitation of word-counting methods is their heavy reliance on the selection of topics and related keywords, which can introduce inherent human bias that skews interpretations. Investors possess wide-ranging concerns regarding a firm's exposure to climate change, which cannot always be captured within a preset framework of topics. The impracticality of anticipating all relevant words by humans also implies that manually determined keywords can inadvertently overlook critical terms for each topic, thereby further biasing the collected climate information.[1] To address this issue, our method maximizes the detection of climate-related discussions by applying LDA to automatically extract frequently mentioned words in IPCC reports as seed words and then employing word2vec to expand this

[1]Even the use of traditional machine learning techniques for keyword detection cannot fully solve this issue as their lack of contextual understanding often results in overdependence on initial seed words.

list by incorporating firm-specific concerns from earnings calls. This method largely reduces the dependence on human judgment and is not restricted to prescribed exposure types. We objectively collect a wide range of climate-related information and facilitate the discovery of more exposure sources than manual selection.

Another fundamental drawback of word-counting methods is their exclusive focus on individual climate keywords that neglect valuable contextual information surrounding them. This results in a vague categorization of climate exposure that cannot be clearly interpreted. Since identical climate words can convey contrasting information in different sentences, creating mutually exclusive word lists for each topic is infeasible. An illustration of this is the use of the phrase "carbon emissions" in the context of a company's efforts to decrease emissions or its concerns about regulation compliance, which exhibit opposing value implications but cannot be distinguished by word-level analyses. To overcome this limitation, we utilize LDA and FinBERT to extend our analysis from single words to entire sentences containing climate terms. Our method automatically classifies sentences into different categories based on the embedded topics and sentiments. This enables us to generate more interpretable categories: firstly, our transition exposure effectively differentiates the risks faced by firms and their efforts for climate mitigation; secondly, we decompose the transition effort exposure into sub-categories that closely align with the hierarchy of greenhouse gas emission mitigation; thirdly, our physical exposure is split into two types with different levels of persistence. These detailed categorizations carry different economic implications and could counteract each other when combined, which might explain the limited empirical findings regarding firm valuations and carbon reduction effects observed in previous research.

Prior literature mainly focuses on the risk side of transition exposure. However, our unsupervised method automatically discovers that firms often spend more time talking about their climate mitigation efforts. Differentiating the two has vital implications for asset pricing as they represent distinct channels through which climate change can impact stock returns. While transition effort exposure might be viewed positively by investors and lead to favorable market reactions, risk exposure tends to receive a negative response as it signals future cost and loss. Grouping them into a single measure could result in misleading or insignificant pricing effects.

Our transition effort exposure can be further decomposed into three categories. *Carbon removal* emphasizes solutions like carbon capture for managing generated carbon emissions. *Renewable energy* depicts firms' adoption of clean energy to cut emissions at the source and is often discussed by utility companies. *Enabling technology* assigns more weight to techniques that improve energy efficiency or promote electrification, ultimately aiming to decrease energy demand and substitute fossil fuels with low-carbon electricity. This granular breakdown is crucial since each effort type exhibits its unique mitigation effects and potential challenges. For instance, *Carbon removal* allows firms to continue using existing equipment or fuels at a lower transition cost but may encounter regulatory hurdles; *Renewable energy* is more effective in reducing emissions but can be affected by weather conditions and natural resources; *Enabling technology* can assist a broader range of companies through the supply chain, but its success depends on market acceptance. By looking at a firm's investments across these categories, investors can better understand the strategies of firms and their real effects on climate mitigation.

The transition risk exposure consists of *Technology challenges* and *Regulation concerns*. The first refers to the difficulties firms face when trying to develop climate technologies, such as impractical R&D, reduced funding, and customer reluctance, all of which render the pace of decarbonization unpredictable. *Regulation concerns* reflect risks from climate regulations. For example, companies that are heavily reliant on fossil fuels may be susceptible to regulations like cap-and-trade systems and carbon taxes. While technology is necessary for creating new solutions and making it possible to reduce emissions, regulation is vital for creating the right incentives and ensuring firms are contributing to the transition. By differentiating between *Technology challenges* and *Regulation concerns*, firms and policymakers can develop more targeted and effective strategies and accelerate the transition while mitigating the associated risks.

Our physical risk exposure is split into *Disaster* and *Weather*. *Disaster* detects natural disasters that inflict acute damage on firms' properties and operations. *Weather* pays attention to prolonged weather patterns such as abnormal temperatures and unusual precipitation levels, which occur more frequently than rare disasters and have chronic impacts. While *Disaster* usually involves significant one-off costs for repair and recovery, *Weather* can involve ongoing and potentially increasing operational costs. Distinguishing between the two is also essential as their different levels of persistence could imply different financial effects.

We conduct various validation tests for our measures. To validate whether transition effort exposure effectively represents firms' actions towards a low-carbon economy, we analyze their correlations with different types of green patents and carbon emissions. We observe that *Carbon removal* displays a positive and significant link with carbon abatement patents. *Renewable energy* is positively associated with patents about nuclear and renewable energy generation. *Enabling technology* is the only measure with a positive relation with innovations designed to improve energy efficiency and advance electrification. On the other hand, the relationship with carbon emissions reveals that *Carbon removal* and *Renewable energy* are more frequently discussed by carbon-intensive firms. This indicates that major emitters are dedicated to either capturing the emissions they produce or replacing traditional energy with renewable sources. Interestingly, while *Carbon removal* shows a significant decreasing effect on changes in carbon intensity, this does not translate to changes in carbon emission levels. It could be attributable to the rebound effect, where fossil-fuel-based techniques improve energy efficiency but could also increase demand (Bolton et al., 2023). In comparison, *Renewable energy* presents as a potent carbon reduction approach because it significantly reduces both carbon intensity and emission levels, indicating a greater reduction or a lesser increase in carbon emissions relative to industry peers. *Enabling technology* is primarily discussed by non-carbon-intensive firms. Although not reducing their own carbon emissions, they facilitate a slowdown in the growth rate of emission levels across all three scopes for their customers through supply chains. The patent activities and carbon reduction effects confirm that these three measures not only represent firms' climate efforts but are also in line with the emission mitigation hierarchy.

To validate our transition risk exposure, we examine their relationship with carbon emissions and observe that firms placing emphasis on these two categories tend to increase their emissions at a faster rate. Furthermore, firms frequently discussing *Regulation concerns* are more likely to experience negative climate incidents related to litigation risk, while *Technology challenges* finds more prominence in discussions within firms facing financial constraints.

To validate our physical risk measures, we employ dummy variables for different types of extreme climate events. *Disaster* shows a significant positive association with the occurrence of hurricanes and floods, which account for over 70% natural disasters in the past century. Conversely, *Weather* is positively correlated to drought and heat waves.

Our validated text-based measures enable us to investigate whether firms' valuations reflect their climate change exposure and further examine if these effects originate from the discount rate or cash flow channels. We note that companies whose earnings call transcripts comprise a higher proportion of transition efforts demonstrate increased valuations. This could be a result of potentially stringent regulations that favor these firms, which subsequently boost their stock prices. By contrast, firms discussing *Technology challenges* and *Regulation concerns* witness a decrease in their values. The discount rate channel is the primary driver of these value effects: value-weighted returns for high-minus-low portfolios that take a long (short) position in the quintile portfolio of the highest (lowest) *Enabling technology*/*Renewable energy*/*Carbon removal* is -0.50% /-0.48%/-0.47%, significant at the 1% level; long-short portfolios based on *Regulation concerns* yields a statistically significant average return of 0.54% per month. These return patterns cannot be explained by common risk factors. Fama-MacBeth regressions with a wider set of controls obtain results that remain economically and statistically significant. Conversely, the cash flow channel plays a less important role as the effects on profit growth become insignificant within a year.

For physical exposure, we observe value-decreasing effects for both *Disaster* and *Weather*. The portfolio sorting demonstrates that expected returns increase with physical exposures. Firms with high exposure are more susceptible to future extreme climate events due to their geographic location and operational sensitivity to weather. Investors seek compensation for the risk associated with holding these vulnerable stocks. This positive correlation is particularly strong for *Weather*, which reflects abnormal weather patterns that can persist for several days or even months. Investors perceive these physical events as more enduring and require higher expected returns compared to *Disaster*. The relation between physical measures and profit growth further corroborates our findings: although both measures are negatively related to changes in profitability, *Weather* displays smaller coefficients but remains significant over a longer period, substantiating its role in capturing the effects of chronic weather patterns.

Overall, firms discussing transition efforts (risks) more frequently have lower (higher) expected returns, correspondingly increasing (decreasing) their firm valuations. On the other hand, firms with larger physical risk exposure measures tend to have higher discount rates, thus lowering their present value.

We investigate whether each of our two methodological advancements translates into the value implications we discern. To isolate the effects of the objective creation of the climate word pool, we use keywords from companion works including Sautner et al. (2023) and Li et al. (2023), and then break the related sentences into seven categories using LDA and FinBERT. To probe the value of our sentence-level categorization that enhances interpretability, we gather sentences pinpointed by our climate terms and consolidate them into a single overall exposure measure. In both experiments, the association with firm valuations loses significance, indicating the economic value of newly identified information through both advantages. Therefore, our contributions are twofold: not only do we improve the method for better precision in constructing climate exposure measures but, more crucially, we obtain important economic insights through these method improvements, shedding light on previously undiscovered channels that climate change can influence corporate valuations.

Literature review: A few text-based climate exposure measures have emerged in the literature. Engle et al. (2020) measure aggregate climate change risk by calculating the correlation between the text content of the Wall Street Journal each month and a fixed climate change vocabulary. Ardia et al. (2020) construct a Media Climate Change Concerns index using news about climate change published by major U.S. newspapers. Faccini et al. (2022) conduct LDA on Reuters climate-change news and uncover risk factors related to natural disasters, global warming, international summits, and U.S. climate policy.

Some studies also focus on constructing firm-level climate change measures using textual analysis. Ceres, a nonprofit organization on climate change, evaluates the length of relevant disclosures in 10-Ks and provides climate risk measures for Russel 3000 firms. However, most companies' climate disclosures in 10-Ks are very brief and use boilerplate language with limited utility for investor communication (Ceres, 2014). Our paper is most closely related to Sautner et al. (2023) and Li et al. (2023), both of which assess climate change exposure by counting the frequency of certain keywords in earnings calls. In contrast, we employ more advanced machine learning techniques, improving climate information detection and categorization, which aids in the discovery of novel economic interpretations.

Our paper contributes to the literature on climate change exposure pricing. Existing literature investigates the pricing effects in equity markets for green versus brown firms, with mixed findings. Bolton and Kacperczyk (2021) argue that firms with higher CO_2 emissions earn higher stock returns, reflecting investor demand for compensation regarding carbon emission risk. Hsu et al. (2022) demonstrate that highly polluting firms are exposed to regulation risk and command higher average returns. Leippold and Yu (2023) find that firms with more green innovations have lower expected returns. Conversely, Pastor et al. (2022) contend that green assets have delivered high returns in recent years due to an unexpectedly strong increase in environmental concerns. Berg et al. (2021) find that stocks with higher ESG performance have higher expected returns. Our study adds to this literature by showing that firms discussing climate mitigation efforts in greater detail enjoy a lower discount rate.

Several studies address the pricing of physical climate risks. Hong et al. (2019) find that drought conditions forecast poor profit growth and stock returns for food companies, suggesting that food stock prices underreact to climate change risks. Huynh and Xia (2021) demonstrate that a firm's stock and bond returns are higher in the subsequent month when exposed to natural disasters, attributing this to investor overreaction. Kruttli et al. (2021) reveal that options of firms with establishments in hurricane landfall regions exhibit large, long-lasting implied volatility increases. We build on this literature by showing that firms exposed to physical risks related to natural disasters and abnormal weather patterns are expected to yield higher returns.

The remainder of our paper is organized as follows. Section 2 introduces the datasets we use for our textual analysis and empirical tests. Section 3 provides a detailed description of our textual measure construction. In Section 4, we validate our measures using various climate exposure proxies. Section 5 compares our measures with existing text-based climate measures. Section 6 investigates the value implications and examines whether the effects are dominated by the discount rate or cash flow channels. We conclude our paper in Section 7.

2 Data

In this paper, we first use textual information from IPCC reports and earnings call transcripts to construct firm-level climate change exposure measures. We then validate these textual measures using disaster dummies from the Spatial Hazard Events and Losses Database for the United States (SHELDUS), green patents from the United States Patent and Trademark Office (USPTO), climate incidents from RepRisk, and carbon emissions from Trucost. Firm-level financial data is obtained from Compustat North America Fundamentals Quarterly database via Wharton Research Data Services (WRDS). Finally, we investigate the asset pricing implications based on monthly returns from the Center for Research in Security Prices (CRSP).

2.1 Textual data

The reports published by the IPCC provide an exhaustive overview of the drivers of climate change, its impacts, future risks, and the ways in which adaptation and mitigation strategies can alleviate these risks. The in-depth exposition within these reports is instrumental in allowing us to extract keywords that are pertinent to climate change. Our analysis encompasses Synthesis reports, which offer an integrated perspective, as well as Special reports that delve into specific issues related to climate change. We have included reports published post the year 2000. Table IA1 enumerates the titles of these reports. We obtain all reports in PDF format from the IPCC website and subsequently convert them into text files utilizing Adobe Acrobat.

Earnings conference call transcripts, available from Thomson Reuters' StreetEvents database as XML files, span from January 1, 2001, to November 3, 2020. Two types of transcripts exist: complete versions that record both the presentation and the Q&A session, and brief versions that include an overview of the presentation and the entire Q&A session. We use the brief transcript only if the complete record is missing.

2.2 Data for validation tests

We obtain information on hazards from the SHELDUS database, which contains quarterly property losses resulting from natural disasters at the county level. Using Census Geocoder and Google Maps API, we obtain firms' county locations from historical headquarters addresses collected from the SEC header of the 10-K filings.

We retrieve patent data from the USPTO[2] and construct three categories of green patents in alignment with the strategy outlined by the Organisation for Economic Co-operation and Development (OECD)[3] and Lanzi et al. (2011). The detailed classifications for these patents are presented in Table IA2. Furthermore, we aggregate data on firms' incidents related to climate change, greenhouse gas emissions, and local pollution from RepRisk. Within this dataset, *severity* represents the magnitude of the incidents' impacts, *reach* quantifies the extent of dissemination of information through various sources, based on metrics such as readership and circulation, and *novelty* indicates whether an incident represents a first-time occurrence for the company.

Data on carbon emission intensity is obtained from Trucost, which computes the level of greenhouse gas emissions relative to total revenue. In this data, Scope 1 emissions refer to direct greenhouse gas emissions resulting from the combustion of fossil fuels and production processes that are owned or controlled by the company. Conversely, Scope 2 emissions are derived from the consumption of purchased electricity, heat, or steam by the company. Lastly, Scope 3 emissions include other indirect greenhouse gas emissions that occur upstream in the value chain.

2.3 Data for asset pricing tests

Our sample comprises firms that are present in both the Compustat database and earnings call transcripts. We link these two datasets by matching company names and elaborate on this matching process in Internet Appendix (refer to Section IA.2.4). Additionally, we gather monthly returns of individual stocks that are traded on NYSE, AMEX, and NASDAQ from the CRSP database, excluding financial firms with SIC classifications between 6000 and 6999.

[2] https://patentsview.org/download/data-download-tables

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

3 Textual measures

We employ natural language processing techniques, including LDA, word2vec, and FinBERT, to create our text-based climate change exposure measures. Figure 1 illustrates our three-step procedure. Firstly, we conduct LDA analysis on IPCC reports to identify seed words related to climate change. Secondly, we train a word2vec model using all earnings call transcripts to generate vector representations for each word in the vocabulary. We then select words in transcripts exhibiting the highest similarity with seed words from IPCC reports and create a climate word pool containing 1,129 words. Thirdly, gathering sentences in transcripts featuring at least one word from the climate word pool, we perform LDA and FinBERT on these climate-focused sentences to determine each transcript's loading on distinct climate topics with different sentiments in real-time.[4]

[Insert Figure 1 here]

3.1 NLP approaches

3.1.1 Step 1: Seed words from IPCC reports

To efficiently analyze the content of firms' earnings call transcripts for insights on climate change, it is imperative to establish a list of climate-related keywords to enable the identification of sentences where climate change is the focal point. For constructing such a word list, the IPCC reports serve as a gold standard due to their comprehensive and authoritative coverage of climate science.

Specifically, we employ the LDA method to distill recurrent and informative terms within these reports.[5] LDA, a generative probabilistic model, posits that documents are composed of a latent topic structure, where each topic is represented as a probability distribution across the vocabulary. While the same topics are inherent in all documents, their proportions may vary

[4]We utilize the Gensim library in Python to train the word2vec and LDA models and execute all Python code on a GV100 server. Section IA.1 provides more descriptions and mathematical deviations for the NLP techniques we used.

[5]We begin by applying standard preprocessing procedures to IPCC reports, including tokenization, lowercasing, lemmatization, and removal of stop words, to enhance computational efficiency. Utilizing the Phraser module in the Gensim library, we identify collocations specific to the IPCC corpus, which capture the combination of words that frequently co-occur. Phrases may possess meanings distinct from the individual meanings of their constituent words (e.g., "New York" vs. "New" and "York"). Following Mikolov et al. (2013), we select bigrams and trigrams with scores above a threshold. The score for a phrase comprising words w_i and w_j is calculated as $score(w_i, w_j) = \frac{count(w_i w_j) - \delta}{count(w_i) \times count(w_j)}$, where δ preventing phrase formation from very infrequent words.

widely among different pieces of text. Within each topic, the words that have been assigned the highest probabilities are often indicative of the central themes that the topic represents. For instance, a topic with high probabilities associated with terms like "solar project", "wind farm", and "sustainable" would typically pertain to renewable energy practices among firms. Compiled by thousands of climate scientists worldwide, IPCC reports represent the most current understanding of climate change. The automated extraction of top keywords from these authoritative climate reports through LDA provides not only a wider range of potential topics than human selection could furnish but also a level of scalability that is impractical to achieve manually.[6]

Furthermore, one may concern about the stylistic differences between formal written documents such as IPCC reports and the more conversational style of earnings call transcripts, we therefore carefully scrutinize the seed words and remove terms that may not be typically used in earnings calls. Figure IA3 lists the 54 climate seed words that made the final cut. These contain a range of themes such as traditional energy sources (e.g., fossil fuel, carbon), alternatives in renewable energy (e.g., solar, wind power), physical events (e.g., flood, precipitation), and emissions-related terminology (e.g., greenhouse gas, combustion, pollution, HFC). These seed words form a robust foundation for analyzing the impact of climate change, circumventing the necessity for conceptualizing climate-related elements from scratch.

3.1.2 Step 2: Word expansion to detect climate sentences in earnings calls

While the seed words extracted from IPCC reports offer a foundation, they might not comprehensively represent the multifaceted ways in which firms are exposed to climate change. This could be attributed to the targeted audiences of the IPCC reports, which are primarily policymakers. In contrast, firms often have more granular and pragmatic considerations regarding the impacts of climate change on their business operations. To enrich our climate word pool and to better discern sentences that encapsulate firm-specific exposures to climate change, we adopt the word2vec methodology on the entire earnings call transcripts.[7]

[6]IPCC reports often cover various topics with extensive document lengths. To more effectively discern these topics, LDA is performed at the chapter level, which results in documents of similar lengths and a higher likelihood of concentration on specific themes. Based on coherence values, we opt for seven topics as shown in Figure IA4a, taking into account both model performance and human interpretability. Nonetheless, the number of topics in this step remains relatively flexible, since the emphasis is placed on the predominant words within the topics rather than the distribution across various topics. For robustness, we test larger numbers of topics and obtain similar dominant words. Hanley and Hoberg (2019) also utilize top words in LDA-generated topics from banks' 10-Ks to detect emerging risks, producing 25 topics in their paper.

[7]For the preprocessing of transcripts, which differs from that of the IPCC reports, we employ Stanza, an open-source Python NLP toolkit. It applies neural networks to raw text inputs, generating sentence-level documents that identify entities such as company names, speaker names, and locations, tagging them

Word2vec is a two-layer neural network to generate word embeddings. It takes a text corpus as input and outputs vector representations for each unique word in the vocabulary. The cosine similarities between vectors serve as a reliable metric for the semantic relatedness of the words these vectors represent. Trained by earnings call transcripts, this step generates an abundance of similar terms that frequently appear in transcripts (e.g., extreme weather, emission reduction, renewable energy) and fewer words that are less common in firm discussions (e.g., sea level, vegetation).[8]

Ultimately, we obtain a list of 1,129 climate-relevant terms, as documented in our Internet Appendix IA.2. This collection contains information that sheds light on not only the impacts of climate change on businesses but also on firms' strategies and efforts to mitigate these effects.[9] For example, the ten terms exhibiting the highest degree of semantic proximity to the phrase "climate change," as determined by cosine similarity within the transcripts, are "global warming," "address climate change," "greenhouse gas," "climate protection," "decarbonization," "energy policy," "environmental issue," "carbon reduction," "fight against climate change," and "clean energy."

3.1.3 Step 3: Anatomy of firm-level climate change exposures

Sentences are climate-related if they include at least one term from the climate word pool. However, the consequences of climate change are varied and multifaceted, ranging from supply chain disruptions due to extreme weather events to new opportunities associated with the transition toward a low-carbon economy. These diverse exposures could have different implications for operations and asset pricing. Therefore, proper categorization is crucial for delving into the nuances of firm-specific climate exposures.

Unlike previous literature that often adopts predefined topics with specific keywords and then counts the frequency of the related keywords within each topic, we take a different approach. Specifically, we employ LDA to analyze climate-related sentences in each transcript. The LDA algorithm autonomously generates five distinct topics based on coherence mea-

accordingly. Named entities are replaced with generic tags to facilitate enhanced semantic learning. Stanza achieves better or comparable performance on various training corpora compared to existing toolkits. After preprocessing, the transcripts are refined into lines of clean sentences.

[8]The extended terms reveal disparities between IPCC reports and earnings calls. For instance, sea-level rise is seldom a central concern for firms, as its direct impact on operations is not immediate. On the other hand, rather than discussing emission volumes, firms may opt to emphasize their progress in electrification as a proxy for efforts in curbing carbon emissions.

[9]To bolster efficiency, hierarchical agglomerative clustering is used to divide seed words into six clusters. The vector representation for each cluster is computed as the mean of vectors corresponding to the words within that cluster. This helps in finding synonyms at the cluster level instead of doing it for each seed word individually.

sures, as depicted in Figure IA4b. Figure IA5 shows the word cloud of the five topics, with the size of words being proportional to their probability within the respective topic.[10] For each sentence, we utilize the trained LDA model to calculate topic loadings and subsequently assign the sentence to the topic that has the highest proportion, provided it surpasses a 50% threshold.[11]

Upon examining the contexts associated with each topic, it becomes clear that sentiment analysis is crucial, particularly for transitional exposures. For instance, within the "carbon" topic, a clear dichotomy is observed where some firms are promoting advancements in carbon capture technology, while others express concerns regarding the rise of carbon taxes. Sentiment analysis, therefore, enables us to differentiate between the opportunities and challenges faced by various firms. For this purpose, we employ the FinBERT model to implement sentiment classification for each sentence. We also apply the ClimateBERT model from Bingler et al. (2023), which is fine-tuned on climate-related content to discern opportunities and risks associated with climate change, for robustness and obtain similar classification results.[12]

Each sentence's category is determined by its topic and sentiment.[13] For example, sentences centered around the topics "renewable" and "technology" with non-negative sentiment are categorized under *Renewable energy* and *Enabling technology* respectively. Conversely, those sharing the same topics but expressing negative sentiment are categorized as *Technology challenges*. Similarly, sentences that talk about the topic "carbon" with differing sentiments are assigned to either the *Carbon removal* or *Regulation concerns* categories. The categories *Weather* and *Disaster* emphasize the negative discussions related to their respective topics.

3.2 Constructed measures

To quantify the extent of climate-related discussions within earnings call transcripts, we calculate the proportion of sentences that is devoted to each of the seven categories. The

[10] Before employing LDA, we exclude firms in the financial sector due to their unique climate considerations which might distort word distributions across topics. We also experimented with a higher number of topics, which produced more granular industry-specific categories. For instance, eight topics yielded one focused on agricultural terms such as "crop," "soil," "fertilizer," and "drought." However, as our aim is to evaluate climate change impact holistically across the market, a five-topic model is optimal.

[11] This can filter out sentences that lack specific details and maybe vaguely related to various categories.

[12] While our baseline empirical results utilize measures derived from FinBERT, we also conduct an analysis using ClimateBERT and observe consistent implications for firm valuations, as documented in Table IA17.

[13] In this paper, we employ the term "topic" to denote the output of LDA, and "category" to represent our final seven exposure types, which take into account both topic and sentiment behind related sentences.

mathematical representation of the measure is as follows:

$$Climate\ exposure_i^k = \frac{\sum_{j \in i} 1_{j,k} \cdot l_j}{\sum_{j \in i} l_j}, \tag{1}$$

where $Climate\ exposure_i^k$ represents the share of discussions related to climate exposure category k in transcript i. l_j denotes the length of sentence j. $1_{j,k}$ is a dummy variable that equals one if sentence j belongs to category k, and zero otherwise.

[Insert Table 2 here]

Table 2 provides detailed descriptions of these seven exposure measures. Moreover, Internet Appendix IA.2.2 provides excerpts from earnings call transcripts, which serve as illustrative examples of climate-related discussions that fall under each category. Table 3 reports the summary statistics for our textual measures.[14] Excluding firms in the financial industry, we have 155,108 firm-quarter observations covering 5,969 unique U.S. firms. *Renewable energy* appears in 10,524 earnings call transcripts, with a mean value of 0.13%. The most commonly discussed category, *Enabling technology*, is present in 20,334 transcripts with an average of 0.15%. The category *Disaster* (*Weather*) has a smaller mean value and is mentioned in 12,626 (15,993) earnings calls. Furthermore, only a small proportion of discussions in transcripts address firms' *Regulation concerns* and *Technology challenges*.

[Insert Table 3 here]

3.2.1 Transition exposure

Our transitional exposure measures are broadly divided into three groups: *Mitigation efforts*, *Technological challenges*, and *Regulatory concerns*. *Mitigation efforts* comprise three distinct categories: *Carbon removal*, *Renewable energy*, and *Enabling technology*. These measures represent the opportunity side of transitional topics, highlighting the actions firms are taking to combat climate change. On the other hand, *Technology challenges* and *Regulation concern* depict the negative aspects of related topics, capturing the difficulties and worries firms face in their transition process.

Technology challenges reflects the difficulties firms encounter when developing green technologies, including impractical R&D, reduced funding, and customer reluctance. For instance, renewable energy production can be influenced by environmental factors like dry

[14] All textual measures are winsorized at the 99.9^{th} percentile to reduce the impact of outliers. We winsorize this way because only 15% of observations have non-zero discussions about climate change. Our empirical results remain if we winsorize at the 1^{st} and 99^{th} percentiles.

conditions, wind volumes, and solar resources, potentially failing to satisfy energy demand. This category also considers the fact that current technologies are not sufficient for the required emission reductions needed by 2050, and further reductions would depend on the development of new technologies. These challenges make the pace of decarbonization unpredictable.

Regulation concerns describes firms' concerns regarding the costs and risks related to environmental regulations. For example, companies that are heavily reliant on fossil fuels may face challenges due to regulations like cap-and-trade systems and carbon taxes. These companies may express concern over the costs of compliance, the risk of obsolescence, and the potential phase-out of certain technologies or fuels.

Lastly, each of the three mitigation effort-oriented measures carries a unique emphasis. *Carbon removal* focuses on solutions for managing generated carbon emissions, *Renewable energy* underscores the use of cleaner energy to cut emissions at the source, and *Enabling technology* spotlights techniques that improve energy efficiency or facilitate electrification. We then delve deeper into how these measures align with the GHG emission mitigation hierarchy shown in Figure 2.

[Insert Figure 2 here]

GHG emission mitigation hierarchy: Most global carbon emissions originate from the energy system. Achieving net-zero emissions mandates a profound transformation in both energy supply and consumption. The first priority in mitigating GHG emissions is to save energy and ensure the least amount of energy is consumed. For example, energy-efficient building design using thermal insulation and innovative materials can decrease energy demand. *Enabling technology* encompasses such techniques, allowing firms or their clients to decrease energy consumption through improved efficiency and product redesign. Prioritizing energy demand reduction often brings immediate benefits, making the subsequent energy transition more manageable in the long run.

The second priority involves substituting carbon-intensive energy with cleaner alternatives. This involves a shift from fossil fuels to renewable sources such as solar, wind, and hydropower for electricity generation and industrial processes. *Renewable energy* reflects this approach, emphasizing firms' competitive positions in clean energy adoption. Another crucial contributor, documented in *Enabling Technology*, is widespread electrification, converting various energy-consuming sectors like transportation and heating to rely on electricity rather than directly combusting fossil fuels. This strategy could reduce GHG emissions if the electricity used is generated from renewable sources.

Some activities cannot currently operate using renewable sources due to the long lifespan of existing assets. This situation underscores the need for fossil-fuel-based carbon reduction strategies incorporated in *Carbon removal*, especially crucial for decarbonizing hard-to-abate sectors with inefficient coal plants and chemical facilities that may contribute significantly to emissions in the coming decades.

In summary, the transition to a low-carbon economy presents both opportunities and challenges. Firms actively engaging in *Mitigation efforts* are likely to experience reduced regulatory risks and potentially gain a competitive advantage. However, *Technology challenges* and *Regulation concerns* represent the hurdles and potential costs that firms face during the transition. Our analysis permits the capturing of the evolution of climate discourse over time. Firms may initially express concerns regarding transitioning to sustainable practices, but as they adapt and innovate, the tone might change to reflect the successful integration of such practices and the realization of their benefits. Our approach can track these temporal changes and provide insights into the dynamic nature of firms' climate strategies.

3.2.2 Physical exposure

The physical climate exposure is divided into two distinct categories: *Disaster* and *Weather*. The *Disaster* category focuses on discussions about natural disasters that cause significant and immediate damage. In the transcripts, this category is characterized by the frequent mention of words such as "hurricane," "storm," and "flood." On the other hand, the *Weather* category emphasizes discussions about abnormal weather patterns which can include extreme cold, heat, or dry conditions. These patterns may have chronic effects on society compared to the immediate impact caused by natural disasters.

Importantly, both categories are centered around physical topics with negative sentiments. This is to mitigate the concern that the measures simply capture general discussions about weather and natural disasters and ensure that they reflect adverse impacts experienced by firms due to these climate events. High values in these measures are often associated with firms that are sensitive to weather changes and are located in areas prone to climatic events. This indicates increased physical risk in the future, especially considering that the frequency and severity of extreme weather events are anticipated to rise due to climate change.

Though one could argue for an alternative approach to capturing physical risk exposure, such as requiring physical climate keywords must be accompanied by "risk" synonymous. We have chosen not to impose this restriction for two reasons. Firstly, implementing such a requirement would significantly narrow down the scope of discussions that are detected as being about physical climate topics. However, our intention is to capture a more comprehensive picture of the physical exposure faced by firms. Secondly, keyword-based methods, such as mandating the presence of the word "risk," are less robust compared to context-aware approaches. The latter, such as the detection of negative concerns through the FinBERT model, are more effective, especially considering that companies are adapting their language and disclosures to accommodate automated processing by AI, as highlighted in the study by Leippold (2023). The Internet Appendix IA.2.3 offers examples of sentences that are misclassified when using the restrictive word method based on "risk" synonyms.

4 Validation of textual measures

To validate our firm-level exposure measures, we take a three-step validation process. Firstly, we verify whether the time-series pattern of our exposure measures aligns with important events related to climate change. Secondly, we evaluate if the cross-sectional variation of our measures corresponds with industry patterns, acknowledging that different industries may be subjected to varying levels of climate risk. Lastly, we examine the categorization of our climate exposure measures to examine whether our constructed physical risks correlate with realized natural disasters, and if our transitional measures such as *Mitigation Efforts* are strongly associated with firms' different types of green patent applications. Moreover, we assess if these *Mitigation Efforts* lead to reductions in carbon emissions for the firms themselves or their clients in the following year.

4.1 Time-series pattern

Transition exposure: As discussed in Section 3, our transition exposure includes five categories in three groups: *Mitigation efforts* (including *Carbon removal*, *Renewable energy*, and *Enabling technology*), *Technological challenges*, and *Regulation concerns*. Figure 3 displays the mean value of each measure across firms over time. Panel A shows that the five transition measures began to rise in the early 2000s, peaking in 2010-2011, immediately following the Copenhagen Summit.[15] Despite a temporary lull in focused discussions during the first quarter of 2020 due to the emergence of COVID-19, the last two years have witnessed a sharp uptick in discussion concerning transitional topics. However, conversations revolving around *Mitigation efforts* vastly overshadow those addressing *Technological challenges* and

[15]This international conference recognized that all countries should take action to reduce CO_2 emissions and keep temperature increases below 2°C.

Regulatory concerns.

Physical exposure: We have two physical risk exposures: *Disaster* and *Weather*. The *Disaster* measure reveals a pronounced, temporary spike in periods when the U.S. endured severe hurricanes and floods, effectively reflecting the time-series fluctuations in overall damages inflicted by natural disasters. The *Weather* measure increases when firms are significantly impacted by adverse weather events such as heavy rains, droughts, abnormal temperatures, or disease outbreaks induced by these conditions. Our NLP procedure inherently distinguishes these two categories, suggesting that firms express their concerns about these distinct types of climate events differently.

[Insert Figure 3 here]

4.2 Industry distribution

To examine industry-level exposure to different climate measures, we average textual measures across firms within the same industry according to Fama-French 49 industry classifications. Table IA4 ranks industries with the top ten climate exposure measures. The diverse distributions across industries for each measure affirm that they reflect different facets of climate exposures.

Transition exposure: Discussions regarding *Carbon Removal* are primarily found within heavy industries that produce traditional energy and materials — such as Chemicals (with a mean of 0.81%), Precious Metals (0.47%), Coal (0.46%), and Petroleum (0.41%). Over 70% of Utility firms discuss *Renewable Energy*, yielding the highest industry-level mean (2.50%) and standard deviation. Coal and Construction firms also have substantial exposure to *Renewable Energy*, highlighting that significant greenhouse gas emitters are deeply invested in renewable energy initiatives. Companies within the Electrical Equipment (1.52%) and Automobile (0.77%) industries emphasize *Enabling Technology*, aligning with the crucial role of electrification in addressing global warming. Utilities and Electrical Equipment firms are more inclined to address *Technology Challenges*, while firms in Coal and Chemical industries exhibit sensitivity towards *Regulation Concerns*.

Physical exposure: The Agriculture industry has the most discussions related to *Weather* and *Disaster*. On average, its *Weather*-related discussions make up 0.75% of their content, which is double that of the second-most talkative industry in this category. Industries like

Metal Mining and Utilities also frequently discuss vulnerabilities to unusual weather patterns, while sectors related to Construction and Transportation tend to focus on natural catastrophes.

We further conduct a variance decomposition to evaluate the relative contributions of aggregate, industrial, and firm-level climate change exposure. This is done by determining the proportion of variation in each measure accounted for by various sets of fixed effects. Table IA5 demonstrates that the bulk of the variation occurs at the firm level, as opposed to the industry or the entire economy, especially in the case of physical exposure.

4.3 The validation tests

In this subsection, we carry out regression analysis to examine the relationship between our constructed climate exposure measures and real-world outcomes such as realized disasters, the application and development of new climate technology, and reductions in carbon emissions. In general, the key is to look for external, objective measures that are expected to correlate with the internal, subjective measures identified from the earnings calls. If these correlations hold up, it provides good evidence that our textual measures are capturing meaningful information about the firms' climate exposures.

4.3.1 Green patents and climate incidents

To validate our transitional climate exposure measures, we regress alternative transition measures on our textual measures using the following equation,[16]

$$AlterTransition_{i,j,t+1} = \sum_{k} \beta_k Climate\ exposure^{k}_{i,j,t} + \gamma X_{i,j,t} + \delta_j + \delta_t + \epsilon_{i,j,t}. \tag{2}$$

First, we examine whether firms that devote higher proportions of their discussion to effort-related measures also invest more heavily in the development and application of climate technology. This is determined through an analysis of patent data from the USPTO. To substantiate the differentiations between these effort measures, we use the classifications of green patents in Table IA2. The patents span various domains. Those concerning carbon abatement address air pollution control, enhancements in fossil fuel efficiency, and carbon capture methods. In contrast, renewable patents are dedicated to renewable and nuclear

[16]When the dependent variable is at the year level, we average the four quarterly textual measures each year to obtain annual proportions of climate measures.

energy. The last category encompasses advances in energy efficiency, electrification, and enabling technologies across different sectors.

[Insert Table 4 here]

The results, as presented in Panel A of Table 4, indicate that the *Renewable energy* measure has a significant and positive correlation with renewable patents, yet a negative correlation with carbon abatement technology. The *Enabling technology* measure is the sole metric with a significantly positive coefficient on enabling patents and also exhibits a positive correlation with carbon abatement technology, reflecting its role in enhancing fossil fuel efficiency. *Carbon removal* displays a positive coefficient on carbon abatement patents. These findings underscore that transitional measures capture various aspects of firms' greenhouse gas emission mitigation efforts, which aligns with the previously discussed hierarchy.

The above validation process primarily focuses on the three aspects of *Mitigation efforts*. However, the question arises as to how we should validate the other two categories of transition measures, namely *Technological challenges* and *Regulation concerns*.

Establishing proxies for challenges experienced during technology development is a complex endeavor. Nevertheless, we partially validate this measure by employing the WW index, which quantifies firms' financial constraints. Our findings suggest that firms discussing *Technology challenges* more frequently tend to be more financially constrained.

With regard to *Regulation concerns*, Panel B examines firms that are frequently implicated in negative incidents related to climate change, greenhouse gas emissions, and local pollution. We employ proxies such as the number, severity, reach, and novelty of climate incidents as dependent variables. Only *Regulation concerns* provide a positive prediction of negative incidents in the subsequent year, while other measures do not exhibit any significant correlation. This implies that firms placing a higher emphasis on *Regulation concerns* in their discussions are more likely to face litigation risk.[17]

4.3.2 Carbon emission reductions

Do firms that discuss their climate efforts manage to reduce their carbon emissions? Table 5 examines the carbon reduction effects of transitional measures, utilizing regression equation (2) and emission data from Trucost.

[17] We conduct identical regressions on physical exposure measures in Table IA7 and find insignificant coefficients, except for a negative correlation between *Weather* and renewable patents. This can be attributed to the sensitivity of renewable energy production to weather conditions.

[Insert Table 5 here]

Panel A evaluates the impacts on a firm's own Scope 1 emissions. Column (1) utilizes carbon intensity in the same year as the earnings call as the dependent variable. The significantly positive coefficients on *Carbon removal* and *Renewable energy* suggest that major carbon emitters are committed to developing technology to capture their emissions or replace traditional energy sources with renewables. The dependent variables in columns (2) and (3) are $\Delta Carbon\ intensity_{i,t+1}$, which is the carbon intensity of firms i in year $t+1$ minus its value in year t, and the growth rate of the carbon emission level, respectively. *Carbon removal* exhibits a significant diminishing effect on changes in carbon intensity, but not on the carbon emission level. This aligns with the rebound effect mentioned in Bolton et al. (2023), implying that improvements in fossil fuel technology do not necessarily reduce carbon emissions as they could also trigger an increase in energy demand. On the contrary, *Renewable energy* appears to be a more effective strategy for carbon reduction, as it significantly decreases the growth in firms' carbon intensity and level in the subsequent year. Interestingly, *Regulation concerns* exhibit a weak effect in increasing carbon emissions.

Unlike the other two measures related to efforts, *Enabling technology* is often discussed by firms that are not carbon-intensive. It shows a negative correlation with carbon intensity and does not have a definitive impact on the firms' own carbon reduction. This could be because such technologies are deployed by the firms' clients to reduce emissions rather than by the firms themselves. Panel B investigates this possibility, replacing the dependent variables with the growth rate of emissions in scopes 1, 2, and 3 from the firms' customers.[18] The results align with the concept of technology spillover effect; *Enabling technology* significantly decreases the growth rate of carbon emissions across all three scopes for the firms' customers.[19]

4.3.3 Realized natural disasters

We validate our physical risk measures by conducting regression analysis with them against various types of extreme weather events from the SHELDUS database:

$$Climate\ exposure^{k}_{i,j,t} = \beta_h Realized\ disaster^{h}_{i,j,t} + \gamma X_{i,j,t} + \delta_j + \delta_t + \epsilon_{i,j,t}, \tag{3}$$

[18] We collect information on customer-supplier relations from the Compustat Segment (WRDS Supply Chain with IDs). Firms are required to disclose the existence or sales to customers representing more than 10% of total firm revenues.

[19] Table IA8 examines the relationship between transition measures and firms' own scope 2 and 3 emissions, noting a similar carbon reduction effect of *Renewable energy*.

where $Climate\ exposure^{k}_{i,j,t}$ represents the proportion of exposure k in firm i's transcripts in industry j at time t. We normalize all textual measures to have a mean of zero and a standard deviation of one for subsequent analyses. $Realized\ disaster^{h}_{i,j,t}$ is a dummy variable for type h extreme climate events, taking a value of one if the firm's headquarters is located in a county impacted by at least one type of natural disaster in the preceding year.[20] The variable $Hurricane$ ($Flood/Drought/Heat$) equals one if the firm experienced a hurricane (flood/drought/heat) in the past year.[21] Control variables include size, leverage, investment, tangibility, sales growth, ROA, and dividend. We also include industry and year-quarter fixed effects and cluster standard errors by industry.

[Insert Table 6 here]

Table 6 Panel A shows a positive and significant association between the *Disaster* measure and *Realized disaster*, *Hurricane*, and *Flood*. The occurrence of a natural disaster leads to a 10% standard deviation increase in *Disaster* discussion. The coefficient of *Hurricane* is larger than that of *Flood*, consistent with the literature indicating that hurricanes generally have more substantial impacts on firms due to their unpredictability and the difficulty in preparing for them. The same regressions conducted for *Weather* in Panel B demonstrate a statistically significant positive correlation with *Drought* and *Heat*. This confirms that the *Weather* measure captures weather events with chronic or weaker impacts. However, transitional measures do not display significant coefficients on realized disasters in Table IA6, indicating that they provide little information about physical climate exposures.

In a word, our validation tests confirm that *Carbon removal*, *Renewable energy*, *Enabling technology*, *Technology challenges*, and *Regulation concerns* are related to different aspects of transition exposure, while *Weather* and *Disaster* contain information about physical exposure.

5 Comparison with existing textual measures

In this section, we compare our textual measures to those developed by Sautner et al. (2023) and Li et al. (2023). Both papers quantify firm-level exposures by counting the frequency of

[20] We select disasters lasting less than 30 days. A county is defined as affected only if property and crop damages exceed 10 million in 2019 constant dollars. According to Huynh and Xia (2021), this threshold is approximately the minimum average damage caused by billion-dollar disasters to each county.

[21] We require hurricanes (floods) to result in damages exceeding 1 million dollars in the firm's county. Considering that abnormal weather patterns cause less damage compared to other acute disaster types, we include drought events causing damages exceeding 100,000 dollars at the county level and heat events resulting in damages over 1000 dollars in the past half year.

certain climate keywords within a set of predefined topics. The keywords of Sautner et al. (2023) are developed by traditional machine learning methods that treat words as isolated entities. Li et al. (2023) depends on manual efforts to curate keywords from sources such as severe weather lists, disaster summaries, and meteorology textbooks. Our methodology distinguishes itself by harnessing contextual information through advanced NLP techniques (LDA, word2vec, and FinBERT) with minimal human intervention and improved economic interpretability.

The method embraced in our research yields two primary benefits. Firstly, our word expansion technique, word2vec, facilitates a considerably broader and more precise word inclusion. This identifies three times the number of climate-related sentences in earnings calls compared to existing studies, emphasizing that omitting such discussions might misrepresent a firm's climate change exposure. Secondly, our study comprehends transcripts at the sentence level. Instead of focusing solely on individual words, we decipher both the topic and sentiment embedded within each climate-focused sentence. This nuanced analysis provides economically meaningful insights into diverse exposure facets, which, if combined without distinction, might mask or offset one another's empirical effects.

Table IA9 summarizes the distinctions between our study and two related papers across three key dimensions: (1) methodology adopted to identify climate keywords objectively and the scope of climate discussion captured; (2) taxonomy of climate sentences that enhance the economic interpretability; (3) empirical findings that are consistent with economic rationales behind each type of measure. Subsequent subsections elaborate on the first two dimensions and also partially examined the third point, which will be further explored in Section 6 by examining if our methodological enhancements translate to significant value implications.

[Insert Table IA9 here]

5.1 Reducing human bias

The first dimension of comparison lies in the methodology for identifying climate-related keywords. We apply the LDA technique on IPCC reports to derive seed words and then employ word2vec to expand this list. In an effort to minimize human intervention, our automated approach adds a layer of objectivity and scalability that is particularly advantageous for large-scale textual analyses. Crucially, word2vec weighs the context in which each term appears when determining whether it is climate relevant, which does not merely increase the number of climate terms but enhances the accuracy of terms detection. This minimizes the risks of omitting vital details or incorporating irrelevant data. For example, phrases such as "harmful emission" and "co2 reduction" directly relate to a firm's environmental footprint, "green hydrogen" and "renewable portfolio" reflect the growing emphasis in firms' strategic planning for climate mitigation, and "bad weather" and "extreme heat" echo the increasing mentions of extreme weather events. These phrases have a large frequency in our sample but might evade detection by a traditional climate dictionary.

We would like to compare our keywords with those in companion papers to document the difference in climate discussion coverage. However, we are unable to obtain the complete keyword lists from both studies. Sautner et al. (2023) provides the top 100 bigrams for each of their measures, namely, $CCExposure$, $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$, cumulating to 323 unique bigrams. Meanwhile, Li et al. (2023) disclose the top 30 keywords for each of their three topics, including *Transition risk*, *Acute risk*, and *Chronic risk*. These top keywords, despite their limited quantity, account for a significant portion of climate word frequencies in the two papers. Hence, our following comparison narrows its focus to high-frequency terms. This comparison, albeit limited, will shed light on the range and depth of climate-related words captured by our methodology.[22]

[Insert Figure 4 here]

It is important to highlight that this paper compiles a more exhaustive list of climate-related terms. Figure 4 presents Venn diagrams comparing our word pool with those from the two companion papers.[23] An impressive 75.78% of bigrams from Sautner et al. (2023) are

[22]One potential concern is that the most frequently used keywords do not contain the full range of information within measures in the two works, which might bias the comparison. We posit that this limitation is tolerable due to two rationales. First, keywords with the highest frequency considerably outweigh the others in terms of occurrences, thus capturing the majority of the information. For instance, the most frequent keyword related to physical topics in Li et al. (2023) appears over a hundred times more often than their 30th word, whereas in Sautner et al. (2023), the top one keyword is mentioned more than thirty times as frequently as their 100th keyword. Similarly, in our study, the keywords with more frequency than that of the 100th word from Sautner et al. (2023) are responsible for identifying 70% of the climate-related sentences, based on which we can also derive empirically significant findings concerning value implications. To ensure a fair comparison, we only consider terms in our pool whose frequencies exceed the top words from the two studies respectively. If our most frequent words exhibit a more comprehensive coverage compared to theirs, it lends credence to the inference that our entire corpus of words is likely more informative. Secondly, this concern diminishes further when examining the advantage of our sentence-level analysis. We aim to highlight identical words, depending on their contextual placement within sentences, can convey varying or even contrasting implications for firms, which does not necessitate a full coverage of the word pool. The fact that sentences containing the same top words convey different details already indicates the inadequacy of a solely word-level analysis.

[23]Differences in frequency for overlapping terms from different datasets arise from different sample lengths and bigram construction methods. While the two studies treat adjacent words as bigrams, we form bigrams only if the frequencies of words appearing together exceed a designated threshold for meaningful phrases. The largest discrepancy lies between the frequency of overlapped unigrams in our research and that of Li et al. (2023). This is largely attributed to most of their unigrams being physically related terms, counted

incorporated into our climate word pool, while only 36.91% of our bigrams are included in their work, not to mention our unigrams and trigrams. Similarly, a scanty 1.25% of unigrams and 39.78% of bigrams from Li et al. (2023)'s work are absent in our pool, while their work only includes 43.17% of unigrams and 18.90% bigrams in our work. The difference in climate word pools further leads to disparate capabilities in pinpointing climate-centric sentences.

A plausible reservation might be that the most frequently used terms in their works, although fewer in numbers, could already capture major climate-related content of earnings call transcripts effectively. If this is the case, the additional climate words in our pool might appear to be redundant. To examine the merit of augmenting the climate word pool, we conduct a comparison analysis to investigate if their words comprehensively cover each of our seven climate categories in Panel (a) of Figure 5. The findings unveil that the sentences pinpointed by their climate terms predominantly gravitate towards transitional themes, especially *Renewable energy* and *Enabling technology*.[24] In contrast, both sets display limited insights into the physical dimensions, as well as into areas like *Technology challenges* and *Regulation concerns*. However, these climate-related categories not covered by them are also important, as, in Section 6, it is shown that the physical aspects along with *Technology Challenges* and *Regulation Concerns* have significant empirical implications on firms' stock market performance. As such, our recognition of these additional terms is not only necessary but proves to be indispensable for a more comprehensive understanding of the climate-related contents in earnings call transcripts.

[Insert Figure 5 here]

Conversely, one may also be curious to know the terms from the companion studies that did not find a place in our climate word pool. Such omissions might spark concerns that our methodology, grounded in NLP techniques, could inadvertently miss certain facets of climate exposure. Should this be the case, it implies that human intervention, as employed in other studies for the selection of climate-related keywords or pertinent topics, could hold merit. To address this concern, we investigate the terms from companion studies absent in our pool to determine if their exclusion equates to a significant informational gap. Table IA10 shows that many of the missing terms either do not fit our criteria for phrase formation or pertain to climate change only in particular contexts, potentially introducing false positives. For example, bigrams like "government india," "obama administration," and "unite nation"

only when appearing alongside risk synonyms in neighboring sentences. In contrast, our approach counts all such terms present in physical sentences bearing a negative sentiment.

[24]According to the correlation matrix in Table IA3, these two categories exhibit a strong correlation with measures from Sautner et al. (2023).

might have ties to climate change discussions, but they are not strictly climate-centric and could relate to various non-climate topics. Similarly, terms like "vehicle manufacturer" and "driver assistance" might align with climate discourse when discussing electric vehicles or fuel efficiency, but not in all context. Incorporating these bigrams might inflate a firm's perceived climate change exposure by including non-climate-centric dialogues, thereby reducing the accuracy of the exposure measurement.

5.2 Enhancing economic interpretability

The second difference comes from the categorization of climate-related sentences to decipher the context surrounding climate-related keywords. Solely depending on word analysis proves insufficient to extract the complex impacts of climate change on firms.[25] For instance, consider two sentences from an energy company's earnings call: "Our investments in carbon capture technologies have reduced our emissions" and "The strict climate regulation on carbon emissions poses continuous pressure on us". Both statements feature similar climate-related words, yet their context and sentiment are quite different. The first sentence might be viewed positively by investors as it portrays the firm as proactive in addressing climate change, possibly leading to favorable market reactions. In contrast, the second one could signal future costs associated with regulatory compliance. A word-level analysis, even when supplemented by sentiment dictionaries, can blur these distinctions and would treat the two sentences the same.

Using LDA and FinBERT algorithms, we classify climate-related sentences into diverse categorizations that extend beyond traditional classifications. Our sentence-level contextual analysis unravels three nuanced layers of granularity compared to existing measures, each of which has economical interpreations. First, we dissect transition exposure into *Mitigation efforts*, *Technology challenges*, and *Regulation concerns*, which have divergent impacts on a firm's valuation and expected returns. Second, we further decompose *Mitigation efforts* into *Carbon removal*, *Renewable energy*, and *Enabling technology*. These categories exert distinctive carbon reduction effects and align closely with the hierarchy of greenhouse gas emission mitigation. Third, we partition physical exposure into *Weather* and *Disaster*, categories that exhibit varied levels of persistence and invoke different investor perceptions in the stock markets.

Figure 5 Panel (b) visually shows that sentences from the topics in Sautner et al. (2023)

[25] An increasing number of studies are integrating sentence-level analysis to more effectively capture contextual information in their target texts. (Kölbel et al. (2021), Leippold and Yu (2023), Giglio et al. (2023), Lopez-Lira and Tang (2023)).

and Li et al. (2023) are distributed across our seven categories, highlighting the coexistence of varied nuances within single topics. For instance, $CCExposure^{Opp}$ intertwines both initiatives and challenges in relation to mitigation strategies, whereas $CCExposure^{Reg}$ combines carbon mitigation efforts with regulatory concerns. While Li et al. (2023) presumes that every sentence containing transition keywords pertains to transition risks, our methodology categorizes most of their identified sentences under mitigation responses. Such discrepancies underscore the need for a more nuanced categorization of climate-related discourse.

Relying exclusively on word restrictions for topic decomposition may be also insufficient. Li et al. (2023) categorize their transition risk exposure into proactive and non-proactive classes based on the presence of a specified set of 16 verbs in the current or adjacent sentences. However, as Panel (c) illustrates, the mere presence of specific verbs in a sentence does not always indicate proactive transition exposure. On the flip side, firms may describe their initiatives without using any of the listed verbs. Internet Appendix IA.2.3 provides more sentence examples. Likewise, the approach to identifying physical risk—based on the presence of "risk" synonyms in the current or adjacent sentences—has its limitations. Firms might discuss the adverse consequences of extreme weather events using a plethora of terms without directly invoking "risk" synonyms. Limiting the analysis to these risk-associated terms substantially narrow the spectrum of physical risk discourse, identifying only about 10% of the initial pool of detectable negative physical sentences that show significant impacts on stock returns.

More importantly, we run validation tests in the form of horse races to compare these measures. Sautner et al. (2023) and Li et al. (2023) also investigate the relation between transition measures and carbon emission level/intensity within industries. We suggest that examining their correlation with changes in carbon emissions can provide more insights into firms' mitigation endeavors. Firms effectively engaging in climate actions should display a negative correlation with changes in carbon emissions, signifying either an amplified reduction or a muted increase in carbon emissions compared to their industry peers, which does not necessarily imply a negative correlation with emission levels/intensity, particularly in scenarios where base emissions are extensive. Table IA11 carries out horse race tests between measures from our study and those of Sautner et al. (2023) with respect to carbon reduction effects. We regress the change in scope 1 carbon emissions on transition exposures controlling industry fixed effects. Whereas our measures suggest that adopting renewable energy sources and enabling technologies contribute to emission reductions for both firms and their clients, Sautner et al. (2023)'s measures do not demonstrate a statistically significant correlation.

Similarly, while our physical measures in Table IA12 reveal a reasonable association with both acute disaster and chronic weather event occurrences, the coefficient for the measure from Sautner et al. (2023) is not statistically significant.

In summary, our comparative analysis—in terms of climate discussion detection, sentence categorizations, and horse-race validations—emphasizes the advanced nature of our approach. By objectively capturing intricate contextual nuances, our methodology greatly enhances both the quality and economic interpretability of insights drawn from climate-focused corporate communications. This not only aids stakeholders in making more informed decisions but also enriches our comprehension of the dynamic relationship between corporate strategies and climate change.

6 Is climate change exposure priced in stock markets?

We now proceed to the main part of our empirical analysis, seeking to identify whether the valuations of firms mirror their exposure to climate change. Furthermore, we examine whether these effects mainly originate from the discount rate or cash flow channels.

6.1 Impact on firm valuations

First, we investigate the relationship between climate change exposure and firm values by performing the following pooled regressions:

$$FirmValue_{it} = \alpha + \sum_{k} \beta^k Climate\ exposure_{it}^k + \gamma \mathbf{X}_{it} + \delta_i + \delta_t + \epsilon_{it}. \tag{4}$$

We consider two measures for firm values: the contemporaneous monthly returns and the logarithm of Tobin's Q in the same quarter as the earnings call.[26] $\mathbf{X}_{it}$ contain size, book-to-market ratio, ROA, investment rate, leverage, tangibility, WW index, and sales growth. Standard errors are clustered by industry and year. We also incorporate firm and year-month (year-quarter) fixed effects for contemporaneous returns (Tobin's Q).

[Insert Table 7 here]

Table 7 Panel A reports the estimation results for transition climate change exposures. The results align well with our expectations. On one hand, the three *Mitigation efforts* measures are significantly and positively correlated with firm values. An increase of one standard

[26] Specifically, we match the returns from March to May of year t with the textual measures on climate exposures extracted from the corresponding quarterly earnings calls in March of year t.

deviation in *Carbon removal*, *Renewable energy*, or *Enabling technology* boosts contemporaneous returns by 0.08%, 0.06%, and 0.09% per month, respectively. Investors appreciate firms' efforts in reducing greenhouse emissions and transitioning towards a lower-carbon economy. They incentivize these actions by purchasing their stocks, resulting in increased contemporaneous returns. Conversely, firms that discuss *Technology challenges* witness a decrease in their values. The inherent challenges make investing in and deploying climate technologies riskier and undermine the predictability of achieving net-zero emission goals. An increase of one standard deviation in *Regulation concerns* leads to a 0.06% decrease in monthly returns. Firms that frequently address this topic face potential threats from stricter climate regulations and, as indicated by our validation test, are more likely to encounter negative climate change-related incidents. Investors tend to divest from stocks involved with environmental issues, leading to a decrease in stock prices. The consistent significant results observed when employing Tobin's Q as the dependent variable reinforce our valuation findings across various transition metrics.

Moreover, given our companion papers, such as Sautner et al. (2023), also construct transition climate measures, we are interested in examining whether our measures capture additional information. To verify this, we include $CCExposure^{Reg}$ and $CCExposure^{Opp}$ from Sautner et al. (2023) as controls in our regression analysis in columns (3) and (4). We find that: (1) their measures reveal no significant correlation with firm values, and (2) our results remain robust. Li et al. (2023) also investigates the valuation effects of their transition risk measure and finds a negative relationship with Tobin's Q when compared within industries; however, this relationship becomes statistically insignificant for the subsequent three years when compared within firms. A key distinction between our studies is that Li et al. (2023) assume all discussions of transition-related keywords in earnings calls relate to transition risk. In contrast, through our sentence-level analysis, we dissect transition measures and identify that a considerable fraction of discussions on transition-related topics actually underline firms' proactive initiatives to tackle climate change. Therefore, these discussions show a significant positive correlation with valuations when comparisons are made within firms.

Panel B applies the same regression analysis to physical climate exposures. A one standard deviation increase in discussion about *Disaster* (*Weather*) corresponds to a decrease of 0.1% (0.12%) in monthly stock returns. A similar negative association is observed when we use Tobin's Q as the dependent variable, confirming that materialized physical risks can result in economic damages and significant value decreases. By incorporating $CCExposure^{Phy}$ from Sautner et al. (2023), the coefficients of our measures persist as negative and significant, thereby reinforcing the merit of our augmented climate word pool and nuanced categorization. Furthermore, we observe that $CCExposure^{Phy}$ displays a positive correlation with firm values, which could imply that this measure might not precisely reflect the information in this domain.

A potential concern regarding our measures is the managers' inclination to portray firm operations in a positive light. To address this, Table IA17 incorporates controls for the firm's overall sentiment and risk measure. While the two controls display significant positive and negative coefficients respectively, the relationships between firm values and both transition and physical climate change exposures persist in their significance.

Then it brings an important question: Do the methodological distinctions in our work translate into the value implications we discern? As discussed in Section 5, our study diverges from related works in its incorporation of more climate-relevant terms and the sentence-level classification of climate-focused discussions. To isolate the effects of each enhancement, we modify one distinction at a time to develop measures, subsequently analyzing their relationship with firm valuations. First, to evaluate the value of our expanded set of climate terms generated through word2vec, we use keywords from both Sautner et al. (2023) and Li et al. (2023) to identify climate-focused sentences. These are then broken down into our seven categories using LDA and FinBERT. Secondly, to probe the effects of our sentence-level categorization, we gather sentences highlighted by our top climate terms, setting aside the granular categorization, and consolidate them into an overall climate exposure index.

Upon applying the new measures to regressions in Table IA18, we find that one or two measures derived from their keywords show slight significance, suggesting the benefits of sentence-level decomposition. However, most coefficients are statistically insignificant, highlighting the essential role of precise climate discussion detection. On the other hand, the measure based on our climate sentences becomes insignificant when grouped together, underscoring the significance of our detailed categorization of climate exposure. In the absence of such nuanced differentiation, the impacts would nullify each other, rendering them statistically insignificant or counterintuitive, as observed in prior research. In short, both methodological improvements play important roles in discerning value implications.

6.2 Discount rate vs. cash flow effects

Next, we attempt to identify whether the climate change exposure measures affect firm values through discount rate or cash flow channels. If climate exposures mainly operate through the discount rate channel, we expect them to predict expected stock returns. On the other hand, if they primarily function through the cash flow channel, the exposures should forecast firms' future profitability.

6.2.1 Discount rate channel

This subsection looks at the discount rate channel and estimates the relation between climate change exposure measures and expected stock returns using portfolio sorting, asset pricing factor tests, and Fama-MacBeth regressions.

Portfolio sorting: To calculate yearly exposures, we average the four quarterly climate risk measures for each firm. Based on these year-end exposure measures, we then categorize firms into quintiles and compare the monthly excess returns across these portfolios from July of year t to June of year $t+1$. Firms with the lowest and highest exposure measures make up Portfolio 1 and 5, respectively, and the high-minus-low portfolios comprise long positions in high exposure firms and short positions in low exposure firms. Our sample period ranges from January 2010 to December 2019.

[Insert Table 8 here]

Table 8 reports the value-weighted raw returns on these portfolios. Panel A reports results for transition climate exposure measures. The three effort-related measures appear to negatively predict future returns. The value-weighted returns for the long-short portfolios based on *Enabling technology*, *Renewable energy*, and *Carbon removal* are -0.50%, -0.48%, and -0.47%, respectively. These results are significant at the 1% level, with the exception of *Renewable energy*. The increased attention to climate change in recent years could prompt the implementation of stringent emission reduction regulations, thereby favoring firms that have previously invested in green technologies. As indicated earlier, firms with more discussions around climate efforts assist in mitigating carbon emissions in the following year, either for themselves or their clients. Due to green preferences or the wish to hedge climate risk, investors tend to acquire these stocks, leading to a rise in their prices and a subsequent decrease in future returns. *Technology challenges*, however, does not display a clear return pattern and has not attracted sufficient investor interest. Lastly, the average excess returns increase from 0.46% to 1.00% when moving from the bottom to the top portfolios sorted based on *Regulation concerns*. Firms in the higher portfolios are more likely to be exposed to regulatory risk and thus bear a higher cost of capital.

Contrarily, Sautner et al. (2022), building on Sautner et al. (2023), estimate risk premiums among S&P 500 stocks but find the realized risk premiums for their climate exposures to be statistically insignificant. They use two option-based proxies to further demonstrate that between 2012 and 2014, stocks with greater climate exposure yielded higher expected returns.

These results are attributable to a larger exposure to climate-related opportunities, contrasting the negative premium for our effort-related transition measures. The discrepancies in our empirical findings can be attributed to the differences in the measures we previously discussed.

Panel B shows that returns increase with physical climate exposures. This trend is especially noticeable for portfolios sorted on *Weather*, which indicate an annual return spread of 4.8%. This difference is statistically significant at the 1% level, taking into account Newey-West corrections with 12 lags for the computation of standard errors. The return pattern for *Disaster* is akin, with returns growing from 0.86% to 1.08% as we move from low to high *Disaster*-exposed stocks. This results in a return spread of 0.22% with t-statistics of 1.91. Given their geographic locations and weather-sensitive operations, firms with high physical exposure measures are more prone to face extreme climate events in the future. Hence, investors demand compensation for holding stocks of these susceptible firms.

Moreover, the slightly varying discount rate implications of the two physical measures may stem from their different levels of persistence. *Disaster* represents low-frequency disasters that can significantly impair firms' property and operations but typically last less than a week. In contrast, *Weather* captures extreme weather patterns that may endure for several days to months. These events have a lasting impact on firms, and their probability has risen in recent years due to climate change. As a result, investors perceive exposure to such physical events as more persistent and consequently demand higher expected returns for *Weather* compared to *Disaster*.

Asset pricing factor tests: Table 8 applies asset pricing factor tests to examine if the relationship between returns and climate exposure measures persists after accounting for known risk factors. It reports alphas for portfolio return regressions against the market factor (column 7), the Fama-French three factors plus momentum factor (column 8), the Fama-French five factors (column 9), and the Hou-Xue-Zhang q-factors (column 10).

The factor-adjusted returns remain negative and significant for portfolios sorted on *Carbon removal* and *Enabling technology*, confirming that the return predictability for these two measures cannot be explained by existing risk factors. The alphas for *Renewable energy* become smaller and statistically indistinguishable from zero after controlling for common factors. Panel D of Table IA13 investigates the relationship between *Renewable energy* portfolio returns and the Fama-French five factors. Firms with high *Renewable energy* are less exposed to the market factor as the loadings decrease with *Renewable energy* at the portfolio level. Furthermore, *Renewable energy* negatively loads on the value factor, suggesting that firms with more *Renewable energy* tend to be growth stocks. These significant loadings partially account for the return relationship between *Renewable energy* and returns, thereby reducing the return spread of portfolios.

On the other hand, the alpha turns positive for *Technology challenges*-sorted portfolios in all factor models due to a negative loading on the market factor. Finally, the return premium for *Regulation concerns* remains positive and significant, though the significance disappears when calculating equal-weighted portfolio returns. This outcome is plausible since larger firms often have higher carbon emission levels and are more exposed to regulatory risks related to climate change.

The alphas remain significantly higher when *Weather* is more frequently discussed. The high-minus-low portfolios yield significantly positive risk-adjusted returns, with the monthly alpha for the *CAPM* model being 0.40%, and 0.43%, 0.51%, and 0.36% for the other three models, respectively. Table IA13 reports the alpha and beta estimations for the Fama-French five-factor models. The results indicate that the long-short portfolio returns are predominantly driven by the long side, as demonstrated by the most significant alphas of the high *Weather* portfolios. Conversely, the return spread for *Disaster*-sorted portfolios becomes insignificant after adjusting for size and value factors. As Tables IA15 and IA16 demonstrate, factor-adjusted returns based on *Disaster* are significantly positive for equal-weighted portfolios. This indicates that the *Disaster* premium predominantly stems from smaller firms, which could experience more pronounced impacts during natural disasters than their larger counterparts.

Generally, most of the cross-sectional return spreads across portfolios sorted on our climate exposure measures cannot be attributed to common risk factors.

[Insert Table 9 here]

Fama-MacBeth regressions Table 9 presents Fama-MacBeth regression results as a robustness check, which explores the combined effects of our climate exposure measures while accounting for several firm characteristics. We run monthly cross-sectional regressions of future excess returns on climate exposure measures and then average the coefficient estimates from these cross-sectional regressions over time. To prevent look-ahead bias, monthly returns from July of year t to June of year t+1 are regressed on the textual measures from year t-1. We include industry fixed effects based on Fama-French 49 industry classifications.

The regression results corroborate the findings from portfolio sorting. While profitability positively predicts future returns and the investment rate negatively does so, in accordance with existing literature, these additional variables do not influence the significance of our climate change exposure measures. Firms with higher physical exposure can expect higher returns in the future, while firms with more discussions on mitigation efforts anticipate lower future returns and costs of capital. Hence, our climate exposure measures are not absorbed by other firm characteristics known to predict stock returns.[27]

6.2.2 Cash flow channel

In the next stage of our analysis, we explore the cash flow channel by estimating the following regression:

$$\Delta ROA_{i,t} = \alpha + \sum_k \beta^k Climate\ exposure^k_{i,t} + \gamma \mathbf{X}_{it} + \delta_i + \delta_t + \epsilon_{i,t}, \tag{5}$$

In this model, $\Delta ROA_{i,t}$ represents the difference between the return on assets of firm i in period t and the ROA in the corresponding quarter of the previous year. $Climate\ exposure^k_{i,t}$ denotes the proportion of climate exposure measure k in firm i's transcript at time t. $\mathbf{X}_{it}$ incorporates size, leverage, tangibility, and dividend. We include firm and year-quarter fixed effects to control for unobservable time-invariant firm characteristics and macroeconomic conditions, respectively.

[Insert Table 10 here]

Table 10 Panel A shows that when firms discuss their climate efforts, they experience an increase in profit growth in the current and following quarters. This could be due to rising demand for their carbon-reducing products and services. Conversely, the estimated coefficients for *Technology challenges* and *Regulation concern* are significantly negative in the current quarter, indicating that firms with high exposures may suffer from decreasing demand or even litigation costs related to environmental issues.

Panel B reveals that physical climate exposure measures are significantly and negatively associated with changes in profitability in quarters t and t+1. This suggests that firms dis-

[27] In our portfolio sorting and Fama-MacBeth regressions, we omit observations post-2019, taking into account the significant impact of COVID-19. Furthermore, recent years have seen a remarkable surge in attention and concern regarding climate change, largely spurred by the Biden administration. To ensure our results are not unduly influenced or distorted by extreme events or outliers, we exclude the sample from these recent years. Additionally, we carry out a supplementary Fama-MacBeth regression analysis on the subsample composed of observations from 2019 onwards. We find that the coefficients are insignificant and some reverse in sign, as detailed in table IA14. This result is consistent with findings from Pastor et al. (2022) and Leippold and Yu (2023), suggesting that firms with a stronger commitment to green initiatives have been outperforming due to the unexpected surge in climate concern.

cussing extreme climate events during their earnings calls subsequently experience a decrease in profitability. The coefficient for *Weather* remains significant in quarter t+2, aligning with the understanding that extreme weather events exert a chronic impact on firms, characterized by smaller coefficients with longer persistence, compared to natural disasters. The associations become insignificant when we replace the dependent variable with $\Delta ROA_{i,t+1,t+4}$, which equals the average of ΔROA from $t+1$ to $t+4$ for firm i. This suggests that the decrease in profitability halts within one year.

To sum up, the cash flow effects appear to be transitory, and the discount rate channel seems to play a more critical role in determining the value implications. Firms discussing climate efforts more frequently enjoy a lower discount rate, which in turn increases their firm valuations. On the other hand, firms that exhibit greater exposure to physical measures tend to have a higher discount rate, thereby lowering their present values.

7 Conclusion

We develop a novel approach that integrates LDA, word2vec, and FinBERT to construct firm-level climate change exposure measures with improved objectivity and deeper economic interpretations. Our model incorporates strategies firms use to combat climate change, represented by *Carbon Removal*, *Renewable Energy*, and *Enabling Technology*. These align well with the greenhouse gas emission mitigation hierarchy. The category of *Technology Challenges* illustrates the difficulties encountered in climate technology development, highlighting areas that need attention as they introduce uncertainty in achieving net-zero emission goals. *Regulation Concerns*, another critical measure, encapsulates firms' worries about stringent climate regulations. We also focus on physical exposure, obtaining *Weather* and *Disaster* that represent distinct extreme climate events, each with differing degrees of persistence and implications for value.

A detailed decomposition of these climate measures is instrumental for discovering previously unexplored value implications associated with distinct exposure components. For example, firms with high physical exposure measures are observed to undergo profit losses and stock price declines. They are also expected to yield higher returns, indicating a profound impact of physical risks on market valuations and the cost of equity. Conversely, firms that actively strive to mitigate global warming - through efforts such as reducing carbon emissions, adopting renewable energy sources, or innovating in climate technologies - tend to create a favorable impression among investors. This positive perception can lead to elevated stock prices and, consequently, lower future returns.

References

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. (2020). Climate change concerns and the performance of green versus brown stocks. *Working paper*.

Berg, F., Kölbel, J., Pavlova, A., and Rigobon, R. (2021). Esg confusion and stock returns: Tackling the problem of noise. *Working paper*.

Bingler, J., Kraus, M., Leippold, M., and Webersinke, N. (2023). How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk. *Swiss Finance Institute Research Paper No. 22-01,*.

Bolton, P. and Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2):517–549.

Bolton, P., Kacperczyk, M. T., and Wiedemann, M. (2023). The co2 question: Technical progress and the climate crisis. *Working paper*.

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

Faccini, R., Matin, R., Skiadopoulos, G., and Skiadopoulos, G. (2022). Dissecting climate risks: Are they reflected in stock prices? *Working paper*.

Giglio, S., Kuchler, T., Stroebel, J., and Zeng, X. (2023). Biodiversity risk. *Working paper*.

Hanley, K. W. and Hoberg, G. (2019). Dynamic Interpretation of Emerging Risks in the Financial Sector. *The Review of Financial Studies*, 32(12):4543–4603.

Hong, H., Li, F. W., and Xu, J. (2019). Climate risks and market efficiency. *Journal of Econometrics*, 208(1):265–281. Special Issue on Financial Engineering and Risk Management.

Hsu, P.-H., Li, K., Tsou, C.-Y., and Tsou, C.-Y. (2022). The pollution premium. *Journal of Finance*, Forthcoming.

Huynh, T. D. and Xia, Y. (2021). Panic Selling When Disaster Strikes: Evidence in the Bond and Stock Markets. *Management Science*, Forthcoming.

Kölbel, J. F., Leippold, M., Rillaerts, J., and Wang, Q. (2021). Ask BERT: How Regulatory Disclosure of Transition and Physical Climate Risks affects the CDS Term Structure. Swiss Finance Institute Research Paper Series 21-19, Swiss Finance Institute.

Kruttli, M. S., Roth Tran, B., and Watugala, S. W. (2021). Pricing Poseidon: Extreme Weather Uncertainty and Firm Return. *Available at SSRN: 3284517*.

Lanzi, E., Verdolini, E., and Haščič, I. (2011). Efficiency-improving fossil fuel technologies for electricity generation: Data selection and trends. *Energy Policy*, 39(11):7000–7014. Asian Energy Security.

Leippold, M. (2023). Sentiment spin: Attacking financial sentiment with gpt-3. *Swiss Finance Institute Research Paper No. 23-11*.

Leippold, M. and Yu, T. (2023). The green innovation premium: Evidence from u.s. patents and the stock market. *Swiss Finance Institute Research Paper No. 23-21*.

Li, Q., Shan, H., Tang, Y., and Yao, V. (2023). Corporate Climate Risk: Measurements and Responses. *Available at SSRN: 3433728*.

Lopez-Lira, A. and Tang, Y. (2023). Can chatgpt forecast stock price movements? return predictability and large language models. *Working paper*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'13, page 3111–3119, Red Hook, NY, USA. Curran Associates Inc.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2022). Dissecting green returns. *Working paper*.

Röder, M., Both, A., and Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, WSDM '15, page 399–408, New York, NY, USA. Association for Computing Machinery.

Sautner, Z., van Lent, L., Vilkov, G., and Zhang, R. (2022). Pricing climate change exposure. *Management Science*, Forthcoming.

Sautner, Z., van Lent, L., Vilkov, G., and Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, n/a(n/a).

Wallach, H., Mimno, D., and McCallum, A. (2009). Rethinking lda: Why priors matter. In Bengio, Y., Schuurmans, D., Lafferty, J., Williams, C., and Culotta, A., editors, *Advances in Neural Information Processing Systems*, volume 22. Curran Associates, Inc.

Yang, Y., UY, M. C. S., and Huang, A. (2020). Finbert: A pretrained language model for financial communications.

Figure 1: **NLP Procedure to Construct Firm-level Climate Exposure**

This figure illustrates the NLP process employed to create various firm-level climate exposure measures. Initially, we apply LDA to IPCC reports, extracting seed words related to climate change. Next, we train a word2vec model on the full earnings call transcripts dataset to identify words similar to the seed words, forming a climate word pool. Finally, we gather sentences from transcripts containing at least one word from the climate word pool and apply both LDA and FinBERT to these climate-focused documents to determine each transcript's loading on distinct climate categories.

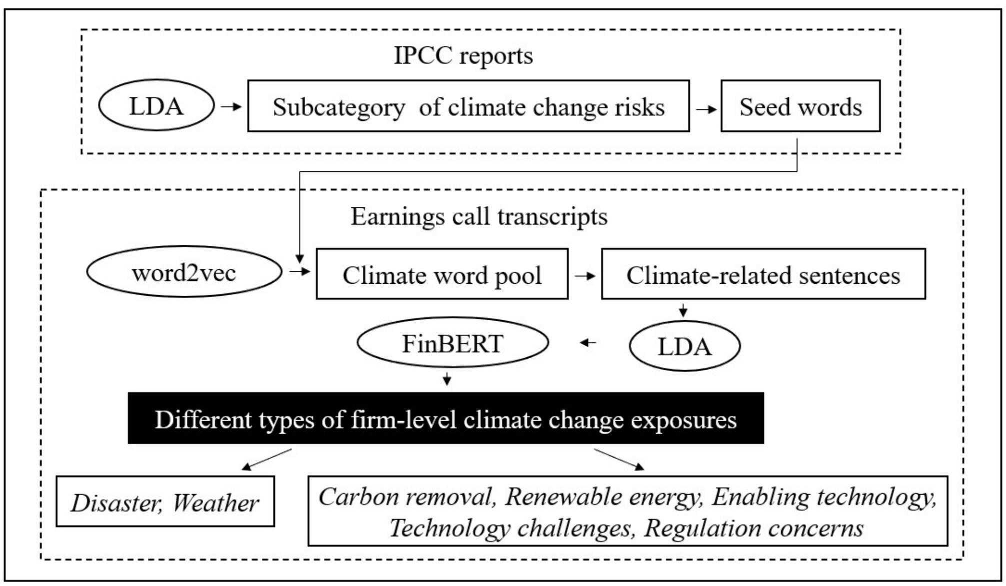


Figure 2: **GHG Emission Mitigation Hierarchy**

This figure depicts the greenhouse gas emission mitigation hierarchy and the placement of our firm-level exposure measures regarding climate efforts in each layer.

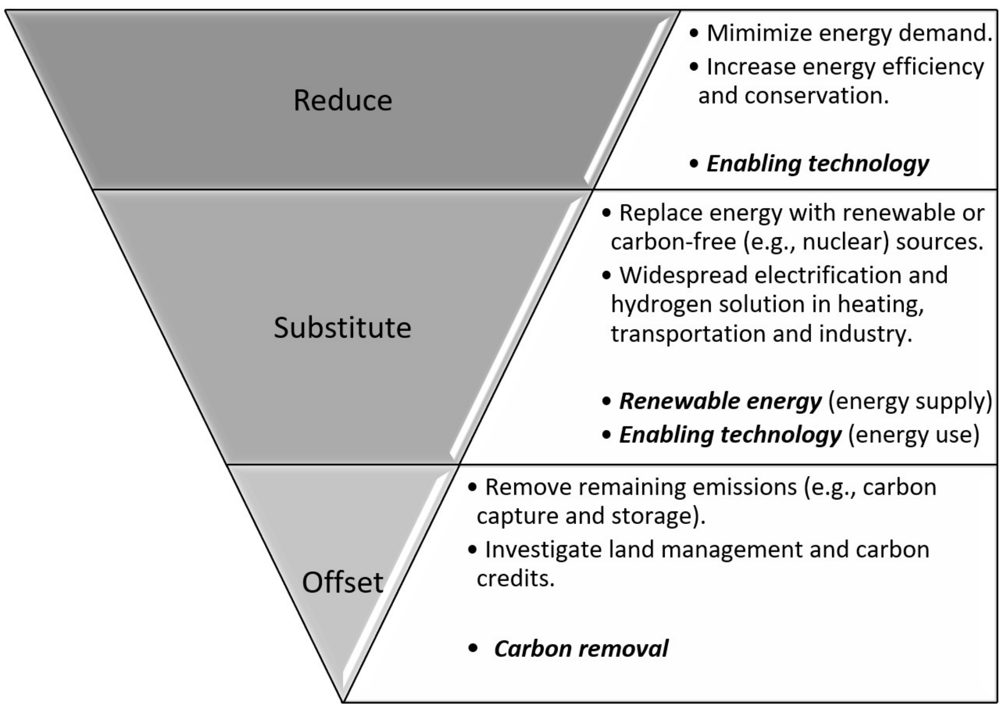
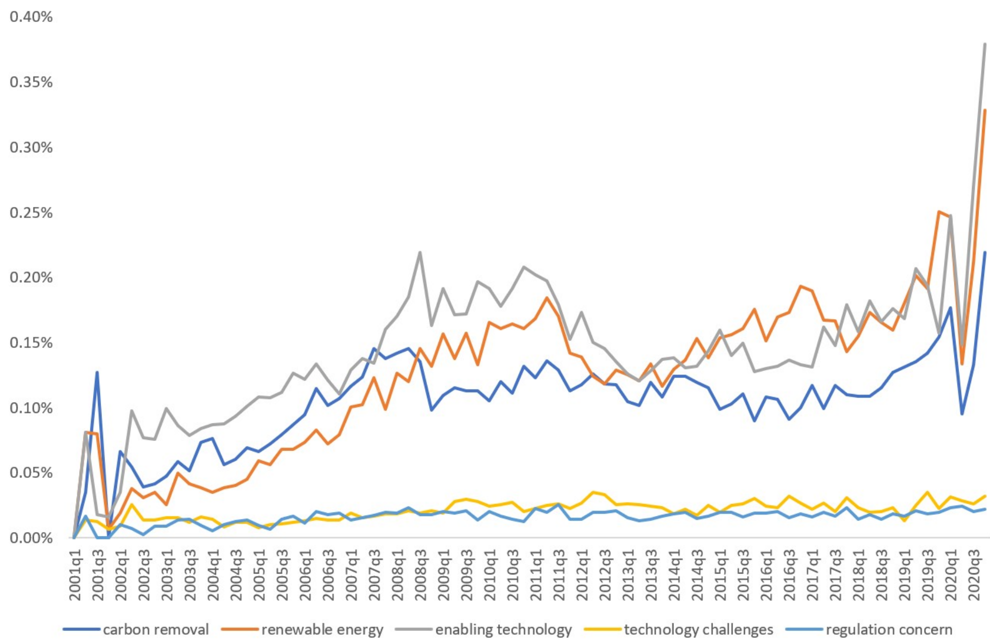
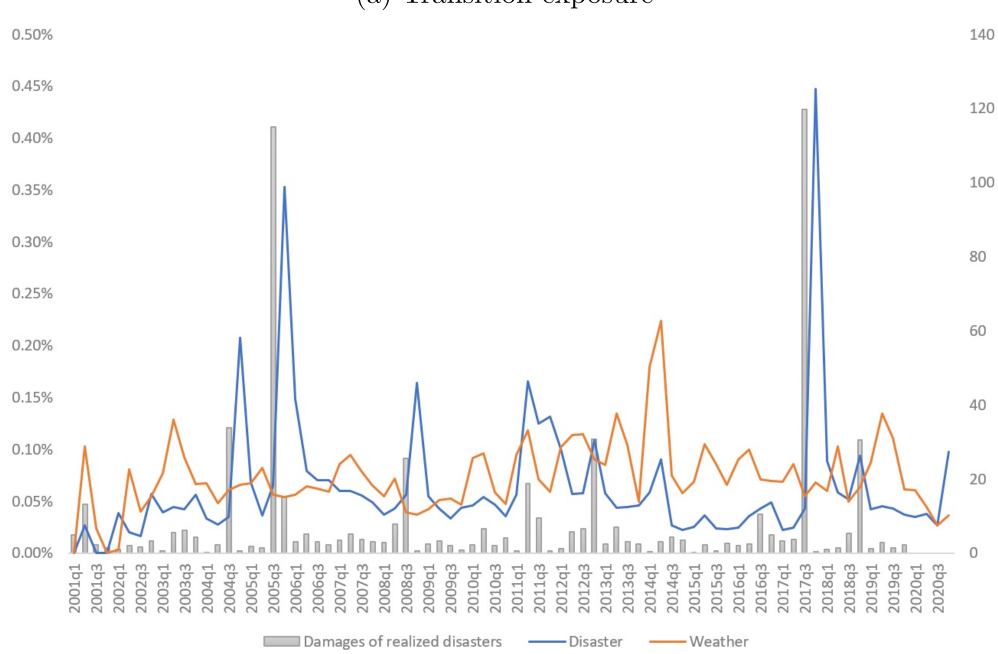


Figure 3: **Average Climate Exposure Measures across Firms over Time**

This figure presents the average textual measures across firms over time. Panel (a) illustrates patterns for transition measures, such as *Carbon removal*, *Renewable energy*, *Enabling technology*, *Technology challenges*, and *Regulation concerns*. Panel (b) displays values for physical climate exposures, including *Disaster* and *Weather*. The grey bars represent the aggregate damages caused by natural disasters in the US each quarter.



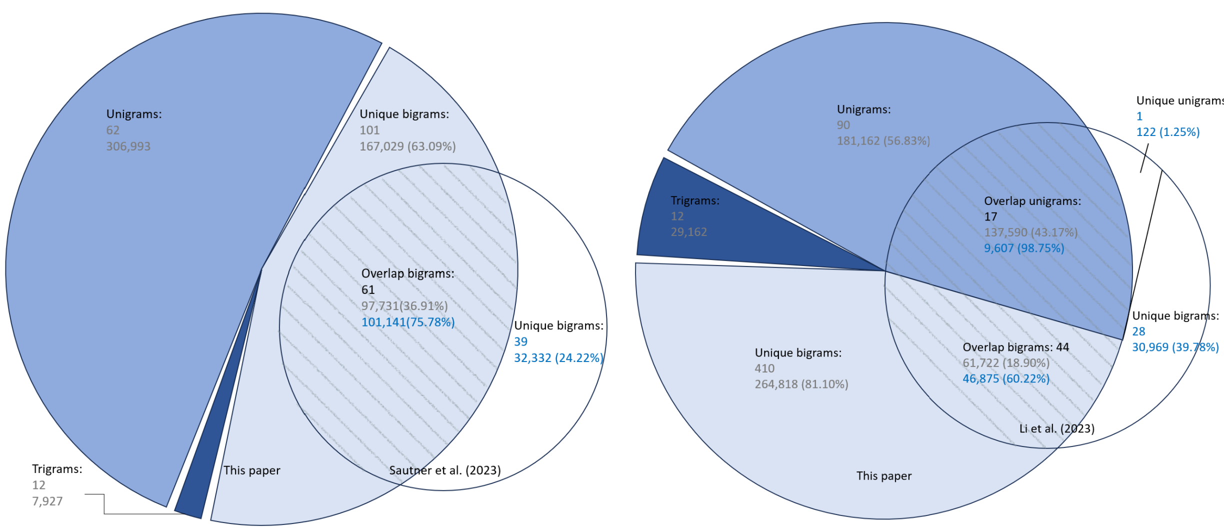
(a) Transition exposure



(b) Physical exposure

Figure 4: **Comparing with Existing Measures: Climate Word Pools**

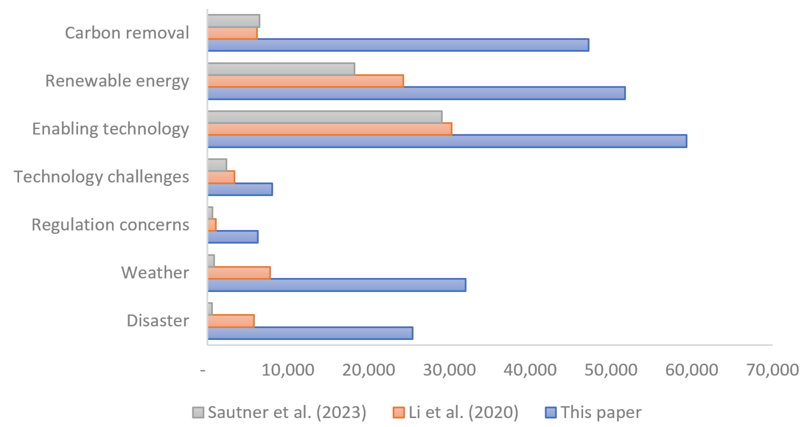
This figure illustrates the Venn diagrams for climate-related words derived from our word pool compared to the top 100 bigrams from Sautner et al. (2023) in Panel (a) and top 90 words from Li et al. (2023) in Panel (b). We chose to include only those terms in our pool with a frequency higher than the 100th (90th) most frequent term in the list of Sautner et al. (2023)(Li et al. (2023)) for a fair comparison. The first row displays the count of unique terms within each group. Frequencies and proportions of overlapping terms are highlighted in gray for our word pool and in blue for theirs.



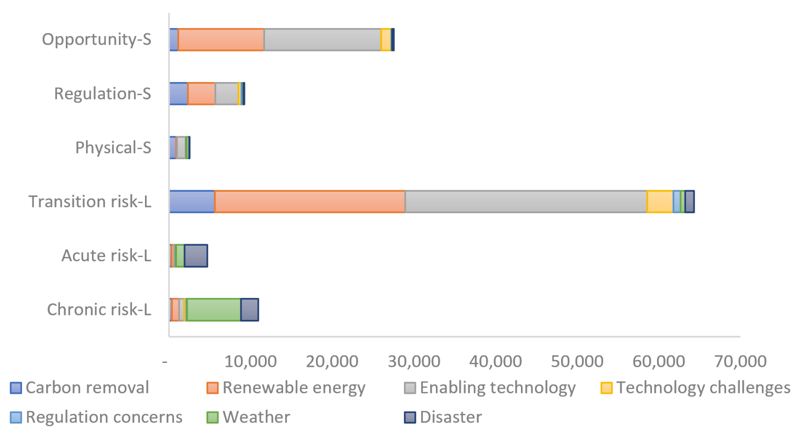
(a) (b)

Figure 5: **Comparing with Existing Measures: Climate Sentence Distributions**

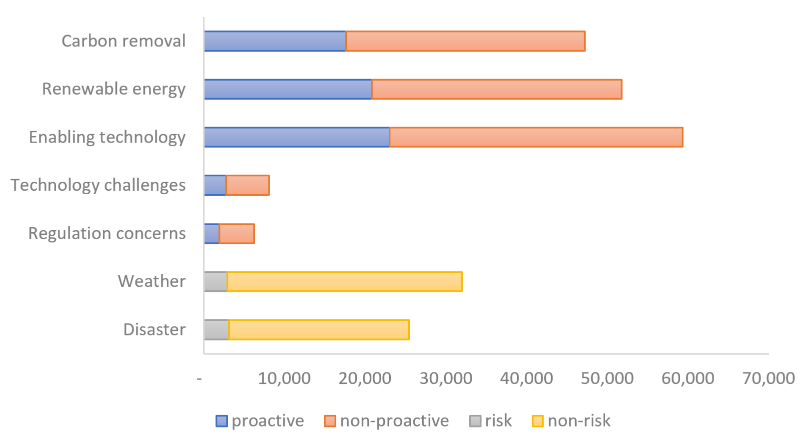
This figure illustrates the distribution of sentences across various categories as classified by our study compared to existing measures. Panel (a) showcases the distribution within our seven categories based on sentences identified using top climate terms from the three papers. Panel (b) presents the categorization of sentences as identified by topic-specific keywords from Sautner et al. (2023) and Li et al. (2023). Meanwhile, Panel (c) depicts the percentage of sentences in each category that incorporate risk synonyms or proactive verbs as defined by Li et al. (2023).



(a)



(b)



(c)

Table 1: **Comparison with Word-Counting Methods**

This table compares the traditional word-counting methods with our approach to construct firm-level climate exposure measures.

	Word-counting methods	**Our method**
Contribution 1: Improved objectivity		
Climate information collection	• Pre-select topics. • Choose keywords for each topic.	• Automatically extract seed words from IPCC reports through LDA. • Expand the word pool using word2vec model.
Weaknesses/advantages	• Introduce human bias and limit the climate information captured.	• Unrestricted by preset topics. • Maximize the detection of climate information to facilitate the discovery of more exposure sources.
Contribution 2: Enhanced economic interpretability		
Contextual understanding	• Count keywords within each preset topic.	• Extend analysis from single words to entire climate-related sentences. • Automatically categorize each sentence based on its topic and sentiment identified by LDA and FinBERT.
Weaknesses/advantages	• The absence of contextual information merges various exposure channels into a single topic. • Value effects for different channels may offset each other when combined and lead to insignificant results.	• Generate more interpretable measures: – Transition exposure contains both firms' risks and efforts. – Subcategory of transition efforts align with emission mitigation hierarchy. – Physical exposure consists of two types with different persistence and market perception. • Each channel exhibits a significant correlation with valuations and carbon reduction in distinct ways.

Table 2: **Descriptions of Text-based Climate Exposure Measures**

This table provides details of the seven climate exposure measures derived from our NLP procedures. Two measures capture information on physical exposures, while the remaining five pertain to transitional topics.

Panel A: Transition exposure

- Mitigation efforts
  - *Carbon removal*: Solutions to reduce or dispose of greenhouse gas emissions from fossil fuels, e.g., biomass with carbon capture and storage.
  - *Renewable energy*: Utilizing renewable energy to cut carbon emissions at the source, e.g., solar, wind, and hydropower.
  - *Enabling technology*: Innovations that assist firms or their clients in enhancing energy efficiency or deploying electrification, e.g., electric vehicles and smart grid.
- *Technology challenges*: Difficulties in green technology development, such as impractical R&D, reduced funding, customer unwillingness, and sensitivity to environmental resources.
- *Regulation concerns*: Uncertainty regarding environmental regulations (e.g., cap and trade system and carbon tariffs) and associated potential threats (e.g., phase-out trends and technological obsolescence).

Panel B: Physical exposure

- *Weather*: Abnormal weather patterns occurring more frequently than rare disasters, with longer-lasting and weaker impacts, including droughts and extreme cold/warm weather.
- *Disaster*: Natural disasters causing acute damage to firms' properties and operations, such as hurricanes, floods, and wildfires.

Table 3: **Summary Statistics**

This table presents summary statistics for variables used in this paper. The variable *Carbon removal (Renewable energy/Enabling technology/Technology challenges/Regulation concerns/Weather/Disaster)* represents the proportion of quarterly earnings call transcripts discussing different categories. *Carbon emission* refers to the annual scope 1 carbon emission level (tons of CO_2 equivalent). *Carbon intensity* is the scope 1 carbon emission scaled by revenue (100 million dollars) at the year level. *Realized disaster* equals one if the firm headquarters is located in a county hit by natural disasters in the past year. The dummy variable *Hurricane (Flood/Drought/Heat)* equals one if the firm experienced hurricanes (floods/drought/heat) in the past year.

Variable	**N**	**Mean**	**Std**	**p5**	**p50**	**p95**
Textual measures						
Carbon removal	155,108	0.11%	0.52%	0	0	0.64%
Renewable energy	155,108	0.13%	0.78%	0	0	0.46%
Enabling technology	155,108	0.15%	0.64%	0	0	0.87%
Technology challenges	155,108	0.02%	0.14%	0	0	0.00%
Regulation concerns	155,108	0.02%	0.12%	0	0	0.00%
Weather	155,108	0.08%	0.32%	0	0	0.53%
Disaster	155,108	0.07%	0.33%	0	0	0.42%
Other variables						
Carbon emission	16,330	16.14	66.16	0.01	0.43	74.08
Carbon intensity	16,330	2.44	8.29	0.02	0.19	12.80
Realized disaster	133,991	14.24%	34.94%	0	0	1
Hurricane	133,991	2.24%	14.79%	0	0	0
Flood	133,991	14.46%	35.17%	0	0	1
Drought	133,991	0.41%	6.36%	0	0	0
Heat	133,991	0.50%	7.02%	0	0	0

Table 4: **Validating Transition Exposure Measures**

This table presents validation test results using various proxies related to transition climate exposure. Panel A uses the number of different types of green patents, while Panel B employs the number, severity, reach, and novelty of negative climate incidents as dependent variables. We average the quarterly textual measures during the year and normalize them to a zero mean and one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, along with industry and year-fixed effects. Standard errors are clustered by industry, and t-statistics are displayed in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)
Panel A: Climate patents and financial constraints				
	Patents			**WW index**
	Carbon abatement	**Renewable source**	**Energy efficiency**	
Carbon removal	0.04***	-0.00	-0.00	-0.59**
	(3.73)	(-0.83)	(-0.12)	(-2.33)
Renewable energy	-0.03**	0.05**	0.02	-0.12
	(-2.24)	(2.23)	(0.92)	(-1.36)
Enabling technology	0.06***	0.01	0.06***	-0.02
	(3.15)	(1.61)	(2.89)	(-0.27)
Technology challenges	0.04**	-0.00	0.01	0.29***
	(2.27)	(-0.49)	(0.57)	(2.95)
Regulation concerns	0.02*	-0.00	0.00	0.07
	(1.96)	(-1.12)	(0.55)	(0.36)
Observations	15,809	15,809	15,809	38,508
R-squared	0.38	0.21	0.46	0.34
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Panel B: Negative climate incidents				
	Number	**Severity**	**Reach**	**Novelty**
Carbon removal	0.02	0.02	0.01	0.02
	(0.80)	(1.10)	(0.63)	(0.87)
Renewable energy	0.02	0.02	0.01	0.00
	(0.95)	(1.64)	(0.89)	(0.31)
Enabling technology	0.05	0.06	0.04	0.05
	(1.39)	(1.61)	(1.22)	(1.35)
Technology challenges	0.00	-0.00	0.01	-0.00
	(0.50)	(-0.28)	(0.72)	(-0.11)
Regulation concerns	0.03***	0.04***	0.03***	0.02***
	(5.62)	(6.86)	(4.35)	(2.86)
Observations	1,997	1,997	1,997	1,997
R-squared	0.43	0.41	0.41	0.35
Control	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Table 5: **Transition Exposures and Carbon Emissions**

This table shows the validation test results using carbon emission data from Trucost. In Panel A, the dependent variables are firms' scope 1 carbon emission intensity, change in intensity, and growth rate of emission level. Scope 1 emissions represent greenhouse gas emissions generated from burning fossil fuels and the production process, which are owned or controlled by the company. Panel B uses the carbon emission growth rate of firms' customers as the dependent variable. We average the quarterly textual measures during the year and normalize them to a zero mean and one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-fixed effects. Standard errors are clustered by industry, and t-statistics are shown in parentheses.*** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

	(1)	(2)	(3)
Panel A: Carbon emissions of firms			
	Carbon intensity$_1$	**Δ Carbon intensity$_1$**	**Δ Carbon emission$_1$**
Carbon removal	0.58***	-4.87**	0.92
	(3.01)	(-2.49)	(1.18)
Renewable energy	2.24***	-14.97***	-0.94***
	(5.40)	(-5.83)	(-2.88)
Enabling technology	-0.67**	1.78	-0.15
	(-2.14)	(1.40)	(-0.29)
Technology challenges	0.60	3.85***	0.27
	(1.35)	(4.52)	(0.45)
Regulation concerns	0.00	4.66**	1.59*
	(0.02)	(2.24)	(1.78)
Observations	13,861	11,380	11,371
R-squared	0.54	0.05	0.50
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Panel B: Carbon emissions of firms' customers			
	Δ Carbon emission$_1$	**Δ Carbon emission$_2$**	**Δ Carbon emission$_3$**
Carbon removal	0.33	-0.65	-0.43**
	(0.63)	(-1.44)	(-2.19)
Renewable energy	-0.19	4.30	-0.40*
	(-1.03)	(1.61)	(-1.89)
Enabling technology	-0.67**	-1.93**	-0.58***
	(-2.33)	(-2.57)	(-3.90)
Technology challenges	-0.29	-0.58	-0.12
	(-0.88)	(-0.66)	(-0.68)
Regulation concerns	-0.42	0.17	-0.05
	(-1.05)	(0.44)	(-0.24)
Observations	5,686	5,684	5,686
R-squared	0.06	0.08	0.20
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Table 6: **Validating Physical Exposure Measures**

This table reports the validation test results using different types of natural disasters. *Realized disaster* is set to one if the firm's headquarters is located in a county affected by natural disasters in the past year. The dummy variable *Hurricane* (*Flood*/*Drought*/*Heat*) equals one if the firm experienced a hurricane (flood/drought/heat) in the past year. Textual measures are normalized to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-quarter fixed effects. Standard errors are clustered by industry. t-statistics are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
Panel A: Disaster					
Realized disaster	0.10***				
	(4.39)				
Hurricanes		0.63***			
		(5.52)			
Floods			0.04**		
			(2.16)		
Droughts				0.02	
				(0.18)	
Heats					0.09
					(1.27)
Observations	122,230	122,230	122,230	122,230	122,230
R-squared	0.09	0.10	0.09	0.09	0.09
Control	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes	Yes
Panel B: Weather					
Realized disaster	-0.02				
	(-1.46)				
Hurricanes		-0.02			
		(-0.67)			
Floods			-0.01		
			(-0.33)		
Droughts				0.24**	
				(2.09)	
Heats					0.06*
					(1.70)
Observations	122,230	122,230	122,230	122,230	122,230
R-squared	0.10	0.10	0.10	0.10	0.10
Control	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes	Yes

Table 7: **Firm Value and Climate Exposures**

This table reports regressions of contemporaneous excess returns and the logarithm of Tobin's Q on quarterly climate exposure measures. Textual measures are normalized to a zero mean and one standard deviation. $CCExposure^{Phy}$, $CCExposure^{Opp}$, and $CCExposure^{Reg}$ are textual measures constructed by Sautner et al. (2023). Other control variables include the natural logarithm of market capitalization, the natural logarithm of book-to-market ratio, ROA, investment, leverage, tangibility, WW index, sales growth, as well as firm and time-fixed effects. Standard errors are clustered by industry and year, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) **Ret**	(2) **Tobin's Q**	(3) **Ret**	(4) **Tobin's Q**
Panel A: Transition exposure				
Carbon removal	0.08**	0.08	0.07*	0.06
	(2.21)	(0.29)	(1.94)	(0.24)
Renewable energy	0.06**	0.51***	0.05*	0.35*
	(2.70)	(2.72)	(2.08)	(1.70)
Enabling technology	0.09**	0.70***	0.08**	0.61**
	(2.71)	(2.76)	(2.22)	(2.42)
Technology challenges	-0.06*	-0.21**	-0.07*	-0.22**
	(-1.95)	(-2.25)	(-1.98)	(-2.28)
Regulation concerns	-0.06*	-0.16	-0.05*	-0.18
	(-1.97)	(-1.47)	(-1.95)	(-1.58)
$CCExposure^{Opp}$			0.03	0.29
			(0.86)	(1.20)
$CCExposure^{Reg}$			0.01	0.07
			(0.37)	(0.39)
Observations	362,979	124,052	344,351	118,606
R-squared	0.24	0.77	0.24	0.77
Control	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
YearMonth FE	Yes	No	Yes	No
YearQtr FE	No	Yes	No	Yes
Panel B: Physical exposure				
Weather	-0.12***	-0.21**	-0.12***	-0.23**
	(-3.52)	(-2.02)	(-3.83)	(-2.15)
Disaster	-0.10***	-0.23**	-0.10***	-0.23**
	(-4.11)	(-2.29)	(-4.45)	(-2.26)
$CCExposure^{Phy}$			0.03*	0.22
			(1.85)	(1.60)
Observations	362,979	106,357	344,351	103,358
R-squared	0.24	0.79	0.24	0.79
Control	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
YearMonth FE	Yes	No	Yes	No
YearQtr FE	No	Yes	No	Yes

Table 8: **Portfolio Sorting and Asset Pricing Factor Tests**

This table reports the value-weighted average returns for five portfolios sorted on the five transition measures in Panel A and the two physical climate exposure measures in Panel B, relative to their Fama-French 49 industry peers. Portfolios are rebalanced at the end of each June, and monthly returns from July of year t to June of year $t+1$ are matched to the yearly textual measures in year $t-1$. The sample spans from July 2009 to December 2019. Columns (7)-(10) present asset pricing factor tests for the High-minus-Low portfolios. We perform time-series regressions of portfolio returns on the market factor in column (7), on the Fama-French three factors plus the momentum factor in column (8), on the Fama-French five factors in column (9), and on the Hou-Xue-Zhang q factors in column (10). t-statistics based on standard errors using the Newey-West correction for 12 lags are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) **Low**	(2) **2**	(3) **3**	(4) **4**	(5) **High**	(6) **HML**	(7) **CAPM**	(8) **FFC**	(9) **FF5**	(10) **HXZ**
Panel A: Transition exposure										
Carbon removal	0.62	0.48	0.86**	0.18	0.14	-0.47***	-0.42**	-0.44**	-0.39**	-0.45**
	(1.39)	(1.05)	(2.22)	(0.41)	(0.29)	(-2.76)	(-2.11)	(-2.22)	(-1.98)	(-2.32)
Renewable energy	1.60***	0.90***	0.98***	0.75***	1.12***	-0.48	0.03	-0.08	-0.14	-0.20
	(3.77)	(3.50)	(3.15)	(3.71)	(5.44)	(-1.47)	(0.09)	(-0.28)	(-0.43)	(-0.81)
Enabling technology	1.47***	1.36***	1.12***	1.40***	0.97***	-0.50***	-0.48**	-0.38**	-0.36**	-0.31*
	(5.45)	(4.28)	(3.71)	(4.59)	(3.47)	(-2.66)	(-2.55)	(-2.24)	(-2.06)	(-1.68)
Technology challenges	1.17***	1.32***	1.19***	1.13***	1.02***	-0.15	0.37	0.28	0.26	0.26
	(3.22)	(4.96)	(4.25)	(4.75)	(5.04)	(-0.58)	(1.62)	(1.57)	(1.31)	(1.37)
Regulation concerns	0.46	0.83**	1.36***	0.60**	1.00**	0.54**	0.62**	0.77***	0.79***	0.91***
	(1.28)	(2.09)	(4.36)	(2.07)	(2.41)	(2.05)	(2.44)	(2.89)	(3.41)	(3.85)
Panel B: Physical exposure										
Weather	0.90***	0.88***	1.05***	1.14***	1.30***	0.40***	0.40**	0.43***	0.51***	0.36**
	(3.64)	(3.82)	(4.48)	(3.93)	(5.22)	(2.65)	(2.34)	(2.69)	(3.25)	(2.10)
Disaster	0.86***	0.73***	0.94***	1.02***	1.08***	0.22*	0.28*	0.22	0.25	0.20
	(3.12)	(2.80)	(3.72)	(3.20)	(3.95)	(1.91)	(1.76)	(1.37)	(1.44)	(1.19)

Table 9: **Fama-MacBeth Regressions**

This table presents Fama-MacBeth regressions of individual stock returns on climate exposure measures. We match monthly returns from July of year t to June of year $t+1$ on the textual measures in year $t-1$. Control variables include the natural logarithm of market capitalization, the natural logarithm of book-to-market ratio, ROA, and investment rate. All independent variables are normalized to a zero mean and one standard deviation. t-statistics based on standard errors using the Newey-West correction for 12 lags are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
			Excess return		
Carbon removal	-0.14***	-0.08*			-0.08**
	(-4.04)	(-1.89)			(-1.99)
Renewable energy	-0.04*	-0.02			-0.02
	(-1.68)	(-0.62)			(-0.58)
Enabling technology	-0.10**	-0.06*			-0.06*
	(-2.36)	(-1.83)			(-1.77)
Technology challenges	-0.01	-0.02			-0.02
	(-0.28)	(-0.77)			(-0.85)
Regulation concerns	0.04	0.03			0.03
	(1.27)	(0.95)			(0.78)
Weather			0.06**	0.05*	0.06**
			(2.16)	(1.83)	(2.10)
Disaster			0.06**	0.06**	0.06**
			(2.42)	(2.26)	(2.34)
logME		0.19*		0.19*	0.19*
		(1.89)		(1.95)	(1.90)
log B/M		-0.05		-0.05	-0.05
		(-0.78)		(-0.84)	(-0.82)
ROA		0.47***		0.47***	0.46***
		(6.16)		(6.26)	(6.16)
Investment		-0.23***		-0.24***	-0.23***
		(-3.52)		(-3.61)	(-3.49)
Observations	270,757	249,578	270,757	249,578	249,578
R-squared	0.08	0.10	0.08	0.10	0.11
Number of groups	120	120	120	120	120
Industry FE	Yes	Yes	Yes	Yes	Yes
YearMonth FE	Yes	Yes	Yes	Yes	Yes

Table 10: **Cash Flow and Climate Exposures**

This table presents relations between climate exposure measures and cash flows. The dependent variable for columns (1) is ΔROA_t, which equals the ROA of time t minus ROA in the same quarter of the previous year. The dependent variable is $\Delta ROA_{t+1,t+4}$ for column (4), which is the mean value of ΔROA for the subsequent four quarters. Textual measures are normalized to a zero mean and one standard deviation. Control variables include size, leverage, investment, tangibility, and dividend, as well as firm and year-quarter fixed effects. Standard errors are clustered by industry, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) Δ **ROA**$_t$	(2) Δ **ROA**$_{t+1}$	(3) Δ **ROA**$_{t+2}$	(4) Δ **ROA**$_{t+1,t+4}$
Panel A: Transition exposure				
Carbon removal	0.11	-0.08	-0.03	-0.10
	(1.21)	(-0.58)	(-0.29)	(-0.86)
Renewable energy	0.17*	0.14**	0.05	0.02
	(1.84)	(2.43)	(0.58)	(0.21)
Enabling technology	0.03	0.21**	0.09	0.07
	(0.24)	(2.04)	(0.69)	(0.61)
Technology challenges	-0.27*	-0.23*	-0.12	-0.09
	(-1.85)	(-2.00)	(-1.18)	(-1.51)
Regulation concerns	-0.24**	-0.09	-0.07	0.01
	(-2.66)	(-1.47)	(-0.78)	(0.10)
Observations	131,228	132,355	132,329	136,424
R-squared	0.11	0.11	0.12	0.21
Firm FE	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes
Panel B: Physical exposure				
Weather	-0.31***	-0.22***	-0.13**	-0.08
	(-5.30)	(-2.87)	(-2.09)	(-1.46)
Disaster	-0.36***	-0.34**	-0.07	-0.07
	(-4.26)	(-2.49)	(-0.98)	(-1.26)
Observations	131,228	132,355	132,329	136,424
R-squared	0.11	0.11	0.12	0.21
Firm FE	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes

Internet Appendix

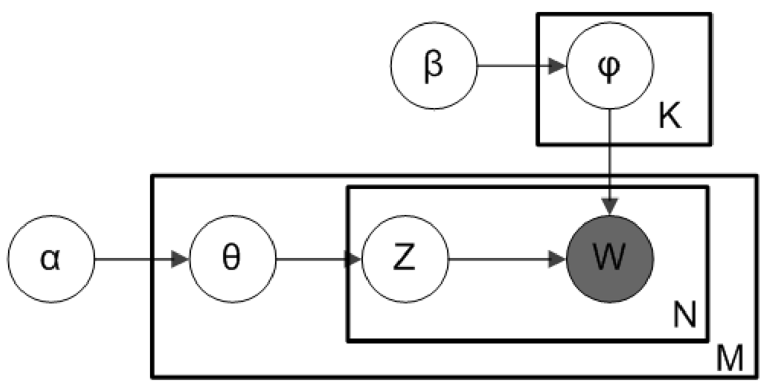
In this Internet Appendix, we provide additional information and supplementary material for our study. Section IA.1 presents a detailed description of the natural language processing techniques we employ, while Section IA.2 enumerates the climate word pool and offers excerpts for categories derived from transcripts. Lastly, Section IA.3 contains supplementary figures and tables related to our textual measures and empirical analysis.

IA.1 Additional details of the natural language processing techniques

IA.1.1 LDA

Latent Dirichlet allocation (LDA)[1] is a generative probabilistic model that considers documents as random mixtures of latent topics, with each topic being a probability distribution over a fixed vocabulary. As in Figure IA1, LDA is a three-level hierarchical Bayesian model. The outer plate represents a corpus containing M documents, while the inner plate indicates

Figure IA1: The graphical model for latent Dirichlet allocation



repeated choices of topic and word for each document. N denotes the number of words in the document, θ_i is the topic distribution for document i, and φ_k is the word distribution for topic k. α and β are the Dirichlet priors on the per-document topic distributions and per-topic word distribution, respectively. Given a corpus of M documents with a vocabulary of W unique words, each document i has a distribution over K topics. θ_i^k represents the share of topic k in document i ($\sum_{k\in K}\theta_i^k = 1$ for each i). All documents share the same set of topics, but each document exhibits those topics in different proportions. φ_k is a W-dimension

[1] See Blei et al. (2003), Blei (2012), Steyvers and Griffiths (2007), and Hoffman et al. (2010).

vector for topic k, where each entry denotes the weight for the corresponding word w in topic k ($\sum_{w \in W} \varphi_k^w = 1$ for each k). The generative process for each document in corpus D can be outlined as follows:

1. Choose $N \sim Poisson(\xi)$.
2. Choose $\theta \sim Dir(\alpha)$.
3. Choose $\varphi_k \sim Dir(\beta)$, where $k \in 1, ..., K$.
4. For each of word w_n, where $n \in 1, ..., N$.
   (a) Choose a topic $z_n \sim Multinomial(\theta)$.
   (b) Choose a word $w_n \sim Multinomial(\varphi_{z_n})$.

LDA's generative process for a single document begins by randomly selecting a distribution over topics (θ_i). Then, for each word in the document, it chooses a topic (z) from the distribution over topics and selects a word (w) from that topic. Words are the only observable variables for texts, and LDA uses them to learn the posterior distribution of latent variables (θ, φ, z) and infer the hidden topic structure, which includes per-document topic distributions and per-topic word assignments. Top probability words within each topic provide insights into the information conveyed by the topic. This inferred topic structure resembles the thematic structure of the corpus.

The Dirichlet distribution functions as a distribution of distributions with a probability density function:

$$p(\theta|\alpha) = \frac{\Gamma(\sum_{i=1}^{k} \alpha_i)}{\Pi_{i=1}^{k} \Gamma(\alpha_i)} \Pi_{i=1}^{k} \theta_i^{\alpha_i - 1}, \theta_i \geq 0, \sum_{i=1}^{k} \theta_i = 1,$$

where α is a k-vector with $\alpha_i > 0$, and $\Gamma(x)$ is the Gamma function. For example, a topic simplex with $k = 3$ forms a triangle, with the three corners corresponding to three distributions over words. For symmetric Dirichlet distributions, small α_i values push the topic distribution to the corners, causing each document to focus on one topic. A large α value centers the distribution around the triangle's center, leading the document to cover all three topics. A similar behavior occurs for a word simplex with Dirichlet prior β. The triangle's centroid represents the uniform distribution over all three words, while the vertices correspond to single words. Typically, α and β are sparse, exerting pressure to allocate words in a document to as few topics as possible.

Given a document w, the key inferential problem is to compute the posterior distribution of the proportions of the topics and the corresponding topic for each word in the document, which can be represented as:

$$p(\theta, z, \varphi | w, \alpha, \beta) = \frac{p(\theta, z, \varphi, w | \alpha, \beta)}{p(w | \alpha, \beta)}.$$

The denominator represents the marginal probability of observations, computed by summing the joint distribution over all possible topic structures. However, this calculation is intractable due to the exponentially large number of possibilities.

One solution is to apply variational inference through an Expectation-Maximization (EM) algorithm. This method posits a parameterized family of distributions over the hidden structure and then identifies the member closest to the posterior. The following variational distribution characterizes these distributions:

$$q(\theta, z, \varphi | \gamma, \psi, \lambda) = \Pi_{k=1}^{K} Dir(\varphi_k | \lambda_k) \Pi_{d=1}^{M} (q(\theta_d | \gamma_d) \Pi_{n=1}^{N} q(z_{dn} | \psi_{dn})),$$

where the Dirichlet parameters γ, λ, and the multinomial parameters ψ are the free variational parameters. Starting from an initial (α, β), the algorithm aims to minimize the Kullback-Leibler divergence between the variational distribution and the true posterior. The iterative algorithm repeats the following two steps until the lower bound on the log likelihood converges:

1. E-step: Find (ψ, γ, λ) to maximize the lower bound.
2. M-step: Given the (ψ, γ, λ) in the last step, find (α, β) to maximize the resulting lower bound.

Another solution is Gibbs Sampling, a form of the Markov chain Monte Carlo (MCMC) method. This approach directly estimates the posterior distribution over z while marginalizing out φ and θ. The sampling considers each word in the document in turn, estimating the probability of assigning the current word to each topic, conditioned on the topic assignments of all other words. This process continues until the sampled values approximate the target distributions. The probability of assigning word i to topic k is:

$$P(z_i = k | z_{-i}, w) \propto \frac{C_{w,k}^{WK} + \beta}{\sum_{w=1}^{W} C_{wk}^{WK} + W\beta} \frac{C_{d,k}^{DK} + \alpha}{\sum_{k=1}^{K} C_{d,k}^{DK} + K\alpha}.$$

z_{-i} is the topic assignments of all other words. C_{wk}^{WK} is the number of times word w is assigned to topic k and C_{dk}^{DK} is the number of times topic k is assigned to some words in document d, both of which exclude the current instance i. The intuition is that words are assigned to topics based on the word's likelihood for a topic and the topic's dominance in a document. It is important to note that samples from different Markov chains will yield different estimations of the latent variables. However, if an appropriate number of topics is chosen and the algorithm is iterated a sufficient number of times, topics will remain stable across runs.

Determining topic numbers: Unsupervised methods like LDA may not always guarantee output interpretability. However, generating coherent topics from a human perspective is crucial for our task, as we aim to understand the trends and types of climate change exposures firms face. Therefore, we determine the number of generated topics, K, based on the coherence measure C_v in Röder et al. (2015), which achieves the highest correlation with all available human topic ranking data. The hypothesis behind coherence is that words in a topic with high probabilities tend to co-occur more frequently in documents. In other words, a high coherence measure indicates a strong correlation to human understandability.

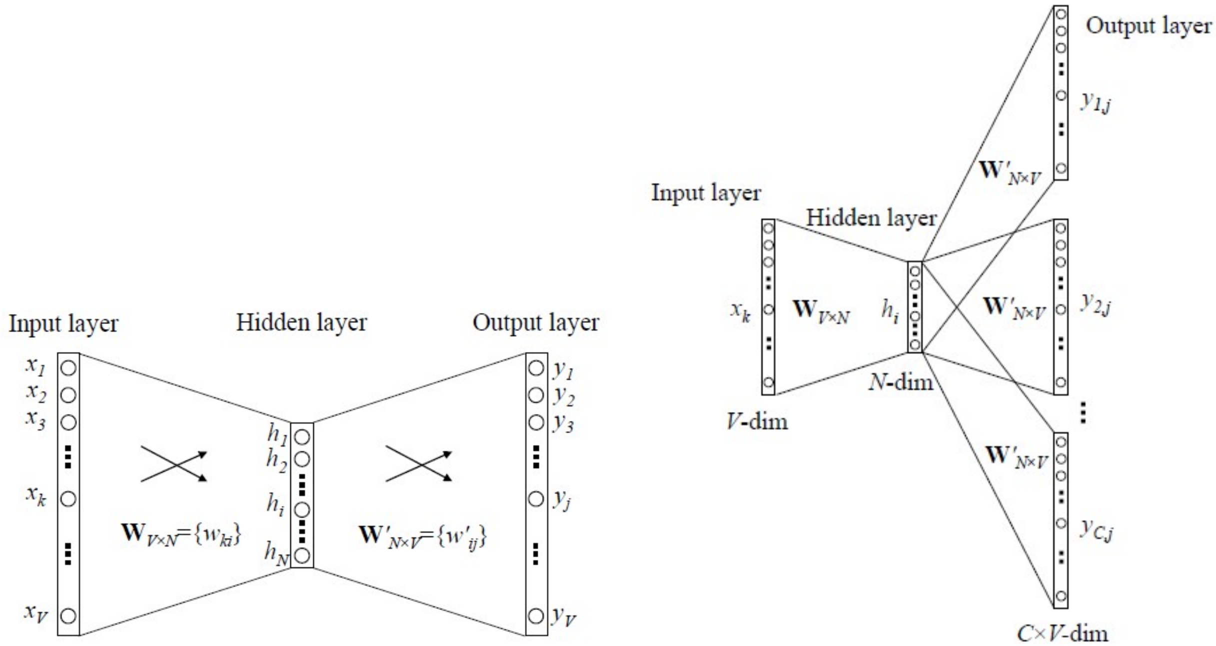
Apart from the number of topics K, LDA has two other hyperparameters: the Dirichlet prior α for topic distributions and the prior β for word distributions for topics. We ask the Gensim library to automatically learn an asymmetric prior from the corpus for both α and β. Wallach et al. (2009) finds that an asymmetric Dirichlet prior over the document-topic distributions has substantial advantages over a symmetric prior.

IA.1.2 Word2vec

Word2vec[2] is a technique that represents words in a vector space, allowing words with similar meanings to have similar representations. This is achieved by training a shallow two-layer neural network with a text corpus to generate a vector for each word in the corpus. Word2Vec can employ one of two architectures: the continuous bag of words (CBOW), which predicts the current word based on context words, or the skip-gram, which predicts context words given the current word. While CBOW can train the model several times faster, the skip-gram produces more accurate results based on large datasets and represents rarely appearing words well. We use the skip-gram model and ask word2vec to produce 300-dimensional vectors for each of the 240,932 words in the vocabulary, including bigrams (two words) and trigrams (three words).

Initially, we consider a simplified model with only one context word for each input word, as shown in Figure IA2a. The input layer is a one-hot $V \times 1$ vector, where V is the size of the vocabulary. The input layer is fully connected to the hidden layer through a weight matrix W of dimensions $V \times N$, where N is a hyperparameter representing the size of the

[2] See Mikolov et al. (2013a), Mikolov et al. (2013), Goldberg and Levy (2014), Rong (2016)



(a) One context per word example (b) The skip-gram model

Figure IA2: Word2vec model

word embeddings. The activation function for hidden layer units is linear, directly passing the weighted sum of inputs to the next layer. The output layer is connected to the hidden layer by another weight matrix, $W'N \times V$. Each neuron in the output layer uses a softmax function to convert uj to y_j, generating the posterior distribution of words. (i.e., y_j on the output layer can be regarded as a probability score of the corresponding word that is the neighbor of the input word). We let v_{w_I} (i.e., input vector) be the I-th row of matrix W, and let v'_{w_j} (i.e., output vector) be the j-th column of matrix W':

$$u_j = v'^T_{w_j} h = v'^T_{w_j}(W^T x) = v'^T_{w_j} v_{w_I},$$

$$p(w_j|w_I) = y_j = \sigma(u_j) = \frac{exp(u_j)}{\sum_{j'=1}^{V} exp(u_{j'})}.$$

v_{w_I} is the N-dimension vector of the associated word of the input layer and is exactly the vector representation we want to obtain for each word in the vocabulary by word2vec. To obtain v_{w_I}, this neural network maximize the probability of observing the actual output word w_O given the input context word w_I with respect to the weights. The loss function is $E = -logp(w_O|w_I)$. Using a stochastic gradient descent, the weight updating equation for hidden-output layer can be written as:

$$w_{ij}^{\prime new} = w_{ij}^{\prime old} - \eta \frac{\partial E}{\partial w_{ij}'},$$

in which $\eta > 0$ is the learning rate and:

$$\frac{\partial E}{\partial w_{ij}'} = \frac{\partial E}{\partial y_j}\frac{\partial y_j}{\partial u_j} \cdot \frac{\partial u_j}{\partial w_{ij}'} = EI_j' \cdot h_i.$$

The weight updating equation for input-hidden layer is:

$$w_{ij}^{new} = w_{ij}^{old} - \eta \frac{\partial E}{\partial w_{ij}},$$

in which

$$\frac{\partial E}{\partial w_{ki}} = \frac{\partial E}{\partial h_i}\frac{\partial h_i}{\partial w_{ki}} = \sum_{j=1}^{V} \frac{\partial E}{\partial u_j}\frac{\partial u_j}{\partial h_i}\frac{\partial h_i}{\partial w_{ki}} = \sum_{j=1}^{V} EI_j' w_{ij}' \cdot x_k = EH_i \cdot x_k.$$

Iteratively updating the parameters through all input-context word pairs will eventually stabilize the weight matrix.

We then consider the case for which one input word has C context words (Figure IA2b), rather than just one. In this case, the loss function becomes:

$$E = -logp(w_{O,1}, ..., w_{O,C}|w_I) = -log\Pi_{c=1}^{C} \frac{exp(u_c, j_c^*)}{\sum_{j'=1}^{V} exp(u_{j'})},$$

in which j_c^* is the index of the actual c-th output context word. The only difference in the deviation is that the derivative of E with respect to u should be summed across all context words in the output layer:

$$EI_j^{skip-gram} = \sum_{c=1}^{C} EI_{c,j}.$$

However, doing these computations, especially for output vectors, is very expensive, which makes it impractical to learn though large training corpora. Using hierarchical softmax or sampling helps mitigate this difficulty. For our purposes, we choose negative sampling, for which we only update a sample of output vectors per iteration. The loss function is changed to:

$$E = -log\sigma(v_{w_O}^{\prime T} v_{w_I}) - \sum_{w_j \in W_{neg}} log\sigma(-v_{w_j}^{\prime T} v_{w_I}),$$

in which $W_{neg} = \{w_j | j = 1, ..., K\}$ is the set of words that are not the context words of the input words (i.e. negative samples), which are sampled based on a noise probabilistic distribution. The objective function aims to maximize the probability of the actual context words while minimizing the probability of the negative samples, given the input words. As we only need to do computations for words that belong to $w_O \cup W_{neg}$ instead of all words in the vocabulary, it will significantly improve computational efficiency.

Cosine similarity between vectors can easily indicate the semantic similarity between the words represented by these vectors. The cosine of two non-zero, n-dimension vectors can be represented as:

$$cos(X, Y) = \frac{XY}{||X|| \cdot ||Y||} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sqrt{\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i^2}},$$

where X_i and Y_i are components of vectors X and Y. The similarity ranges from -1, meaning exactly the opposite, to 1, meaning exactly the same, while in-between values indicate intermediate similarity and 0 indicates orthogonality.

IA.1.3 FinBERT

The Bidirectional Encoder Representations from Transformers (BERT) model is designed to pre-train deep bidirectional representations from unlabeled texts. It does not require pre-processing procedures on documents and can be fine-tuned with just one additional output layer to create state-of-the-art models for classification. We use the FinBERT model from Yang et al. (2020), pre-trained with financial documents including earnings call transcripts, 10Ks, and analyst reports, to implement the sentiment classification task. Unlike traditional algorithms that assume a bag-of-words structure and ignore word order, FinBERT is a contextual model that generates each word's representation based on the entire input text. Yang et al. (2020) documents that FinBERT significantly outperforms the LM sentiment dictionary approach and other machine learning algorithms such as Naive Bayes.

IA.2 Additional details of textual measures

IA.2.1 Climate word pool in transcripts

Seed words - 1:

emission, ghg, emission reduction, environmental, combustion, ozone, ghg emission, greenhouse gas, pollution, thermal

Similar words - 1:

emission, greenhouse gas, emission reduction, reduce emission, greenhouse gas emission, co2 emission, carbon emission, combustion, ghg emission, pollution, pollutant, reduction greenhouse, nox emission, particulate emission, reduce greenhouse gas emission, air emission, nox sox, nitrogen oxide, reduce co2 emission, air pollutant, sox nox, nitrogen oxide emission, gas emission, emission co2, ozone, reduce carbon dioxide emission, harmful emission, emission greenhouse, air pollution, air quality, greenhouse gas reduction, ghg, so2 nox, sulphur dioxide, flue gas, low emission, nox so2, reduce carbon emission, co2 reduction, mercury emission, environmental, carbon dioxide emission, exhaust emission, reduce nox emission, clean air, carbon reduction, reduction carbon emission, reduce greenhouse gas, reduce greenhouse, reduction greenhouse gas emission, emission profile, sulfur dioxide, environmental emission, greenhouse emission, improve air quality, sulfur oxide, sulfur bunker, carbon footprint, hazardous air pollutant, so3, particulate matter, s02, greenhouse gas emission reduction, reduce nox, energy efficiency so2, diesel exhaust, reduction greenhouse gas, nox mercury, carbon injection, carbon emission reduction, emission monitoring, sulfur emission, emission controls, emission control, methane emission, stricter environmental regulation, low carbon emission, diesel particulate, air pollution rule, energy conservation, carbon intensity, environmental footprint, reduce pollution, so2 emission, energy consumption, ash disposal, c02 emission, reduce carbon footprint, sulfur dioxide emission, nitrous oxide, dioxide emission, multi pollutant, air water, carbon intensive, environmental standard, reduction co2 emission, noise emission, stringent emission, coal combustion, environmentally, waste water, acid gas, environmental compliance, ultraclean, pm2.5, co2 emission reduction, powder activate, low nox, clean diesel, nox reduction, carbon monoxide, fugitive emission, nox control, water consumption, emission standard, contaminant, clean water, ambient air, nox burner, hazardous air, reduce energy consumption, depollution, highly pollute, co2 footprint, electrical efficiency, pulverize coal, air toxic, monofill, water usage, exhaust gas, renewable mandate, meet increasingly stringent, nox particulate, environmental regulation, comply environmental, energy savings, water discharge, stringent environmental, ash pond, stricter environmental, environmental pollution, emissions, gas recirculation, coal ash, fuel efficiency, coal fire station, pollution controls, reduce co2 footprint, energy save, coal burn

Seed words - 2:

hfc, hcfc, cfc

Similar words - 2:

hcfc, hfc, cfc, chlorofluorocarbon, hcfcs, cfcs, hydrofluorocarbon, refrigerant

Seed words - 3:

sea level, vegetation, global warming, climate change, deforestation, biodiversity, ecological

Similar words - 3:

climate change, biodiversity, deforestation, global warming, ecological, overfishing, environmental issue, climate protection, vegetation, environmental sustainability, endangered species, sea level, water conservation, endangerment, pollinator, water pollution, environmental social, water scarcity, environmental impact, social environmental, fight against climate change, water resource, address climate change, environmental stewardship, environmental protection, revegetation, environmental concern, environmentalist, soil, wetland, environmental constraint, wildlife, environmental awareness, reforest, plankton, environmentally sensitive, environmental responsibility, greening, aquifer, social economic, water supplies, ground water, minimize environmental impact, economic social environmental, energy policy, energy independence, environmentally sensitive area, carbon neutrality, forest, storm water, siltation, greener, desertification, plastic waste, habitat, environmentally responsible, environmentalism, rainwater

Seed words - 4:

fossil fuel, low carbon, biomass, hydrogen, carbon dioxide, energy, natural gas, hydrocarbon, co2 capture, biofuels, bioenergy, natural resource, co2, carbon, sustainability, sustainable

Similar words - 4:

co2, biomass, carbon dioxide, low carbon, biofuel, carbon, carbon capture, alternative fuel, clean energy, renewable energy source, green hydrogen, green energy, renewable fuel, clean coal, biomass energy, renewable electricity, co2 capture, renewable natural gas, renewable hydrogen, waste energy, carbon fuel, clean burn, coal gasification, renewable diesel, coal gas, nonpollute, cleaner burn fuel, energy waste, biodiesel, coal gas fire, desulphurize, carbon capture sequestration, nuclear energy, renewable source, biomethane, biomass fuel, bioenergy, biorenewable, decarbonization, biomass power, decarbonize, clean energy source, carbon sequestration, cleaner burn

Seed words - 5:

renewable energy, wind energy, geothermal, power plant, grid, wind turbine, turbine, solar, wind power, hydropower

Similar words - 5:

photovoltaic power, [ner:b-quantity] [ner:e-quantity] solar, solar biomass, nuclear power plant, gas power plant, nuclear plant, conventional generation, conventional power plant, wind solar power, coal fire power, thermal generation, geothermal plant, thermoelectric, wind power solar power, gas fire generation, solar installation, [ner:b-quantity] [ner:e-quantity] wind power, hydroelectric power, hydro wind, thermal solar, biomass plant, hydro power plant, coal fire generation, large scale solar, hydropower plant, renewable energy generation, electric generation, solar renewable, offshore wind power, green power, renewable resource, renewable power generation, wind resource, electrical generation, wind power project, hydroelectrical, gas fire plant, biomass fire, geothermal power, wind generator, cogeneration plant, wind tower, hydro generation, photovoltaic plant, shore wind, hydroelectricity, electrical power, electric energy, solar power plant, electric generating, [ner:b-quantity] [ner:e-quantity] offshore wind, ner:b-quantity [ner:e-quantity] wind farm, combine heat power plant, natural gas fire plant, renewable portfolio, gas fire power, waste heat recovery, electric motor, electric vehicle, coal fire power generation, geothermal power plant, renewable energy resource, hydro project, biomass generation, onshore wind farm, thermal power generation, wind turbine generator, fossil power, nuclear generation, hydro nuclear, solar power, hydroelectric plant, thermopower, natural gas fire power, heat power, hydro facility, coal power plant, coal fire power station, hydroplant, solar industry, thermal power station, nuclear hydro, renewable energy sector, offshore wind project, renewable energy certificate, ner:b-quantity [ner:e-quantity] wind turbine, distribute solar, gas fire power station, wind geothermal, coal nuclear, coal fire boiler, waste heat, renewable source energy, biomass facility, electricity transmission, cogeneration facility, solar energy storage, wind solar project, solar power project, offshore wind turbine, thermosolar, cycle gas fire, landfill gas energy, solar solar, ner:b-quantity [ner:e-quantity] solar power, hydroelectric power plant, nuclear coal, hydro electric, geothermal project, nuclear power station, nuclear power generation, hydroelectric facility, gas fire combine, utility scale solar power, electrical vehicle, renewable renewable, electricity gas, biomass power plant, electric power grid, hybrid bus, nuclear fuel, intermittent renewable, conventional power generation, natural gas transmission, electrical power generation, air quality control, ner:b-quantity [ner:e-quantity] coal fire, waste water treatment, renewable rfp, solar rooftop, clean reliable, gas fire power generation

Seed words - 6:

flood, disaster, fire, natural disaster, extreme event, precipitation, drought, weather

Similar words - 6:

flooding, storm, weather event, hurricane, flood, drought, natural disaster, severe weather, earthquake, rain, heavy rain, tornado, snowstorm, typhoon, weather, bad weather, extreme weather, wildfire, hurricane hit, adverse weather, severe storm, tornado, storm hit, rainstorm, severe rain, rainfall, polar vortex, snow storm, ice storm, rain storm, power outage, torrential rain, rain flooding, hurricane wildfire, storm tornado, heavy rainfall, weather related, storm activity, hurricane tropical, hurricane earthquake, hailstorm, flooding event, thunderstorm, tropical storm, forest fire, devastating hurricane, storm flood, hail storm, rain flood, drought condition, flooding occur, catastrophic event, tornadic, blizzard, winter weather, heavy snow, earthquake tsunami, winter storm, derecho, catastrophic storm, severe weather event, weather pattern, tsunami, snow, wind damage, unusual weather, earthquake flood, heavy rain flooding, snowfall, wind storm, bush fire, snow ice storm, extreme weather condition, rainfall flooding, harsh winter, extreme heat, bushfire, hurricanes, hurricane tropical storm, extensive flooding, hot weather, severe drought, cold weather, record rainfall, hurricane season, tornado hail, earthquake hurricane, dry weather, wet weather, precipitation, severe flooding, freezing weather, dry condition, storm surge, freezing temperature, hurricane strike, severe cold, terrible weather, natural calamity, adverse weather event, mudslide, extreme cold, devastating fire, hailstone, weather disruption, record snowfall, snowy weather, weather relate, extreme winter weather, mother nature, quake, ice snow, super storm, devastating flood, northeaster, hurricane storm, historic flooding, devastating storm, devastating earthquake, snow rain, unseasonably cold, weather condition, major snowstorm, convective storm, recent flooding, snow ice, exceptionally warm, severe winter, cold winter, hurricane flood, hot dry, tsunamis, national disaster, [ner:s-norp] earthquake tsunami, heavy snowfall, rain snow, mild winter, active hurricane season, cyclone, inclement weather, noreaster, catastrophic weather, severe drought condition, warm weather, excessive rain, worst drought, mild weather, rainy, drought affect, fire flood, water damage, downpour, electrical outage, bitter cold, storm damage, snowmelt, cool wet weather, cold temperature, frost, storm hurricane, wetter normal, wet winter, abnormally cold, hail, cool rainy, weather northeast, superstorm, shoaling, severe cold weather, natural catastrophe, hot temperature, extreme weather event, extremely cold temperature, torrential, worst weather, windstorm, hurricane hit [ner:b-loc] [ner:i-loc], rain fall, volcanic eruption, lightning strike, wet spring, winter condition, abnormally wet, wettest, snow cold, cold snap, unusual weather pattern, wet condition, wind hail, rainy season, unseasonably warm temperature, come ashore, cause flooding, [ner:b-quantity] [ner:e-quantity] snow, extreme drought, hurricane impact, rainy weather, unseasonable, average rainfall, poor weather, [ner:s-norp] flood, brushfire, hurricane make landfall, snowdrift, unusually cold, crop failure, storm storm, extremely wet, catastrophic weather event, storm related, hail damage, meteorologically, lack rain, tragic earthquake, normal rainfall, snow event, unusually cool, severe snowstorm, storm season, spring weather, major hurricane hit, extreme cold weather, cool weather, significantly warmer normal, due heavy rain, cool spring, downburst, unseasonably wet, hot summer, extremely cold weather, hot dry summer, waterlogging, seismic event, excessive rainfall, extreme winter, excessive heat, monsoon, cold wet, superstorm [ner:s-person], fire claim, wet cool, extremely wet weather, dry weather condition, severe weather condition, harsh weather, damage hurricane, abnormally warm, severe hurricane, unusually hot, wintery, frigid, wild fire, unusually warm, tough winter, weather wise, winter weather condition, worst storm, ner:b-loc [ner:e-loc] hurricanes, unseasonably warm, hurricane damage, snow pack, snow melt, wet cold weather, severe ice storm, non-hurricane, dry winter, cool summer, cold snow, hot dry weather, impact natural disaster, wet season, cyclone season, catastrophe weather, worst winter, flood damage, [ner:s-norp] earthquake, warm winter, extremely warm, wintry weather, cold snowy, rain rain, damage storm, cold rainy, extreme cold temperature, bad weather condition, water shortage, tsunami earthquake, unseasonably cold weather, coldest [ner:s-date], ice condition, flood water, warm temperature, [ner:b-gpe] [ner:e-gpe] earthquake, unseasonal rain, impact hurricane, recent hurricane, driest, due extreme weather, difficult weather condition, hurricane affect, climatology, hurricane related, extremely mild, unusual weather condition, cooler normal, summer weather, drought situation, tsunami hit, lack snow, snow fall, inundation, deep freeze, cold wet weather, fire season, harsh weather condition, warm weather condition, cool wet, warmer weather, volcano, colder winter, unusually severe weather, stormier, rainiest, milder normal, warmest [ner:s-date], abnormal weather condition, coldest, volcano eruption, blustery, hotter normal, due flooding, coldness, poor weather condition, unseasonably warm weather, colder wetter, warmer usual, wet weather condition, snowier, unusually cold weather, unusually wet weather, rockslide, landslide, climatic, cooler weather, cloudburst, unseasonal weather, affect hurricane, horrible weather, unseasonable weather, severe ice, aftermath earthquake, affect earthquake, floodwater, hurricane hurricane, extremely cold, wet summer, severe winter weather, cooler wetter, unusually warm weather, alga bloom, wetter weather

IA.2.2 Excerpts for different categories from transcripts

Carbon removal

<u>RENTECH INC - Construction Materials - 20090210:</u>

Slide 4 is a reminder of what we have to consider in making our fuel with low or no net carbon greenhouse gas emissions. We use renewable carbon that originates from carbon dioxide, that's CO2 in the atmosphere. We too are capturing carbon from the atmosphere for our raw material, we just do it through plants, the things that grow. When we measure our carbon footprint, the majority of our emissions come from the fossil based natural gas and fossil-based electricity use. Eliminating the fossil carbon from energy sources makes a major reduction in our carbon footprint. Carbon dioxide is captured by corn. With their advanced farming techniques, they are capturing and sequestering carbon in the soil. We calculate about 2 pounds of carbon dioxide per gallon of jet fuels actually sequestered in the soil based on past studies, and I think that we shall see increases going well beyond that because farmer techniques have improved further.

<u>CODEXIS INC - Pharmaceutical Products - 20110505:</u>

We're making very strong technical progress in carbon capture with Alstom, and in April we announced an agreement with Alcoa to develop our carbon capture technology to capture CO2 emissions for treatment of bulk site residue, a byproduct of aluminum manufacturing. Alcoa and the US department of energy are funding this program as part of a doe initiative to convert co2 industrial emissions into useful, renewable products. The growth was driven by revenue from our Alstom partnership on carbon capture, as well as continued enzyme evolution service work for pharmaceutical customers.

Renewable energy

ALLIANT ENERGY CORP – Utilities - 20190503:

In the first quarter, we also achieved a major milestone accelerating our progress to a cleaner energy with our upland prairie wind farms being placed in service in late March. Our wind resources performed as expected in the first quarter, achieving an average capacity factor of 36%. This certification recognizes our commitment to environmental stewardship and our collaboration with local communities and landowners. I'm also pleased to share that we were notified that 2 of our recently added generating facilities, the Marshalltown combined cycle natural gas facility and our Dubuque solar project are 2 of 3 finalists for the innovation in sustainable engineering award that will be awarded at the international conference on sustainable infrastructure this fall.

AES CORP (THE) – Utilities – 20200228:

We also signed 2.8 gigawatts of renewable contracts, bringing our backlog to 6.1 gigawatts. This pace puts us on track to nearly triple our portfolio of renewables in operation to 22 gigawatts by the end of 2024 versus 2016. To accelerate a broader adoption of clean energy, we are delivering innovative energy solutions through our leading platforms. We spent the last several years positioning AES to lead the energy transition and cementing our place as a top renewables developer throughout the western hemisphere. In these markets, we see growth in clean energy of 30 gigawatts per year. By capitalizing on our competitive position and the dynamics favoring clean power generation, we have had great success in increasing our backlog of signed ppas. In 2019, we signed long term ppas for 2.8 gigawatts, of which approximately half is wind, 40% is solar and 10% is energy storage.

Enabling technology

FUELCELL ENERGY INC - Electrical Equipment - 20200910:

We highlight many of our well recognized customers currently using our technology platforms, including our extended stack life fuel cells and combined heating power systems that enable micro grids and leverage multiple fuel types, including carbon neutral biofuels. We are working toward commercializing our solid oxide platform capabilities to deliver hydrogen production through highly efficient electrolysis, long duration hydrogen-based energy storage and zero carbon hydrogen power generation, which will support the increasing penetration of intermittent renewable technologies around the world by providing a way to store the power generated by renewables for use when needed. we are excited about the progress we are making sure fulfilling our purpose of enabling the world to live a life empowered by clean energy. Obviously, as we look across the 4 major areas of opportunities, we 're focused on from distributed generation, distributed hydrogen, carbon capture and electrolysis long duration energy storage, we see a market opportunity and momentum really building around hydrogen and the role that hydrogen will play and the transformation of the energy grid.

CATALYTICA ENERGY SYS INC - Machinery - 20030213:

This new system provides a more cost effective and environmentally friendly means to service the heating, cooling and electricity needs of the 650,000 square foot facility. Our technology is designed to enable medium and heavy-duty diesel engines to meet the EPA's mandated 2007 and 2010 emissions standards, which require a 50 to 90 percent reduction in NOx over today's standards. A good example of the large market opportunity for such a product is the Texas emissions reduction plan , under which the state government aims to spend as much as $375m over a two year period to reduce pollution from diesel engines to avoid severe sanctions being imposed by the federal EPA.

Technology challenges

EDP-ENERGIAS DE PORTUGAL SA – Utilities – 20170503:

This was characterized by being the driest first quarter of the year in the Iberian peninsula over the last 5 years, resulting in a 54% reduction in year on year in our hydro generation in Iberia when compared to the very rainy first quarter. Hydro production fell by 50% and coal and gas production increased by 59%. Wind production was also down by 14% with the wind resources in Portugal being in line with long term average versus 16% above the long term average in the first quarter' 16. This ebitda was particularly penalized by very low hydro volumes in Iberia, 48% below historical average and also low wind volumes across key geographies, which stood 7% short of the long term average unusual figure.

MILLENNIUM CELL INC – Electrical Equipment – 20040217:

In addition, we expect that the capability of hydrogen on demand technology and hydrogen fuel cells to enable the design of new devices that have been so far impractical due to the limitations of battery power and difficulties in perfecting direct methanol fuel cells.

ENERGY CONVERSION DEV – Electronic Equipment – 20031015:

The decrease was primarily the result of reduced funding from Chevron Texaco or Texaco ovonic hydrogen systems, ovonics fuel sold which was now self-funded and Texaco ovonic battery systems. The loss from operations increased to $33million, from22$ million in the prior year, because of an almost $ 5.700,000 increased operating loss for ECD segment, which includes machine building, optical memories, fuel cell technology and support services for hydrogen storage provide fragmented computer and primarily due to a higher investment in product logging as company received lower third party funding.

DISTRIBUTED ENERGY SYS CORP – Construction – 20040804:

We find that our challenge in sales and marketing of proton is not so much the reality of our metrics but customer perceptions of the challenges of converting from incumbent methods of hydrogen supply, delivered hydrogen in cylinders, delivered hydrogen in tube trailers and asking them to convert to a significantly different kind of onsite technology. I think that what people are recognizing is that as much as there is a lot of interest in alternative energy, the challenges of making it practical and having it contribute to the energy problems that we are reading about on page 1, is that the gap is quite significant.

Regulation concerns

HUDSON TECHNOLOGIES INC – Wholesale – 20090814:

EPA 's proposed regulations which, when finalized, would limit HCFC production in 2010 to just 80% of the EPA's projected demand thereby creating a 20% supply shortfall. Moreover, the Montréal protocol mandates deeper reductions in HCFC production by 2015 with all production of HCFC 22 phased out by 2020. The proposed legislation is expected to create a mandatory cap and trade system and mandate the phaseout of HFC refrigerants.

MOSAIC CO – Chemicals - 20110331:

These risks and uncertainties include but are not limited to risks and uncertainties arising from changes in environmental and other governmental regulation, including greenhouse gas regulation and implementation of the US environmental protection agency's numeric water quality standards for the discharge of nutrients into Florida lakes and streams.

HUDSON TECHNOLOGIES INC – Wholesale - 20090306:

Such factors include, but are not limited to, possible technological obsolescence of existing products and services, possible reduction of the carrying value of long lived assets, estimates of the useful life of its assets, and potential environmental liability. As a result of this climate bill, R134a and all of the next generation refrigerants may be phased out.

CALGON CARBON CORP – Chemicals – 20080723:

In the meantime, power plants with an April 2015 mass compliance date are continuing to eject powder carbon in order to meet the requirements of the regulation. You don't know who's going to be running the economic policy, so you can imagine that completely uncertainty, who 's going to be within Fernandez' team on the energy policy side.

Weather

ADECOAGRO SA - Food Products - 20120516:

As you know our business in its nature is exposed to weather and we have gone through tough weather conditions during this harvest season. We have measures such as different timing for seeing our growth, diversification of products and geography, and know this technology that contributes to mediate this climatic risk that we are exposed to. In the case of soybeans, we expect final yields to be approximately 5% below the previous last year as a result of the drought experienced during November and December of 2011 in the humid pampas and the recent drought suffered in the northwest of Argentina during February and March. December rains are critical in this region since the corn planted in September and the growth plant flowering, a growth phase in which the plants ' water requirement speak. Therefore, the dry weather cost irreversible damage on early corn planted area.

BIG 5 SPORTING GOODS CORP – Retail - 20030212:

We faced not only what turned out to be a sluggish holiday shopping season, but also warm and dry weather conditions in all of our major markets that were decidedly unfavorable to our business. Unseasonably warm weather conditions clearly influenced product performance. Sales of virtually all cold weather-related products were significantly down which led to sales in our total apparel category being down in the mid to high single digit range for the quarter.

Disaster

CONN'S INC – Retail - 20081126:

Sales were negatively impacted by two hurricanes and mandatory evacuations for most of the gulf coast. The storms resulted in 144 store days lost in the month of September. In addition, sales velocity slowed as soon as the storms entered the gulf of Mexico, and remained negatively impacted until many days after landfall.

HONDA MOTOR CO LTD - Automobiles and Trucks - 20111031:

Revenue decreased by jpy366 billion, down by 16.3% from the second quarter last year to jpy1,885.8 billion, due mainly to decreased revenue in the automobile business, mainly caused by supply chain disruptions from the earthquake. With regard to the impact of flooding in Thailand, we have suspended our production since the 4th of October and from the 8th of October, our automobile factory located in Ayutthaya province has been flooded. We also suspended our motorcycle production from the 11th of October to the 31st of October, due to supply disruption, as some of our suppliers in Thailand have been affected by the flooding.

IA.2.3 Other sentence examples

Sentences mislabeled by the LM sentiment dictionary (Sentiment classified by FinBERT is shown at the end of each sentence).

- Mislabeled as positive:

  1. Year to date gross profit was $737,000, down from $8 million a year ago due to decreased activity in our higher margin well enhancement operation, warm weather impact, startup costs in our water management and dirt hauling, and the effects of lower revenue on our fixed cost base. (negative)
  2. And then on mercury, obviously that's the wild card because the only we have co benefits associated with the scrubbers, but we have no other mercury removal technology installed at the plant. (neutral)
- Mislabeled as negative:

  1. The active locomotive fleet was up 3% for the first quarter of 2016, primarily to help minimize the impact of weather-related challenges, and we had approximately 1,400 locomotives in storage at the end of the first quarter. (positive)
  2. Those two OEMs are building quality product and meeting new emissions targets, and I believe that they will restore share they are losing. (positive)
- Mislabeled as neutral:

  1. Food-service sales of $108.9 million increased 9% from the 2004 quarter as hotter than normal weather in July and August drove sales of replacement equipment, especially ice machines. (positive)
  2. Be aware as well that although California's been inundated with a significant amount of precipitation this year that the bulk of the hydropower that serves that area is in the Pacific Northwest, and that area is actually below its normal precipitation levels for the year. (negative)
  3. As for the hurricanes in Florida, I think weather-related issues clearly resulted in a drop in tonnage a little bit in the September quarter and then into October and November, as those hurricanes had impact. (negative)

Sentences labeled as Regulation in Sautner et al. (2023) but classified differently in our paper.

- Regulation - Mitigation efforts:

  1. This award is a recognition of our efforts to educate consumers and employees regarding the benefits of energy efficiency and promote energy-efficient products.

2. It's great to be part of such an innovative facility, which incorporates an ultra-high concentrator solar PV plant to supply power for the desalination process, providing for reduced operational cost and no harmful gas emissions, a great win for our team based on our strong product performance in corrosion resistance, optimized design for maintenance and the ability to serve the global market.

- Regulation - Regulation concern:

  1. I guess, is permitting that much more difficult, environmental standards that much more difficult than it's holding it back?

  2. Part of the litigation challenging out in California was brought by the poet group and it was entitled poet 1, challenging that the incorporation of biodiesel into the mix of fuel alternatives or options in California did not reduce NOx and therefore was adverse to the California environmental quality act.

Word restriction of "risk" synonyms.

- Sentences with "risk" synonyms but not talking about risk:

  1. It's going to be a long, hot summer for everyone involved, and I want to thank everyone in advance for all the work that no doubt we're all going to be doing this summer to make this year a great year for us.

  2. If we put aside Toshiba, which I don't want to relate to the company itself, of course, no doubt that the calamity that happened in Japan is helping in the growth of green energy in Japan.

- Sentences without "risk" synonyms but talking about risk:

  1. We also experienced a deferral and cancellation of food service equipment orders in the gulf coast region due to widespread hurricane damage.

  2. Be aware as well that although California's been inundated with a significant amount of precipitation this year, the bulk of the hydropower that serves that area is in the Pacific Northwest and that area is actually below its normal precipitation levels for the year.

Word restriction of "proactive" verbs.

- Sentences without "proactive" verbs but talk about proactively addressing transition risk:

  1. The wind energy, we can apply all these platforms to this industry to gain a competitive advantage.

  2. To us, success goes hand in hand with enabling California's clean energy economy, a topic I'm particularly passionate about.

- Sentences with "proactive" verbs but do not talk about proactively addressing transition risk:

  1. Likewise, we anticipate slow and steady project development growth in Europe and emerging technical difficulty as the price of wind energy decreases and becomes most cost competitive for both offshore and onshore installations.

  2. Allete Clean Energy's net income decreased by $400,000 due to lower wind resources added facilities, partially offset by lower operating and maintenance expenses compared to the same period in 2016.

IA.2.4 Linking earnings call transcripts to Compustat

We link the earnings call transcripts with Compustat according to the company names to match textual measures with financial data later. Thomson Reuters (TR) notes that if company A is acquired by company B, all historical transcripts of company A will be assigned to the company B identifier. As a result, using the company name directly provided by TR may not be appropriate. As suggested by Li et al. (2020), we extract company names from the event titles, which remain unchanged for conferences held before the acquisition.

Over 90% of the event titles can be reformatted or transformed into the format "QX 20XX COMPANY Earnings Conference Call" (X represents a number). We remove transcripts with poorly organized event titles. To facilitate matching between company names from transcripts and Compustat, we standardize the names in both datasets using consistent rules, allowing us to capture as many matches as possible. For company names that do not perfectly match, we use Python to identify the top five similar names in Compustat and hire research assistants to manually verify each pair.

IA.3 Supplementary figures and tables

Figure IA3: **Clusters of Climate Seed Words from IPCC Reports**

This figure reports seed words extracted from IPCC reports. Using hierarchical agglomerative clustering, we group them into six clusters to efficiently identify similar words.

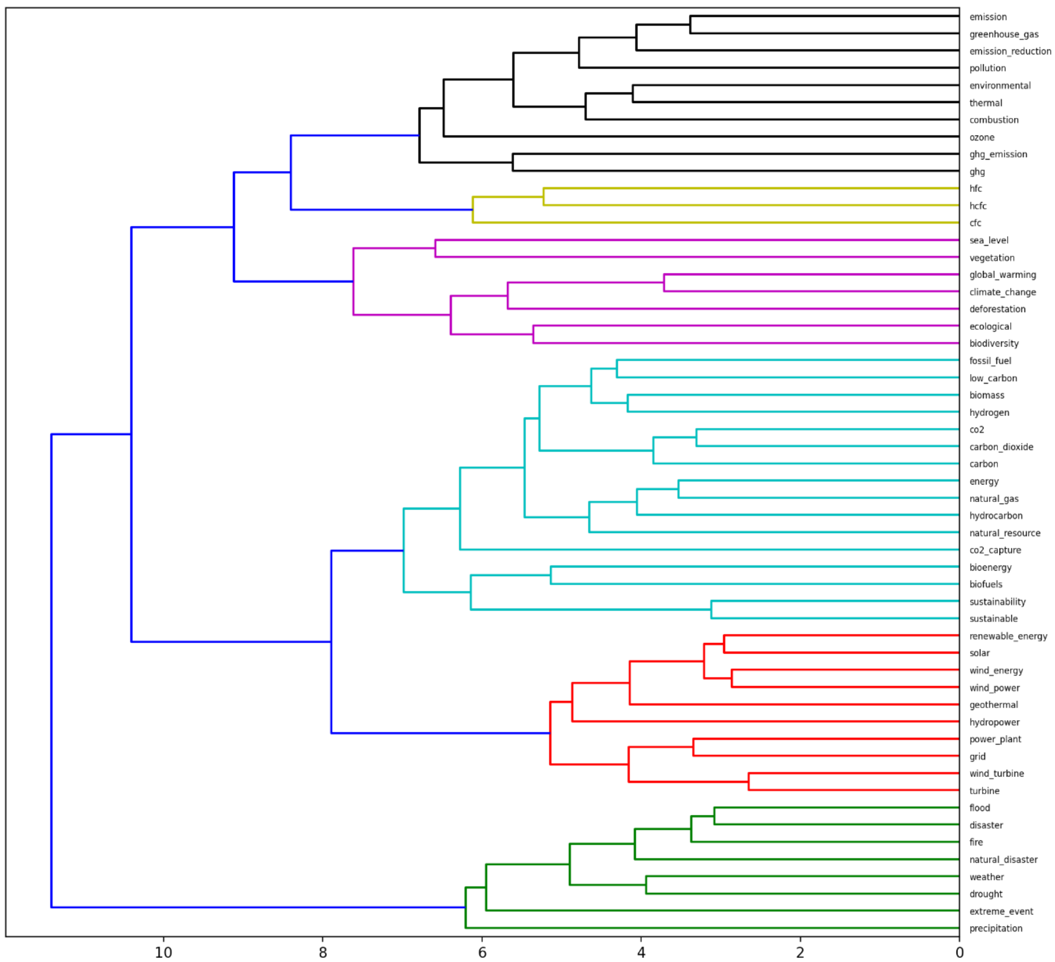
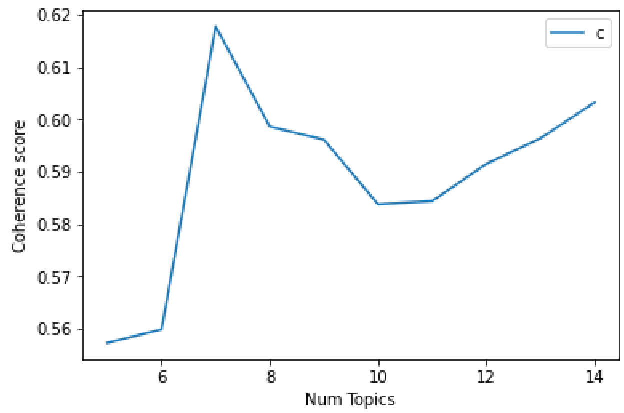
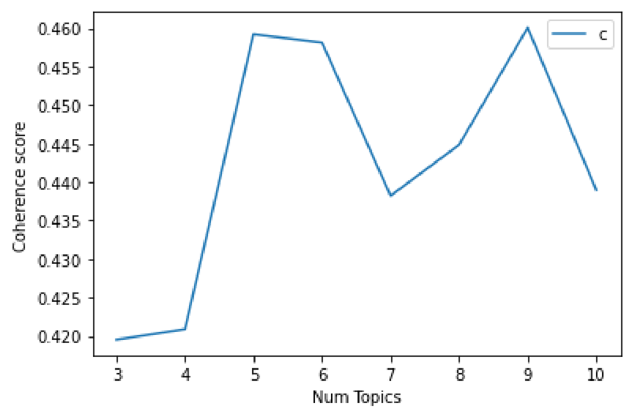


Figure IA4: **Coherence Scores for Different Topic Numbers**

This figure displays the coherence scores associated with different numbers of topics when applying LDA to IPCC reports and earnings call transcripts. High coherence measures correspond to a greater correlation with human understandability. To balance topic interpretability and model performance, we opt for generating seven topics for IPCC reports and five topics for earnings call transcripts.



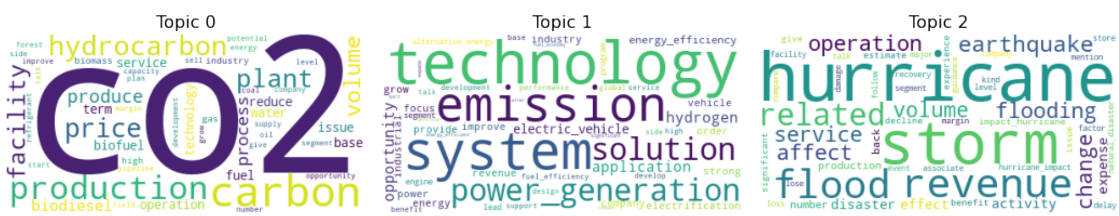
(a) IPCC reports



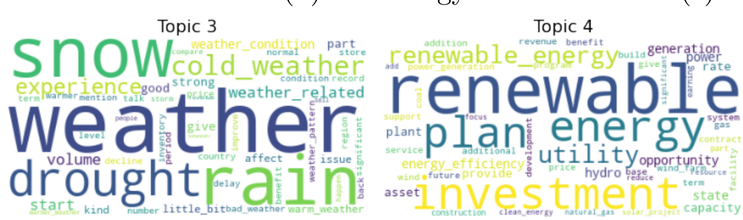
(b) Earnings call transcripts

Figure IA5: **Word Cloud for Topics Generated by LDA in Transcripts**

This figure presents the word cloud for the five topics generated through LDA on climate-related transcripts. Larger words signify higher probabilities within the associated topic.



(a) carbon (b) technology (c) disaster



(d) weather (e) renewable

Table IA1: **IPCC Reports**

This table displays the titles and publication dates of reports from the Intergovernmental Panel on Climate Change (IPCC). It encompasses both Synthesis reports, offering a comprehensive and integrated perspective, and Special reports, concentrating on a particular aspect of climate change.

Titles	Time
Emissions scenarios	2000 March
Land use, land-use change, and forestry	2000 March
Methodological and technological issues in technology transfer	2000 March
TAR Climate Change 2001: Synthesis Report	2001 October
Carbon dioxide capture and storage	2005 March
Safeguarding the ozone layer and the global climate system	2005 March
AR4 Climate Change 2007: Synthesis Report	2007 September
Renewable energy sources and climate change mitigation	2011 April
Managing the risks of extreme events and disasters to advance climate change adaptation	2012 March
AR5 Synthesis Report: Climate Change 2014	2014 October
Global warming of $1.5°C$	2018 October
Climate change and land	2019 August
The ocean and cryosphere in a changing climate	2019 September

Table IA2: **Green Patent Classifications**

This table presents classifications for green patents used in our validation tests, based on the OECD strategy.

Carbon abatement
1. Environmental management:
1.1. Air pollution abatement
1.2. Water pollution abatement
1.3. Waste management
4. Climate change mitigationn - energy:
4.3. Combustion technologies with mitigation potential (using fossil fuels)
5. Capture, storage, sequestration or disposal of greenhouse gases
6. Climate change mitigation - transportation:
6.1.1. Conventional vehicles
8. Climate change mitigation - waste
9. Climate change mitigation - production:
9.1.1. Reduction of greenhouse gas
9.2. Technologies relating to chemical industry
9.3. Technologies relating to oil refining and petrochemical industry
9.4. Technologies relating to the processing of minerals
9.5. Technologies relating to agriculture
Efficiency-improving fossil fuel technologies for electricity generation from Lanzi et al. (2011)
Renewable sources
4. Climate change mitigationn - energy:
4.1. Renewable energy generation
4.4. Nuclear energy
Electrification and energy efficiency
4. Climate change mitigationn - energy:
4.5. Efficiency in electrical power generation, transmission or distribution
4.6. Enabling technologies in energy sector
6. Climate change mitigation - transportation:
6.1.2. Hybrid vehicles
6.1.3. Electric vehicles
6.1.4. Fuel efficiency-improving vehicle design
6.2. Rail transport
6.3. Air transport
6.4. Maritime or waterways transport
6.5. Enabling technologies in transport
7. Climate change mitigation - buildings
9. Climate change mitigation - production:
9.1.2. Process efficiency
9.6. Technologies in the production process for final industrial
9.7. Climate change for sector-wide applications
9.8. Enabling technologies

Table IA3: **Correlation Matrix**

This table presents correlation matrix for our measures along with those from Sautner et al. (2023).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Carbon removal	1										
(2) Renewable energy	0.11	1									
(3) Enabling technology	0.10	0.16	1								
(4) Technology challenges	0.06	0.39	0.21	1							
(5) Regulation concerns	0.37	0.04	0.04	0.04	1						
(6) Disaster	0.03	0.04	0.00	0.03	0.05	1					
(7) Weather	0.06	0.06	-0.01	0.05	0.06	0.09	1				
(8) CCExposure	0.23	0.68	0.54	0.38	0.09	0.03	0.05	1			
(9) CCExposure^{Opp}	0.10	0.65	0.49	0.34	0.03	0.03	0.02	0.86	1		
(10) CCExposure^{Reg}	0.30	0.42	0.31	0.24	0.13	0.02	0.03	0.55	0.33	1	
(11) CCExposure^{Phy}	0.09	0.04	0.08	0.03	0.04	0.03	0.10	0.14	0.06	0.06	1

Table IA4: **Industry Distribution (Top Ten) of Climate Measures**

Fama-French 49 Industries	**Mean**	**Std**	**>0 obs.**	**obs.**
Panel A: Carbon removal				
14 Chemicals	0.81%	1.99%	1454	3627
1 Agriculture	0.53%	1.90%	72	300
27 Precious Metals	0.47%	0.80%	539	1109
29 Coal	0.46%	0.67%	264	512
30 Petroleum and Natural Gas	0.41%	1.03%	2851	7796
28 Non-Metallic and Industrial Metal Mining	0.38%	0.70%	387	936
19 Steel Works Etc	0.28%	0.63%	565	1847
31 Utilities	0.27%	0.59%	1631	5153
17 Construction Materials	0.25%	1.15%	476	2876
39 Business Supplies	0.23%	0.60%	427	1892
Panel B: Renewable energy				
31 Utilities	2.50%	3.09%	3675	5153
29 Coal	0.47%	0.78%	227	512
18 Construction	0.39%	1.09%	482	2135
22 Electrical Equipment	0.32%	0.89%	630	2718
17 Construction Materials	0.11%	0.78%	211	2876
21 Machinery	0.11%	0.58%	547	5794
28 Non-Metallic and Industrial Metal Mining	0.09%	0.34%	116	936
14 Chemicals	0.09%	0.45%	348	3627
25 Shipbuilding, Railroad Equipment	0.09%	0.30%	57	450
20 Fabricated Products	0.08%	0.26%	34	255
Panel C: Enabling technology				
22 Electrical Equipment	1.52%	2.63%	1538	2718
23 Automobiles and Trucks	0.77%	1.34%	1305	2633
20 Fabricated Products	0.69%	1.60%	98	255
21 Machinery	0.53%	1.43%	1996	5794
38 Measuring and Control Equipment	0.33%	0.78%	982	3490
14 Chemicals	0.31%	1.04%	1038	3627
18 Construction	0.28%	1.05%	532	2135
24 Aircraft	0.24%	0.64%	251	1047
37 Electronic Equipment	0.23%	0.76%	2340	12257
19 Steel Works Etc	0.23%	0.58%	502	1847

Fama-French 49 Industries	Mean	Std	>0 obs.	obs.
Panel D: Technology challenges				
31 Utilities	0.25%	0.56%	1460	5153
29 Coal	0.15%	0.38%	114	512
22 Electrical Equipment	0.10%	0.31%	428	2718
21 Machinery	0.06%	0.27%	458	5794
20 Fabricated Products	0.06%	0.28%	20	255
23 Automobiles and Trucks	0.05%	0.20%	260	2633
18 Construction	0.05%	0.21%	176	2135
19 Steel Works Etc	0.03%	0.16%	109	1847
25 Shipbuilding, Railroad Equipment	0.03%	0.13%	28	450
24 Aircraft	0.03%	0.13%	62	1047
Panel E: Regulation concerns				
29 Coal	0.13%	0.35%	107	512
14 Chemicals	0.13%	0.39%	609	3627
27 Precious Metals	0.08%	0.31%	131	1109
28 Non-Metallic and Industrial Metal Mining	0.06%	0.21%	116	936
30 Petroleum and Natural Gas	0.06%	0.31%	712	7796
19 Steel Works Etc	0.06%	0.29%	203	1847
39 Business Supplies	0.04%	0.20%	125	1892
1 Agriculture	0.04%	0.24%	14	300
31 Utilities	0.04%	0.18%	365	5153
2 Food Products	0.04%	0.19%	160	2362
Panel F: Weather				
1 Agriculture	0.75%	0.98%	189	300
28 Non-Metallic and Industrial Metal Mining	0.33%	0.69%	323	936
31 Utilities	0.29%	0.64%	1750	5153
40 Shipping Containers	0.23%	0.47%	194	571
7 Entertainment	0.23%	0.66%	425	2090
29 Coal	0.21%	0.43%	159	512
17 Construction Materials	0.21%	0.50%	732	2876
2 Food Products	0.19%	0.60%	576	2362
43 Retail	0.19%	0.58%	1814	8536
3 Candy & Soda	0.16%	0.37%	99	366
Panel G: Disaster				
1 Agriculture	0.20%	0.62%	64	300
18 Construction	0.18%	0.60%	423	2135
44 Restaurants, Hotels, Motels	0.16%	0.58%	420	2953
25 Shipbuilding, Railroad Equipment	0.16%	0.64%	56	450
17 Construction Materials	0.15%	0.52%	451	2876
33 Personal Services	0.14%	0.55%	257	2075
31 Utilities	0.13%	0.42%	846	5153
30 Petroleum and Natural Gas	0.13%	0.44%	1241	7796
23 Automobiles and Trucks	0.12%	0.63%	246	2633
40 Shipping Containers	0.12%	0.35%	92	571

Table IA5: **Variance Decomposition**

This table provides a variance decomposition of each climate change exposure measure. Regressions are estimated at the firm-yearQtr level. In Panel A, the table reports the incremental R^2 from adding a specific fixed effect. In Panel B, the table decomposes the variation into a firm fixed effect and a residual component. Columns (1) to (7) report the results for categories *Carbon removal*, *Renewable energy*, *Enabling technology*, *Technology challenges*, *Regulation concerns*, *Weather*, *Disaster* respectively.

	(1) CR	(2) RE	(3) ET	(4) TC	(5) RC	(6) Weather	(7) Disaster
Panel A: Incremental R^2							
YearQtr Fixed Effect	0.19%	0.36%	0.37%	0.15%	0.06%	0.97%	4.03%
Industry Fixed effect	9.25%	32.34%	12.59%	10.09%	4.32%	6.92%	2.02%
Industry × Year Fixed Effect	1.26%	4.61%	0.97%	0.75%	0.61%	1.12%	2.16%
"Firm-level"	89.30%	62.69%	86.07%	89.01%	95.01%	90.99%	91.79%
Sum	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Panel B: Fraction of Variation							
Firm Fixed Effect: Permanent differences across firms within sector	54.27%	64.04%	57.88%	24.64%	25.80%	27.99%	14.63%
Residual: Variation over time in the identity of firms within industries most affected by climate exposure	45.73%	35.96%	42.12%	75.36%	74.20%	72.01%	85.37%
Sum	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Table IA6: **Robustness Check: Transition Measure and Realized Disasters**

This table presents the validation test results using different types of natural disasters. The dependent variable is the sum of the five transition-related climate exposure measures. *Realized disaster* equals one if the firm's headquarter is located in a county hit by natural disasters in the past year. The dummy variable *Hurricane* (*Flood*/*Drought*/*Heat*) is assigned a value of one if the firm experienced a hurricane (flood/drought/heat) in the past year. Textual measures are normalized to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-quarter fixed effects. Standard errors are clustered by industry. t-statistics are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
	Transition exposure				
Realized disaster	-0.03				
	(-1.00)				
Hurricanes		-0.02			
		(-0.50)			
Floods			-0.02		
			(-0.76)		
Droughts				-0.05	
				(-1.05)	
Heats					-0.07**
					(-2.59)
Observations	122,230	122,230	122,230	122,230	122,230
R-squared	0.31	0.31	0.31	0.31	0.31
Control	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes	Yes

Table IA7: **Robustness Check: Physical Measures and Green Patents**

This table shows the validation test results using proxies for transition measures. The dependent variables are the number of various green patents in columns (1) to (3), and the number of climate incidents in column (4). We average the quarterly textual measures during the year and normalize them to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-fixed effects. Standard errors are clustered by industry, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) Carbon abatement	(2) Renewable energy	(3) Energy efficiency	(4) Incident numbers
Weather	0.00	-0.01**	-0.00	-0.01
	(0.24)	(-1.99)	(-0.04)	(-0.87)
Disaster	0.01	0.01	0.02**	0.01
	(1.16)	(1.04)	(2.17)	(0.48)
Observations	15,809	15,809	15,809	1,997
R-squared	0.36	0.21	0.46	0.43
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Table IA8: **Robustness Check: Transition Measures and Carbon Emissions**

This table shows the validation test results using scope 2 and 3 carbon emission data from Trucost. We average the quarterly textual measures during the year and normalize them to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year fixed effects. Standard errors are clustered by industry, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)
Panel A: Scope 2			
	Carbon intensity	Δ Carbon intensity	Δ Carbon emissions
Carbon removal	0.07***	-0.22*	-0.21
	(5.50)	(-1.68)	(-0.55)
Renewable energy	-0.01	-0.44**	-0.59
	(-0.70)	(-2.55)	(-0.66)
Enabling technology	0.00	0.14	0.85
	(0.12)	(1.32)	(0.80)
Technology challenges	-0.02***	-0.03	-0.35
	(-4.72)	(-0.30)	(-0.49)
Regulation concerns	0.02	0.40**	0.07
	(1.24)	(2.20)	(0.19)
Observations	13,861	11,380	11,377
R-squared	0.39	0.03	0.40
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Panel B: Scope 3			
	Carbon intensity	Δ Carbon intensity	Δ Carbon emissions
Carbon removal	0.06**	-0.36	-0.07
	(2.15)	(-1.13)	(-0.44)
Renewable energy	0.05	-1.82***	-0.37***
	(1.65)	(-4.20)	(-2.86)
Enabling technology	-0.05	-0.19	-0.03
	(-1.54)	(-1.24)	(-0.21)
Technology challenges	0.04*	0.99***	0.38*
	(1.78)	(3.76)	(1.96)
Regulation concerns	0.05	0.23	0.07
	(1.62)	(0.85)	(0.61)
Observations	13,861	11,380	11,380
R-squared	0.66	0.21	0.88
Industry FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes

Table IA9: **Comparison with Existing Textual Measures**

This table compares the three papers that develop firm-level climate exposure measures, focusing on word collection, category identification, and empirical findings.

	Sautner et al. (2023)	**Li et al. (2023)**	**This paper**
Word collection	• Manually select the initial seed words. • Expand the word list by machine learning techniques without considering context.	• Manually construct a word list by reading external sources like Wikipedia.	• Automatically extract seed words from IPCC reports through LDA. • Expand the word list using the context-driven word2vec model.
Category identification	• Calculate the proportion of bigrams for each preset topic. – *General*. – *Opportunity*, *Regulation*, and *Physical*.	• Count the physical keywords that appear near risk synonyms. – *Acute* and *Chronic*. • Count transition words without the necessity of being adjacent to a risk synonym. • Use a list of proactive verbs to decompose *transition risk* measures into *proactive* and *non-proactive* components.	• Classify climate-related sentences based on underlying topics and sentiments identified by LDA and FinBERT. • Count the proportion of climate sentences for each category: – Physical: *Disaster*(acute), *Weather*(chronic). – Transition: *Mitigation efforts* (*Carbon removal*, *Renewable energy*, *Enabling technology*), *Technology challenges*, *Regulation concerns*.
Emission reduction	Not detectable.	Not detectable.	*Renewable energy* (*Enabling technology*) helps firms enhance their own (customers') reduction in carbon emission levels.
Value implications	Sautner et al. (2022): • No realized risk premiums for climate exposures. • Option-based expected return proxies show a positive relation with climate exposure from 2012 to 2014, attributing to the *Opportunity* topic.	• No value examination for physical risks. • Insignificant negative association between transition risks and firm values over the subsequent three years when compared within firms.	• Firms discussing climate efforts more frequently enjoy a lower discount rate, increasing firm values. • Firms with larger physical exposure have higher expected returns, lowering their present value.

Table IA10: **Comparison to Sautner et al. (2023): Climate-related Bigrams**

This table contrasts the top-100 bigrams in Sautner et al. (2023) with our climate word pool. Panel A presents the words from our word pool with a higher frequency than their 100^{th} bigram but not found in Sautner et al. (2023). Panel B displays the words from Sautner et al. (2023) not included in our word pool. Column N indicates the bigrams' frequency in the corresponding corpus. Appendix IA.2 provides the complete list of our climate word pool, including unigrams and trigrams.

Bigrams	N	Bigrams	N	Bigrams	N
Panel A: Words from our word pool that are absent in Sautner et al. (2023)					
energy efficiency	10173	water quality	1129	storm related	934
weather related	7698	electricity gas	1120	ice storm	929
weather condition	7125	green energy	1102	storm activity	899
wind farm	6862	carbon credit	1089	renewable diesel	895
cold weather	5408	renewable fuel	1083	environmental compliance	892
bad weather	3508	drought condition	1056	weather disruption	559
fuel efficiency	3300	adverse weather	1055	solar plant	558
warm weather	3217	heavy rain	1036	environmental social	551
wind turbine	3164	renewable portfolio	1028	wind tower	551
natural disaster	3084	solar wind	1019	mother nature	548
alternative energy	3019	emission control	1008	storm hit	521
environmental impact	2778	wind generation	995		
impact hurricane	2685	renewable generation	984		
weather pattern	2646	nuclear fuel	958		
wind project	2637	hot summer	937	Total: 90,740	
Panel B: Words from Sautner et al. (2023) that are absent in our word pool					
new energy	4544	vehicle manufacturer	740	driver assistance	545
government india	1201	future energy	737	promote use	536
battery electric	1121	motor control	726	farm project	531
integrate resource	1052	electric bus	709	laser diode	528
world population	977	fast charge	675	deliver energy	526
obama administration	957	cell power	657	manage energy	522
unite nation	925	energy team	650	invest energy	521
provide energy	851	energy application	615	capacity energy	512
efficient solution	839	help state	604		
need clean	821	power agreement	586		
energy star	793	supply energy	585		
design use	777	source power	575		
area energy	777	compare conventional	560		
charge station	762	effort energy	560		
major design	747	pass house	559	Total: 30,903	

Table IA11: **Comparison with Existing Measures: Carbon Emission Reductions**

This table shows the validation test results using carbon emission data from Trucost, controlling for textual measures constructed by Sautner et al. (2023). We average the quarterly textual measures during the year and normalize them to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-fixed effects. Standard errors are clustered by industry, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) Carbon intensity_1	(2) Δ Carbon intensity_1	(3) Δ Carbon emissions_1	(4) Δ Carbon $\text{emissions}_1^{customers}$
Carbon removal	0.38	-3.79**	0.72	0.39
	(1.59)	(-2.61)	(0.91)	(0.40)
Renewable energy	2.53***	-17.74***	-1.85**	0.28
	(4.31)	(-3.76)	(-2.67)	(0.59)
Enabling technology	-0.50**	0.23	-0.54	-1.10**
	(-2.20)	(0.21)	(-1.03)	(-2.12)
Technology challenges	0.59	3.92***	0.23	-0.20
	(1.29)	(4.28)	(0.39)	(-0.43)
Regulation concerns	-0.01	4.65**	1.52	-0.66
	(-0.05)	(2.33)	(1.67)	(-0.84)
CCExposure^{Opp}	-0.93*	6.85	0.50	-0.36
	(-1.73)	(1.30)	(0.70)	(-0.69)
CCExposure^{Reg}	0.74**	-4.47*	1.13	0.38
	(2.33)	(-1.84)	(1.07)	(0.81)
Observations	13,777	11,323	11,314	4,569
R-squared	0.55	0.06	0.50	0.04
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes

Table IA12: **Comparing with Existing Measures: Realized Disasters**

This table presents the validation test results using different types of natural disasters. The dependent variable is the physical climate exposure measure from Sautner et al. (2023). *Realized disaster* equals one if the firm's headquarter is located in a county hit by natural disasters in the past year. The dummy variable *Hurricane* (*Flood*/*Drought*/*Heat*) is assigned a value of one if the firm experienced a hurricane (flood/drought/heat) in the past year. Textual measures are normalized to a zero mean and a one standard deviation. Control variables include size, leverage, ROA, investment, tangibility, and dividend, as well as industry and year-quarter fixed effects. Standard errors are clustered by industry. t-statistics are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
			$CCExposure^{ph}$		
Realized disaster	0.00				
	(0.01)				
Hurricanes		-0.01			
		(-0.45)			
Floods			-0.03**		
			(-2.08)		
Droughts				-0.01	
				(-0.17)	
Heats					0.02
					(0.49)
Observations	117,724	117,724	117,724	117,724	117,724
R-squared	0.01	0.01	0.01	0.01	0.01
Control	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
YearQtr FE	Yes	Yes	Yes	Yes	Yes

Table IA13: **Fama-French Five Factor Tests**

This table presents Fama-French five factor tests for portfolios sorted on each climate exposure measure relative to their industry peers. Standard errors are estimated by using the Newey-West correction of 12 lags, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	L	2	3	4	H	HML	L	2	3	4	H	HML
	Panel A: Disaster						Panel B: Weather					
α	-0.25**	-0.42**	-0.12	0.06	-0.00	0.25	-0.20	-0.23**	-0.14	0.17	0.31**	0.51***
	(-2.10)	(-2.28)	(-0.90)	(0.34)	(-0.02)	(1.44)	(-1.43)	(-2.01)	(-0.79)	(0.76)	(2.08)	(3.25)
MKT	1.01***	0.99***	0.98***	0.88***	0.97***	-0.04	0.97***	0.98***	1.02***	0.83***	0.90***	-0.07
	(27.97)	(15.22)	(35.46)	(24.43)	(8.98)	(-0.41)	(21.23)	(15.73)	(23.14)	(18.75)	(11.43)	(-0.87)
SMB	0.14**	0.15	0.16**	0.04	0.17*	0.02	0.11**	0.22**	0.22**	0.13	0.26***	0.15
	(2.07)	(1.33)	(2.34)	(0.69)	(1.82)	(0.25)	(2.14)	(2.53)	(2.42)	(1.32)	(3.36)	(1.55)
HML	-0.01	-0.15***	0.08	0.05	-0.05	-0.04	-0.08	-0.09	-0.15**	-0.21***	0.01	0.09
	(-0.07)	(-2.65)	(0.86)	(0.30)	(-0.41)	(-0.38)	(-0.98)	(-1.59)	(-2.44)	(-3.33)	(0.12)	(0.78)
RMW	0.08	0.34**	0.13	-0.02	0.22	0.14	0.29**	0.15*	0.35***	0.21**	0.14	-0.15
	(0.73)	(2.35)	(1.16)	(-0.16)	(1.66)	(0.73)	(2.57)	(1.73)	(3.01)	(2.23)	(1.07)	(-0.75)
CMA	0.37*	0.38*	0.40***	0.74***	0.20	-0.18	0.27**	0.16	0.62***	0.57***	0.30**	0.02
	(1.92)	(1.91)	(3.21)	(4.10)	(0.91)	(-0.76)	(2.22)	(1.31)	(5.31)	(6.15)	(2.26)	(0.14)
	Panel C: Carbon removal						Panel D: Renewable energy					
α	-0.39**	-0.40**	-0.05	-0.73***	-0.78***	-0.39**	0.46	-0.27	-0.07	-0.14	0.32*	-0.14
	(-1.99)	(-2.09)	(-0.21)	(-3.39)	(-3.29)	(-1.98)	(1.54)	(-1.15)	(-0.30)	(-0.80)	(1.70)	(-0.43)
MKT	1.10***	1.02***	0.95***	1.03***	1.06***	-0.03	0.95***	0.90***	0.83***	0.66***	0.60***	-0.36***
	(16.71)	(16.21)	(17.99)	(12.86)	(10.54)	(-0.42)	(16.48)	(9.47)	(7.39)	(11.22)	(8.10)	(-5.16)
SMB	0.26***	0.09	0.12	0.01	0.05	-0.20	0.05	-0.03	-0.00	0.08	-0.05	-0.10
	(3.33)	(0.91)	(1.19)	(0.07)	(0.37)	(-1.53)	(0.33)	(-0.15)	(-0.01)	(0.65)	(-0.22)	(-0.44)
HML	-0.13	-0.17	-0.20	-0.10	-0.22	-0.09	0.08	-0.27**	-0.10	-0.13	-0.27*	-0.35***
	(-1.06)	(-0.90)	(-1.02)	(-0.54)	(-1.31)	(-1.07)	(0.83)	(-2.08)	(-0.65)	(-1.25)	(-1.89)	(-2.91)
RMW	0.32*	0.05	0.32*	0.26	0.06	-0.26	0.11	0.22**	-0.00	0.25**	0.10	-0.02
	(1.96)	(0.27)	(1.93)	(1.29)	(0.42)	(-1.64)	(0.72)	(2.25)	(-0.01)	(2.23)	(0.69)	(-0.15)
CMA	0.26	0.57***	0.74***	0.32	0.53***	0.27**	-0.28	0.52**	0.80***	0.60**	0.69**	0.97***
	(1.59)	(2.78)	(3.11)	(1.58)	(3.28)	(2.51)	(-0.93)	(2.02)	(2.89)	(2.58)	(2.24)	(4.82)
	Panel E: Enabling technology						Panel F: Technology challenges					
α	0.12	0.15	-0.20	0.17	-0.23*	-0.36**	-0.11	0.21	0.34	0.29*	0.15	0.26
	(0.96)	(1.05)	(-1.39)	(1.10)	(-1.76)	(-2.06)	(-0.49)	(1.22)	(1.42)	(1.85)	(0.82)	(1.31)
MKT	1.09***	0.97***	1.06***	0.97***	0.99***	-0.10*	1.02***	0.87***	0.64***	0.63***	0.64***	-0.38***
	(24.52)	(20.67)	(21.65)	(37.82)	(23.03)	(-1.91)	(9.73)	(10.20)	(7.26)	(9.68)	(8.99)	(-5.56)
SMB	-0.05	0.13	0.17***	0.09*	0.15	0.20	0.15	0.13	0.17	-0.05	0.04	-0.11
	(-0.61)	(1.62)	(2.99)	(1.69)	(1.25)	(1.44)	(1.34)	(1.43)	(1.49)	(-0.29)	(0.22)	(-0.97)
HML	-0.11**	-0.15***	0.03	-0.06	-0.04	0.08	-0.18	-0.02	-0.15	-0.28***	-0.22*	-0.03
	(-2.52)	(-2.63)	(0.46)	(-1.06)	(-0.41)	(0.80)	(-1.39)	(-0.17)	(-1.34)	(-2.71)	(-1.90)	(-0.27)
RMW	0.19**	-0.01	0.23***	0.22***	0.02	-0.18	0.06	0.28*	0.25	0.26**	0.26**	0.20
	(2.03)	(-0.14)	(3.61)	(2.69)	(0.12)	(-1.28)	(0.47)	(1.84)	(1.32)	(2.28)	(2.40)	(1.49)
CMA	-0.09	0.30***	0.09	0.33***	0.01	0.10	0.44	0.47**	0.58***	0.44*	0.60**	0.16
	(-0.75)	(3.03)	(0.90)	(3.16)	(0.08)	(0.59)	(1.58)	(2.42)	(2.95)	(1.67)	(2.27)	(0.77)
	Panel G: Regulation concerns											
α	-0.86***	-0.33	0.16	-0.48*	-0.07	0.79***						
	(-3.17)	(-1.01)	(0.71)	(-1.72)	(-0.26)	(3.41)						
MKT	1.09***	0.99***	0.91***	0.91***	0.91***	-0.17						
	(9.35)	(7.53)	(13.79)	(9.39)	(19.90)	(-1.61)						
SMB	0.13	0.05	0.24**	0.09	0.31***	0.19						
	(0.77)	(0.56)	(2.20)	(0.61)	(3.34)	(1.40)						
HML	0.10	0.34	-0.17	0.37**	0.26	0.16						
	(0.74)	(1.64)	(-1.35)	(2.49)	(0.98)	(0.94)						
RMW	0.09	0.08	0.26	0.20	-0.21	-0.30						
	(0.66)	(0.39)	(1.52)	(1.10)	(-0.72)	(-0.96)						
CMA	0.16	-0.05	0.89***	0.04	0.43	0.27						
	(0.68)	(-0.16)	(3.89)	(0.16)	(1.22)	(0.80)						

Table IA14: **Fama-MacBeth Regressions for 2020**

This table presents Fama-MacBeth regressions of individual stock returns on climate exposure measures after 2019. We match monthly returns from July of year t to June of year $t+1$ on the textual measures in year $t-1$. Control variables include the natural logarithm of market capitalization, the natural logarithm of book-to-market ratio, ROA, and investment rate. All independent variables are normalized to a zero mean and one standard deviation. t-statistics based on standard errors using the Newey-West correction for 12 lags are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
	Excess returns				
Carbon removal	0.09	0.06			0.07
	(1.15)	(0.87)			(0.92)
Renewable energy	-0.01	0.01			0.01
	(-0.21)	(0.16)			(0.09)
Enabling technology	0.04	0.02			0.02
	(0.22)	(0.09)			(0.10)
Technology challenges	-0.05	-0.08*			-0.08*
	(-1.01)	(-1.89)			(-1.98)
Regulation concerns	0.07	0.08			0.07
	(0.65)	(0.71)			(0.64)
Weather			-0.01	-0.02	-0.02
			(-0.06)	(-0.20)	(-0.31)
Disaster			0.03	0.04	0.05
			(0.42)	(0.59)	(0.75)
logME		-0.31		-0.32	-0.31
		(-0.64)		(-0.66)	(-0.64)
logB/M		-0.13		-0.13	-0.13
		(-0.70)		(-0.69)	(-0.70)
ROA		0.11		0.11	0.10
		(0.25)		(0.25)	(0.25)
Investment		0.10		0.11	0.10
		(0.74)		(0.76)	(0.75)
Observations	60,230	54,739	60,230	54,739	54,739
R-squared	0.11	0.15	0.15	0.10	0.15
Number of groups	24	24	24	24	24
Industry FE	Yes	Yes	Yes	Yes	Yes
YearMonth FE	Yes	Yes	Yes	Yes	Yes

Table IA15: **Robustness Check: Portfolio Sorting**

This table reports the equal-weighted average returns for five portfolios sorted on the five transition climate exposure measures in Panel A and the two physical measures in Panel B, relative to their Fame-French 49 industry peers. Portfolios are rebalanced at the end of each June, and monthly returns from July of year t to June of year $t+1$ are matched to the yearly textual measures in year $t-1$. The sample spans from July 2009 to December 2019. t-statistics based on standard errors using the Newey-West correction for 12 lags are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1) Low	(2) 2	(3) 3	(4) 4	(5) High	(6) HML
Panel A: Transition exposure						
Carbon removal	0.90*	0.54	0.93**	0.56	0.33	-0.57***
	(1.83)	(0.90)	(2.06)	(1.07)	(0.76)	(-3.48)
Renewable energy	0.98**	1.02***	0.87**	1.01***	0.27	-0.70***
	(2.55)	(3.56)	(2.36)	(2.90)	(0.69)	(-3.32)
Enabling technology	1.22***	1.23***	1.16***	0.95**	0.80*	-0.42**
	(3.65)	(2.97)	(2.97)	(2.16)	(1.96)	(-2.58)
Technology challenges	1.09***	1.29***	1.05**	1.14***	0.67*	-0.43
	(2.91)	(3.35)	(2.55)	(2.89)	(1.83)	(-1.65)
Regulation concerns	0.30	0.36	0.36	0.36	0.42	0.12
	(0.44)	(0.63)	(0.50)	(0.66)	(0.82)	(0.31)
Panel B: Physical exposure						
Weather	0.99***	1.22***	0.99***	1.05***	1.36***	0.37*
	(2.83)	(3.15)	(2.98)	(2.97)	(3.82)	(1.81)
Disaster	0.65	0.66	0.73	0.72	0.91	0.26*
	(1.28)	(1.34)	(1.50)	(1.42)	(1.64)	(1.85)

Table IA16: **Robustness Check: Asset Pricing Factor Tests**

This table presents asset pricing factor tests for the High-minus-Low portfolios sorted on the seven climate exposure measures relative to their Fama-French 49 industry peers. We perform time-series regressions of equal-weighted portfolio returns on the market factor in column (1), on the Fama-French three factors plus the momentum factor in column (2), on the Fama-French five factors in column (3), and on the Hou-Xue-Zhang q factors in column (4). t-statistics based on standard errors using the Newey-West correction for 12 lags are shown in parentheses. *** p<0.01, ** p<0.05, and * p<0.1.

	(1) CAPM	(2) FFM	(3) FF5	(4) HXZ
Panel A: Transition exposure				
Carbon removal	-0.52***	-0.45**	-0.39**	-0.50***
	(-3.12)	(-2.50)	(-2.16)	(-2.86)
Renewable energy	-0.64***	-0.61***	-0.58**	-0.62***
	(-2.87)	(-2.69)	(-2.39)	(-2.63)
Enabling technology	-0.38**	-0.33*	-0.32*	-0.32*
	(-2.19)	(-1.95)	(-1.72)	(-1.74)
Technology challenges	-0.30	-0.22	-0.20	-0.21
	(-1.00)	(-0.83)	(-0.65)	(-0.77)
Regulation concerns	0.14	0.18	0.33	0.14
	(0.36)	(0.46)	(0.97)	(0.37)
Panel B: Physical exposure				
Weather	0.47**	0.52**	0.53**	0.49**
	(2.25)	(2.53)	(2.38)	(2.27)
Disaster	0.28**	0.33**	0.38**	0.34**
	(2.02)	(2.13)	(2.49)	(2.35)

Table IA17: **Robustness Check: Firm Value and Climate Exposures**

This table reports regressions of contemporaneous excess returns on quarterly climate exposure measures. Columns (1)-(2) employ measures derived through the implementation of ClimateBERT as a substitution for FinBERT within our NLP procedure for sentiment classification. Columns (3)-(6) add the overall sentiment score and risk score in the earning call as control variables. Textual measures are normalized to a zero mean and one standard deviation. Other control variables include the natural logarithm of market capitalization, the natural logarithm of book-to-market ratio, ROA, investment, leverage, tangibility, WW index, sales growth, as well as firm and time-fixed effects. Standard errors are clustered by industry and year, and t-statistics are shown in parentheses.*** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

	(1)	(2)	(3)	(4)	(5)	(6)
	FirmValue					
	ClimateBERT		Sentiment		Risk	
Carbon removal	0.05*		0.07*		0.08**	
	(1.89)		(1.89)		(2.27)	
Renewable energy	0.06**		0.01		0.05*	
	(2.48)		(0.47)		(1.96)	
Enabling technology	0.09***		0.08**		0.14***	
	(2.76)		(2.10)		(3.72)	
Technology challenges	-0.06**		-0.04*		-0.07***	
	(-2.46)		(-1.73)		(-2.93)	
Regulation concerns	-0.03		-0.03		-0.06**	
	(-1.09)		(-1.15)		(-2.00)	
Weather		-0.13***		-0.07***		-0.13***
		(-4.44)		(-3.32)		(-5.67)
Disaster		-0.11***		-0.04**		-0.11***
		(-4.76)		(-2.26)		(-5.45)
Sentiment			1.13***	1.12***		
			(34.21)	(36.81)		
Risk					-0.16***	-0.13***
					(-5.13)	(-4.51)
Observations	359,083	359,083	359,083	359,083	359,083	359,083
R-squared	0.24	0.24	0.24	0.24	0.24	0.24
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
YM FE	Yes	Yes	Yes	Yes	Yes	Yes

Table IA18: **Comparison with Existing Measures: Firm values**

This table reports regressions of contemporaneous excess returns on quarterly climate exposure measures. Columns (1)-(4) execute regressions utilizing measures that are constructed by using top keywords from Sautner et al. (2023) and Li et al. (2023) to identify climate-focused sentences. Column (5) gathers sentences highlighted by our top climate terms and consolidates them into an overall climate exposure index without granular categorization. Textual measures are normalized to a zero mean and one standard deviation. Control variables include the natural logarithm of market capitalization, the natural logarithm of book-to-market ratio, ROA, investment, leverage, tangibility, WW index, sales growth, as well as firm and time-fixed effects. Standard errors are clustered by industry and year, and t-statistics are shown in parentheses.*** p<0.01, ** p<0.05, and * p<0.1.

	(1)	(2)	(3)	(4)	(5)
	FirmValue				
	Sautner et al. (2023)		Li et al. (2023)		Ours
Carbon removal	0.03		0.04*		
	(0.96)		(1.77)		
Renewable energy	0.02		0.01		
	(0.62)		(0.36)		
Enabling technology	0.02		0.03		
	(0.75)		(1.44)		
Technology challenges	-0.01		-0.03		
	(-0.47)		(-1.66)		
Regulation concerns	0.02		-0.01		
	(1.10)		(-0.61)		
Weather		-0.01		-0.03*	
		(-1.00)		(-1.86)	
Disaster		-0.01*		-0.03	
		(-1.87)		(-1.48)	
Climate overall					-0.01
					(-0.30)
Observations	353,589	353,589	353,589	353,589	353,589
R-squared	0.23	0.23	0.23	0.23	0.23
Firm FE	Yes	Yes	Yes	Yes	Yes
YM FE	Yes	Yes	Yes	Yes	Yes

IA References

Blei, D. M. (2012). Probabilistic topic models. *Commun. ACM*, 55(4):77–84.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, 3(null):993–1022.

Goldberg, Y. and Levy, O. (2014). word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method.

Hoffman, M., Bach, F., and Blei, D. (2010). Online learning for latent dirichlet allocation. In Lafferty, J., Williams, C., Shawe-Taylor, J., Zemel, R., and Culotta, A., editors, *Advances in Neural Information Processing Systems*, volume 23. Curran Associates, Inc.

Li, K., Mai, F., Shen, R., and Yan, X. (2020). Measuring Corporate Culture Using Machine Learning. *The Review of Financial Studies*, 34(7):3265–3315.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. In Bengio, Y. and LeCun, Y., editors, *1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., and Dean, J. (2013b). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'13, page 3111–3119, Red Hook, NY, USA. Curran Associates Inc.

Rong, X. (2016). word2vec parameter learning explained.

Steyvers, M. and Griffiths, T. (2007). Probabilistic topic models. *Handbook of latent semantic analysis. Lawrence Erlbaum Associates Publishers.*, page 427–448.