Firm-level Climate Change Exposure

Finance Working Paper N° 686/2020
March 2023

Zacharias Sautner
Frankfurt School of Finance and Management and ECGI

Laurence van Lent
Frankfurt School of Finance and Management

Grigory Vilkov
Frankfurt School of Finance and Management

Ruishen Zhang
Shanghai University of Finance and Economics

ECGI Working Paper Series in Finance

Firm-level Climate Change Exposure

Working Paper N° 686/2020
March 2023

Zacharias Sautner
Laurence van Lent
Grigory Vilkov
Ruishen Zhang

We thank Stefan Nagel, an anonymous Associate Editor, two referees, Artur Hugon, Marcin Kacperczyk, Bryan Kelly, Kelvin Law, Christian Leuz, Tim Loughran, Quentin Moreau, and Ane Tamayo for helpful comments. We are grateful to Aakash Kalyani for preparing the green jobs data set and to Markus Schwedeler for sharing his code. Participants at AFA 2022 Meetings, NBER, NYU Stern (PhD Classes in Empirical Household Finance), University of Zurich, University of Dusseldorf, Stockholm Business School, SUFE, Bocconi University, Columbia University, CEIBS, LSE, Xiamen University, Duke Kunshan University, University of Miami, Tingshua University, Dongbei University of Finance and Economics, Chinese University of Hong Kong, and University of Hong Kong provided helpful feedback. Funding for this project was provided by DFG Project ID 403041268 - TRR 266 (Van Lent and Zhang), the Institute for New Economic Thinking (Van Lent), Shanghai Pujiang Program (Zhang), the 111 Project (B18033)(Zhang), and National Natural Science Foundation of China (Project ID 72202128, 72241405, Zhang). Parts of this project were conducted when Sautner was visiting NOVA School of Business and Economics and UCSD Rady School of Management (funded in part by the Norwegian Finance Initiative).

Abstract

We develop a method that identifies the attention paid by earnings call participants to firms' climate change exposures. The method adapts a machine learning keyword discovery algorithm and captures exposures related to opportunity, physical, and regulatory shocks associated with climate change. The measures are available for more than 10,000 firms from 34 countries between 2002 and 2020. We show that the measures are useful in predicting important real outcomes related to the net-zero transition, in particular, job creation in disruptive green technologies and green patenting, and that they contain information that is priced in options and equity markets.

Keywords: Climate change, climate risk, conference calls, institutional investors

JEL Classifications: G18, G32, G38, Q54, Q55

Zacharias Sautner*
Professor of Finance
Frankfurt School of Finance and Management
Adickesallee 32-34
60322 Frankfurt am Main, Germany
phone: +49 69 154008 755
e-mail: z.sautner@fs.de

Laurence van Lent
Professor of Accounting and Economics
Frankfurt School of Finance and Management
Adicksallee 32-34
Frankfurt am Main 60322, Germany
phone: +49 69 154008531
e-mail: L.vanLent@fs.de

Grigory Vilkov
Professor of Finance
Frankfurt School of Finance and Management
Adicksallee 32-34
Frankfurt am Main 60322, Germany
phone: +49 (069) 154008-842
e-mail: g.vilkov@fs.de

Ruishen Zhang
Assistant Professor of Accounting
Shanghai University of Finance and Economics, Institute of Accounting and Finance
Guoding Road 777, Yangpu District
Shanghai 200433, China
phone: +86 18637904884
e-mail: zhangruishen@sufe.edu.cn.

*Corresponding Author

Firm-Level Climate Change Exposure

ZACHARIAS SAUTNER, LAURENCE VAN LENT, GRIGORY VILKOV,

and RUISHEN ZHANG*

ABSTRACT

We develop a method that identifies the attention paid by earnings call participants to firms' climate change exposures. The method adapts a machine learning keyword discovery algorithm and captures exposures related to opportunity, physical, and regulatory shocks associated with climate change. The measures are available for more than 10,000 firms from 34 countries between 2002 and 2020. We show that the measures are useful in predicting important real outcomes related to the net-zero transition, in particular, job creation in disruptive green technologies and green patenting, and that they contain information that is priced in options and equity markets.

*Zacharias Sautner, Laurence van Lent, and Grigory Vilkov are at Frankfurt School of Finance & Management. Ruishen Zhang is at Shanghai University of Finance and Economics. The climate change exposure data are available at https://doi.org/10.17605/OSF.IO/FD6JQ. We thank Stefan Nagel, an anonymous Associate Editor, two referees, Artur Hugon, Marcin Kacperczyk, Bryan Kelly, Kelvin Law, Christian Leuz, Tim Loughran, Quentin Moreau, and Ane Tamayo for helpful comments. We are grateful to Aakash Kalyani for preparing the green jobs data set and to Markus Schwedeler for sharing his code. Participants at AFA 2022 Meetings, NBER, NYU Stern (PhD Classes in Empirical Household Finance), University of Zurich, University of Dusseldorf, Stockholm Business School, SUFE, Bocconi University, Columbia University, CEIBS, LSE, Xiamen University, Duke Kunshan University, University of Miami, Tingshua University, Dongbei University of Finance and Economics, Chinese University of Hong Kong, and University of Hong Kong provided helpful feedback. Funding for this project was provided by DFG Project ID 403041268 - TRR 266 (Van Lent and Zhang), the Institute for New Economic Thinking (Van Lent), Shanghai Pujiang Program (Zhang), the 111 Project (B18033)(Zhang), and National Natural Science Foundation of China (Project ID 72202128, 72241405, Zhang). Parts of this project were conducted when Sautner was visiting NOVA School of Business and Economics and UCSD Rady School of Management (funded in part by the Norwegian Finance Initiative). We have read *The Journal of Finance* disclosure policy. Sautner is a Regular Research Visitor at the ECB and a member of the Sustainability Council of Lampe Asset Management (since November 2022). He has no other disclosure to make. Van Lent, Vilkov, and Zhang have no conflicts of interest to disclose.

Correspondence: Zacharias Sautner, Frankfurt School of Finance & Management, Adickesallee 32-34, 60322 Frankfurt am Main, Germany; e-mail: z.sautner@fs.de.

Climate change will profoundly affect the way business is conducted. Scientists have developed complex models that estimate the effect of greenhouse gas emissions on the global climate. At the same time, however, little evidence exists on the degree to which climate change impacts jobs, innovation, and risk-sharing in capital markets. One key challenge in estimating these impacts is that it is difficult to measure how individual firms are affected by climate change (Giglio, Kelly, and Stroebel (2021)), as the effects are multifaceted, originating from multiple sources. For instance, while physical climate changes and regulations implemented to combat global warming can impose costs on some firms, climate change can provide opportunities for other firms, such as those operating in renewable energy, electric cars, or energy storage. It is therefore important to develop disaggregated measures that capture this variation across firms. The measures should also reflect market participants' assessments about how climate change affects individual firms. Such information is important to consider in a finance context given the critical role that market participants play in the resource allocation and price discovery process.

In this paper we make progress on this front by using transcripts of earnings conference calls to construct time-varying measures of how call participants across the globe view firms' exposures to different facets of climate change. Earnings calls are key corporate events in which financial analysts listen to management and ask questions about current and future developments material to the firm (Hollander, Pronk, and Roelofsen (2010)). We interpret these measures as capturing the attention financial analysts and management devote to climate change topics at a given point in time. A benefit of these measures is that they reflect "soft" information originating from information exchanges between managers and analysts.[1]

[1]This feature allows us to provide economic insights beyond those derived from existing firm-level exposure measures based on "hard" information (e.g., carbon emissions, extreme local weather events). Note that the exchanges are not limited to soft information but might also discuss specific quantitative data or restate "hard" information in conversational terms. Prior literature provides important insights

To construct the climate change exposure measures, we build on recent work using quarterly earnings calls as a source for identifying firms' various risks and opportunities (Hassan et al. (2019, 2021, 2023a), Jamilov, Rey, and Tahoun (2021), Hassan et al. (2023b)). These studies use the proportion of the conversation during an earnings call that relates to a particular topic to capture the firm's exposure to that topic. We follow these papers in defining "exposure" to an issue as the share of the conversation in a transcript devoted to that topic.[2] We depart from these papers, however, along two dimensions. Firstly, our measures capture the market's perception of a firm's exposure to various upside or downside factors related to climate change, namely physical threats, regulatory interventions, and technological opportunities. Second, to mitigate the challenges of identifying "niche languages" that use specific wordings, particularly in the context of climate change, where language use varies among policymakers, journalists, and financial market participants (Webersinke et al. (2021)), we develop a new method that adapts the keyword discovery algorithm proposed in King, Lam, and Roberts (2017) to construct four related sets of climate change bigrams in earnings calls. The first captures broadly defined aspects of climate change. The remaining three measures cover specific climate change "topics:" opportunities, physical shocks (e.g., sea level rise), and regulatory shocks (e.g., carbon taxes, cap and trade markets). We then use these four sets of bigrams to construct firm-level measures reflecting call participants' topical attention. In particular, the measures count the frequency of specific climate change bigrams in a transcript, scaled

into the relations between "hard" information and real and financial outcomes at the firm level (e.g., Bolton and Kacperczyk (2021, 2022, 2023), Ilhan, Sautner, and Vilkov (2021) or De Haas and Popov (2022) for carbon emissions, and Kruttli, Roth Tran, and Watugala (2021), Hong, Li, and Xu (2019), Addoum, Ng, and Ortiz-Bobea (2020), or Pankratz and Schiller (2021) for weather events).

[2]This definition of "exposure" is different from how risk exposure is defined in the asset pricing literature. Our measure is not intended to capture the covariance with aggregate fluctuations. Hassan et al. (2019) discuss the relationship between these two areas of literature.

by the number of bigrams.[3] The algorithm only requires human input to specify a short list of initial keywords associated with climate change. Our sample covers data from over 10,000 firms in 34 countries between 2002 and 2020.

We conduct several validation exercises to verify our methodology. First, we consider the face validity of the climate change bigrams. Second, we follow Baker, Bloom, and Davis (2016) and perform a structured human audit in which 18 graduate students independently coded over 2,000 transcript text fragments. Both of these exercises suggest that our algorithm reliably captures bigrams identifying climate change discussions. Third, our exposure measures are robust to excluding one keyword at a time from the initial keywords list. Fourth, our keyword search-based measures substantially improve the identification of climate change discussions relative to an alternative approach using the initial keywords only. And fifth, we find plausible industry patterns in the exposure measures. When we aggregate exposure to the industry level, the sector with the highest overall exposure is Electric, Gas, & Sanitary Services (utilities), followed by Construction (top-ranked firms build power generation systems or solar projects) and Transportation Equipment (top-ranked firms build fuel-cell or zero-emission vehicles). Utilities top the exposure ranking for opportunity and regulatory shocks, which indicates that this sector faces both opportunities (e.g., renewable energy) and regulatory risks (e.g., carbon taxes).[4]

[3]We also construct "sentiment" measures, which count the relative frequency of climate change bigrams that occur in the vicinity of positive and negative tone words (Loughran and McDonald (2011)), and "risk" measures, which count the relative frequency of climate change bigrams mentioned in the same sentence as the words "risk," "uncertainty," or their synonyms.

[4]That firms with heightened regulatory risks also exhibit climate-related opportunities is consistent with Cohen, Gurun, and Nguyen (2021), who document that several major electricity, oil, and gas firms are not only large CO_2 emitters, but also innovators in green technologies. This finding is consistent with how analysts view sectors with high regulatory risks (e.g., "Morgan Stanley: 'Second wave of renewables' to drive 70 GW of coal retirements," *S&P Global Market Intelligence*, December 20, 2019).

Our results reveal sizeable within-industry variation for all measures, which indicates that firms benefit or suffer from climate change to various degrees. A case in point is the comparison between TotalEnergies and ExxonMobil. While TotalEnergies and ExxonMobil have similar regulatory exposures, TotalEnergies scores more than seven times higher in terms of measured opportunities. This divergence in perceived prospects is consistent with differences in the perceived extent to which these firms embrace renewable energy and the net-zero transition into their business models (Pickl (2019)).

In a final validity check, we find that climate exposure positively correlates with carbon emissions and Engle et al.'s (2020, EGKLS) index of public climate change attention. The association with emissions stems from regulatory and opportunity exposure (since physical exposure is unrelated to emissions).[5] The effect of public attention also arises from positive associations between EGKLS's index and the opportunity and regulatory exposure measures.

We apply our measures to shed light on the nature of climate change exposure among our sample firms. Perhaps surprisingly, as climate change is often seen as an aggregate risk factor associated with global changes in the physical climate, its within-sector impact is far from uniform. A variance analysis that separates the relative contributions of aggregate, sectoral, and firm-level exposure by including the corresponding sets of fixed effects shows that between 70% and 96% of the variation in the exposure measures plays out at the firm level. Only half of this firm-level variation is persistent, suggesting that firms within an industry are exposed to climate change to varying degrees over time. Thus, the effects of climate change are heterogeneous across firms even within an industry. This result is consistent with the idea that many factors that affect a firm's ability to adapt to

[5] This result may also reflect the fact that some firms' emissions provides opportunities by supporting the transition to a greener economy (e.g., producers of building materials that make houses more energy-efficient). Such "enabling activities" are also explicitly included in the EU Taxonomy, which identifies activities that help reach the EU's climate targets.

a greener economy exhibit large firm-level components (e.g., managerial skill, financing constraints).

We interpret the large share of firm-level variance as capturing economically meaningful heterogeneity and argue that a firm's idiosyncratic climate change exposure is the key driver of this heterogeneity. That being said, a plausible alternative is that part of the variation reflects idiosyncratic measurement error. Several tests dispel this alternative for several reasons. First, as discussed below, we report robust associations between our measures and green job creation, green innovation, and risk-related outcomes. Second, following Hassan et al. (2019), we directly quantify the amount of measurement error contained in the firm-level variation. Approximately 5% to 10% of the variation in measured exposure can be attributed to measurement error. The implied measurement error at the firm level is about two percentage points higher than that for the overall variation. Although we interpret these results with due caution, they suggest that measurement error in the firm-level dimension is higher than that in the overall panel, but only modestly so.

Having bolstered confidence that the firm-level variation in measured climate change exposure is meaningful, we apply it to four real and financial market outcomes. In the first two applications, we demonstrate that climate change exposure predicts green-tech hiring and green patents, two key drivers of the low-carbon transition. Using data compiled from Burning Glass by Bloom et al. (2021), we establish that firms with higher measured climate exposures create more jobs in disruptive green technologies over the subsequent year:[6] a one-standard-deviation increase in climate change exposure is associated with a 109% increase in green jobs in the following year. This overall effect originates from more job creation at firms exhibiting higher measured opportunities and regulatory exposures.

[6] Our data do not cover all jobs potentially related to climate change, but they do identify job postings with potential to have a lasting and meaningful real impact, as Bloom et al. (2021) only consider job creation in "disruptive" technologies (e.g., solar or battery technology).

The results for green-tech job creation extend to green patenting. A one-standard-deviation increase in climate change exposure is associated with a 72% increase in the number of green patents in the following year. Once more, this finding stems from firms with higher opportunities and regulatory exposures. High-exposure firms are not simply recruiting more across fields. They are also not more innovative, in general. In fact, firms with higher exposure hire less in nongreen-tech areas and generate fewer nongreen patents.

The remaining two applications relate climate change exposure to financial market outcomes. We first show that measured exposure is related to risks and risk premiums in the options market. Such relationships are plausible, as policy uncertainty surrounding regulation, including climate policy uncertainty, is priced in options (Kelly, Pastor, and Veronesi (2016), Ilhan, Sautner, and Vilkov (2021)). Likewise, there is plenty of uncertainty surrounding green technology or renewable energy investment. Realizing these opportunities leads to significant gains if successful or large losses if unsuccessful. It is therefore plausible that measured exposure relates to investors' propensity to hedge extreme climate risks and/or gamble on climate outcomes. Indeed, for options written on stocks with high overall exposure, the tail regions are relatively more expensive. Effects are similar at firms with high opportunity exposure, for which investors are willing to pay a (variance risk) premium. In comparison, effects are smaller but still statistically significant for firms with high regulatory exposure. This finding corroborates the view that some firms with high regulatory exposure face downside risks and upside potential (due to their innovation activity).

We also document the conditional pricing of a factor that reflects innovations to the aggregate level of climate change exposure. Firms with higher betas to this factor face higher uncertainty related to future developments in climate-related areas and, as a result, earn higher returns.[7] Our estimation applies the approach of Gagliardini, Ossola, and Scaillet (2016), which performs well when—as in our case—the cross-section is large relative to the time series. We obtain a positive average conditional risk premium on the factor, and, more importantly, find large time-series variability in the risk premium.[8]

Our keyword discovery approach of extracting climate-related information from text offers an alternative approach to contemporaneous papers that try to accomplish the same task by relying on other advances in natural language processing (NLP). All of this work, including ours, is based on the understanding that standard NLP methods are not well suited for "niche languages," that is, specialized, highly technical vocabulary that varies substantially across textual sources (Webersinke et al. (2021),Varini et al. (2020)). These frictions are exacerbated when the wordings associated with a topic are complex, ambiguous, and fast moving. A promising approach among these alternatives is to use pre-trained language models to learn word patterns in the language. When implementing this pre-training approach in a specific domain of interest (e.g., climate change), rather than using large generic corpora, researcher have found some promising results (Kölbel et al. (2022), Bingler, Kraus, and Leippold (2022)). Work is ongoing on these problems. Which approach works best in the context of climate finance is ultimately an empirical matter.

A valid question is whether our approach delivers meaningful gains above and beyond an alternative, off-the-shelves approach. Our main argument is that keyword discovery

[7]Our primary objective is to show that climate attention in earnings calls is linked to systematic risk, with shocks to such attention being priced in the cross-section. We do not want to propose a new factor to be added to the factor zoo, and we do not try to use a conditional model framework to explain asset pricing anomalies (Lewellen and Nagel (2006)).

[8]A caveat of all four applications is that any evidence of our measures' ability to predict real and financial outcomes is a success only if the true relationship exists in the data. We therefore face the usual joint-hypothesis problem between the quality of our measures and the true economic model generating the data.

is useful when the language of interest is not common. We illustrate this claim by constructing, for comparison purposes, alternative exposure measures using a list of pre-specified keywords from EGKLS. These keywords appear more frequently in earnings calls than the bigrams we identify, probably because EGKLS's set also contains unigrams and more general terms. However, several of EGKLS's unigrams are part of our top-100 list of bigrams, and exposure measures based on the pre-specified keywords correlate positively with our measures. Beyond these correlations, a question is why the approaches differ. As mentioned above, our measures have the benefit of capturing context-specific jargon used in specialized economic environments (earnings calls), while an approach using pre-specified keywords better captures broader discussions (e.g., in news media in the case of EGKLS's keywords). In addition, our approach adjusts the vocabulary over time, while using pre-specified keywords fixes this vocabulary ex ante.[9] Finally, especially for the topic-based measures, it is easier to identify initial seed bigrams than to develop keyword lists from authoritative texts.

Most closely related to our paper is the contemporaneous work by Li et al. (2021, LSTY), who also use earnings calls to identify climate risks. We diverge from their work in terms of our method, focus, and sample. More specifically, LSTY use a pre-specified training library to identify climate risk words, which, we argue, is unlikely to uncover the exact language used in earnings calls to discuss climate change (see also Varini et al. (2020)). In addition, while LSTY focus on physical and regulatory risks among U.S. firms, we provide a more comprehensive analysis based on a global sample and include upside opportunity effects of climate change. Based on a textual analysis of 10K reports, Baz et al. (2022) document that firms with more regulatory climate change exposure experience positive stock return effects after the 2016 Trump election.

[9]Time-series variation in true (unobservable) climate change exposure, especially over long horizons, is more likely to be picked up by such an "evolutionary" approach. Indeed, the selection of pre-specified keywords may become obsolete over time with changing technologies or climate change concerns.

Since making our data available, our measures have been related to a series of real and financial outcomes. This "out-of-sample" evidence is reassuring, as it indicates that the measures capture meaningful variation across firms and do not reflect mostly noise. On the real side, as in our paper, von Schickfus (2021) illustrates more green patenting when the overall measure and the opportunity measure are higher, and Li, Lin, and Lin (2022) show that the overall measure predicts depressed overall innovation. Furthermore, our overall measure positively relates to cash holdings (Heo (2021)) and explains how strongly U.S. firms' emissions declined in response to the EPA's 2010 Greenhouse Gas Reporting Program (Tomar (2021)). Our physical measure is related to physical risk disclosure in 8K filings (Gostlow (2021)), and the opportunity measure relates to firms' carbon risk management (Duong et al. (2021)). On the financial side, our physical measure is associated with lower leverage after the Paris Agreement (Ginglinger and Moreau (2022)). Mueller and Sfrappini (2022) show that after regulatory climate risks become salient, bank lending is skewed towards firms with high regulatory exposure in the U.S., but away from such firms in the EU. We provide additional evidence in Sautner et al. (2022) that our measures are priced in equity markets, and Kölbel et al. (2022) show that the overall measure is negatively associated with CDS spreads after the Paris Agreement. Di Giuli et al. (2022) find that investors' propensity to vote for climate proposals after experiencing hot temperatures is higher at firms with more overall climate change exposure. Heath et al. (2022) find that SRI funds invest less in firms with higher overall climate change exposure. Our keyword dictionary is used by Hail, Kim, and Zhang (2021).

The rest of the paper proceeds as follows. Section I describes the data. Section II presents our method to quantify firm-level climate change exposure. Section III validates the exposure measures. Section IV presents a variance decomposition of the exposure measures and addresses measurement error. Section V presents four applications of the exposure measures. Section VI concludes.

I. Data

A. Data on Earnings Conference Calls

We use transcripts of quarterly earnings calls held by publicly listed firms to construct time-varying measures of the attention paid by call participants to firm-level climate change exposure. The measures are constructed using the entire earnings call, including both the management presentation and the Q&A session with analysts.[10] The transcripts are collected from the Refinitiv Eikon database. We use the complete set of English-language transcripts from 2002 to 2020. Unless indicated otherwise, as most of our other data vary at the year level, we average quarterly transcript-based measures for each firm. We exclude countries with 150 or fewer firm-year observations and drop SIC codes 9900-9999 ("Non-classifiable"). Our final sample includes 86,152 firm-year observations from 10,673 firms headquartered in 34 countries. Variable definitions are presented in the Appendix.[11]

B. Data on Carbon Emissions

Some tests use data on carbon emissions (*Total Emissions*), calculated as the sum of Scope 1 and Scope 2 emissions, from S&P Global Trucost. These data include emissions reported by firms and emissions estimated by Trucost. We use emission levels, rather than intensities, as emission levels are associated with a risk premium (Bolton and Kacperczyk

[10]We also provide tests based on the measured exposure constructed from the Q&A session only. The Q&A part is less scripted and may be less subject to strategic disclosure incentives than the presentation part. In some calls, analysts ask no questions (we would calculate a climate change exposure of zero in these cases). However, zero-question calls are a nonrandom event, and treating these calls as if the firm is unexposed to climate change likely introduces bias (Chen, Hollander, and Law (2014).)

[11]Table IA.I in the Internet Appendix provides the distribution of firm-years across countries. The Internet Appendix is available in the online version of the article on the *Journal of Finance* website.

(2021, 2023)), are the prime target of policy and investor initiatives aiming to achieve net-zero emissions, and are directly linked to carbon budgets (Bolton, Kacperczyk, and Samama (2021)). Further, many firms have witnessed strong investor opposition on reporting emission intensities. To link the emissions data with our sample firms, we apply a series of matching variables based on the following order: (i) GVKEYs, (ii) ISINs, (iii) exact names, (iv) fuzzy names, and (v) tickers plus the first two ISIN digits. We can match 33,789 firm-years with the emissions data (4,999 unique firms from 34 countries between 2004 and 2020).[12]

C. Data on Public Attention to Climate Change

We borrow an index developed by EGKLS to capture how public climate change attention varies in the time series. The *WSJ CC News Index* is constructed by measuring news about climate change in the *Wall Street Journal* (WSJ). To quantify the intensity of climate news coverage, EGKLS compare the WSJ's news content to a corpus of authoritative texts on climate change. The resulting measure reflects the fraction of the WSJ dedicated to the topic of climate change each day (we use average annual values). For our sample, *WSJ CC News Index* is available from 2002 to 2017.

D. Data on Green-Tech Jobs

Job data related to important green technologies come from Bloom et al. (2021). These authors use textual analysis to identify 29 disruptive technologies over the past decades, of which four are broadly related to climate change ("hybrid vehicle electric car," "lithium battery," "solar power," and "fracking"). Our data from Bloom et al. (2021)

[12] Table IA.II illustrates that Trucost data coverage is higher for firm-years with higher climate change exposure, larger, more profitable, and less-R&D-intense firms, and non-U.S. firms. The higher climate change exposure scores are expected given that Trucost caters to clients in need of climate risk data (especially risks related to emissions).

contain online job postings by firms related to these four technologies. We refer to the jobs related to these technologies as green-tech jobs.[13] The data do not cover all jobs potentially related to climate change, but do identify those green jobs that, by Bloom et al.'s (2021) construction, have a lasting and meaningful ("disruptive") real impact. Bloom et al. (2021) obtain these data from Burning Glass (BG), which aggregates online job postings using "spider bots" from job boards or employer websites.[14] We match these data by GVKEY and year. Jobs data are available for U.S. firms for 2007 and 2010 to 2020.

The measure *#Green-Tech Jobs* is the number of postings for disruptive green-tech jobs in a firm-year. We assume that no green-tech job was posted if a firm-year does not indicate disruptive green-tech job creation in the BG database. (The results are robust to only considering firm-years within the BG database; many firm-years in BG also show zero green-tech postings). Some tests use *#Nongreen-Tech Jobs*, the number of job postings related to nongreen disruptive technologies in a firm-year. We observe disruptive green job postings in 5.4% of firm-years, and conditional on *#Green-Tech Jobs* being nonzero, the average (median) number of green-tech jobs is 38 (3). The top-5 firms in the cumulative count of new green-tech jobs include Tesla, Sunrun, First Solar, Sunpower Corp, and Viviant Solar.

[13] It is unclear ex ante whether fracking has positive or adverse environmental effects. More specifically, Acemoglu et al. (2019) argue that shale gas has the short-term benefit of lower emissions, when compared to conventional fossil fuels. However, the shale gas boom may lead to less innovation in other emission-reducing technologies in the long run. Furthermore, fracking has negative climate effects due to emission leakage. Our results are robust to excluding fracking jobs.

[14] BG data are also used by Darendeli, Law, and Shen (2022) to measure green hiring. Campello, Gao, and Xu (2021) additionally use BG data, though not in a climate context.

E. Data on Green Patents

To identify green patents, we collect patent data from the Google Patents (GP) database. This database is also used by Kogan et al. (2017) and Kelly et al. (2021). To identify "green" patents, we follow Cohen, Gurun, and Nguyen (2021) and apply an OECD classification that identifies patents with the potential to address environmental problems. A description of how the OECD classifies patents into technology groups is provided by Haščič and Migotto (2015). Green patents include patents on emission abatement technologies, renewable energy, and energy storage. As in Kogan et al. (2017), we use name matching to match patent assignee names to sample firms.[15] Patent data are available for U.S. firms from 2002 to 2019 (GP coverage for 2020 was still limited at the time of writing).

The measure *#Green Patents* is the number of green patents filed in a firm-year. We assume that no green patenting occurred if we are unable to identify a green patent in GP for a firm-year (results are robust to relaxing this assumption). Consistent with Acemoglu et al. (2019), new green patents are relatively rare—we observe green patenting in only 1.4% of firm-years. However, the distribution is highly skewed. If we consider observations within GP, then green patenting is observed in 15% of firm-years. Conditional on green patenting being nonzero, the average (median) number of green patents equals 8.5 (2). The top green patent producer is Caterpillar, with 1,364 green patents over the sample period.[16] We also use the total number of nongreen patents filed (*#Nongreen Patents*).

[15]We track the timing of an invention by matching patents using the priority year, that is, the effective date of a patent filing (De Haas and Popov (2022)). While the "filing date" corresponds to when a patent application is filed at the patent office, the "priority date" is when the novelty of an invention is established.

[16]Caterpillar traditionally manufactured diesel engines and mining equipment, but moved into selling photovoltaic or energy storage technology. The firm also ranks in the top-10 in Cohen, Gurun, and Nguyen's (2021) sample; the slight ranking divergence is due to different sample periods.

F. Data on Risks and Risk Premiums in the Options Market

Data on option-implied variables are from the Volatility Surface File of Ivy DB OptionMetrics. In these tests we focus on S&P500 firms, for which data on liquid options are available. We match options data through the historical CUSIP link of OptionMetrics. We construct six measures: implied variance ($IVar$), implied skewness ($ISkew$), implied kurtosis ($IKurt$), implied volatility slopes (*SlopeD* and *SlopeU*), and variance risk premium (VRP). The variable construction process is detailed in Section II of the Internet Appendix. The high frequency of the option-implied measures allows us to use quarterly values of $CCExposure$.[17]

G. Data on Risk Premiums in the Equity Market

Our tests examining the climate change exposure factor use monthly data on the standard factors from Ken French's data library. Term and default spread data are from the St. Louis Fed's FRED library. The term spread is the difference between the 10-year and three-month Treasury constant maturity data series (variable *T10Y3MM*). The default spread is the difference between the Baa and Aaa corporate bond yield (*BAA10YM* and *AAA10YM*). Book-to-market ratio data (defined in log terms as in Fama and French (2008)) come from Compustat North America. Term and default spreads and the book-to-market ratio for each firm are centered and standardized in the time series, and then used as instruments for conditional risk premium estimation. We restrict the risk premium tests to S&P500 firms with more than 28 monthly returns (out of 228) during our sample period.

[17] To avoid look-ahead bias, we match quarterly exposure values covering earnings calls in quarter t (typically discusses events of quarter $t-1$) with option-implied measures from the last day of quarter t.

H. Financial Statement Data

Data on firm financial variables (e.g., total assets, debt, CAPEX, R&D, or cash holdings) are from Compustat North America and Compustat Global.

II. Quantifying Firm-level Exposure to Climate Change

A. Discovery of Climate Change Bigrams

To quantify exposure to climate change, we build on Hassan et al. (2019, 2021, 2023a). Extracting climate-related information from text sources is challenging (Webersinke et al. (2021)). Methods using training libraries or pre-specified word lists do not cope well with the niche language used to describe climate change.[18] In addition, discussion in earnings calls considers climate change together with topics such as regulation, tax credits, technological breakthroughs, and performance. This results in substantial ambiguity about when the discussion is genuinely about climate change. Finally, vocabulary used to discuss climate change is fast moving, changing to reflect shifting opinions, regulations, and innovations related to climate change.

To address these challenges, we adapt the keyword discovery algorithm proposed in King, Lam, and Roberts (2017).[19] This algorithm does not require a comprehensive "climate change" training library, but rather only a small set of "initial" bigrams (see Table IA.III). These initial bigrams are chosen because they relate unambiguously to climate change. The algorithm then uses these initial bigrams to search for new bigrams that also likely indicate climate change conversation and searches directly in the transcripts.

[18] That said, researchers have used the SEC Climate Disclosure Search tool, which looks for pre-specified keywords in SEC filings, to develop a measure of climate risk (Berkman, Jona, and Soderstrom (2019)).

[19] Details, including how we define the set of initial bigrams, are presented in Section I of the Internet Appendix.

Because each initial bigram is connected to a specific group of new bigrams discovered through the search algorithm, one can easily decompose the measure of climate change exposure into its constituent parts based on the presence of these bigrams. The initial bigrams allow the algorithm to identify sentences that focus unambiguously on climate change. The algorithm then extracts "features" by relying on supervised learning methods. Features are bigrams beyond the initial set predicting climate change from the identified sentences. Finally, the algorithm constructs a model predicting whether a sentence is related to climate change. We apply this prediction model to sentences that do not include any initial bigrams and then assess whether the predicted sentences are related to climate change. To discover new climate change bigrams, we reverse-engineer the machine-learning (ML) process and trace back the bigrams that best discriminate climate-change-related sentences from other sentences. The resulting set of climate change bigrams $\mathbb{C}$ includes the initial bigrams and the newly identified bigrams.

That our approach generates meaningful climate change bigrams based on the initial bigrams is helpful for several reasons. First, it extends the rather broadly specified initial bigrams into more specialized word combinations.[20] Second, $\mathbb{C}$ includes the names of several power stations and wind farms (e.g., "kibby wind" or "coughlin power"), which are of interest to call participants that discuss the climate change exposure of these facilities' operators. These bigrams illustrate the challenge of using training libraries or pre-specified word lists to identify climate change talk—few researchers have the detailed field knowledge to recognize the relationship between these words and climate change.

Our approach allows us to adapt the bigram-search algorithm to discover three unique sets of bigrams from $\mathbb{C}$ that capture opportunities as well as regulatory and physical climate shocks. Toward this end, we feed a set of initial bigrams reflecting these three

[20]For example, "rooftop solar" and "photovoltaic panel" come from the initial bigram "solar energy," while "nuclear power" and "event fukushima" come from "renewable energy," and "tesla battery" and "hybrid plug" come from "electric vehicle."

topics to the search algorithm. We then allow the algorithm to discover bigrams related to the topic of interest. Table IA.IV lists the initial bigrams used for the topic search. We construct new initial bigrams for these topics by hand-picking appropriate bigrams from the top-500 bigrams discovered after the first generic, nontopic-specific bigram search. We then re-run the search algorithm to find a broader set of bigrams for each topic. As the topic-based algorithm yields some general climate change bigrams, we drop bigrams appearing in more than one topic to guarantee that we do not have overlapping topic measures. In the final stage, we take the intersection between $\mathbb{C}$ and each set of topic bigrams to obtain the sets of opportunity, regulatory, and physical climate change bigrams (i.e., $\mathbb{C}^{Opp}$, $\mathbb{C}^{Reg}$, and $\mathbb{C}^{Phy}$), respectively.

B. Construction of Climate Change Exposure Measures

Using the bigram sets, we construct measures of climate change exposure for each transcript. We interpret these measures as capturing the attention devoted to climate change topics by call participants at a point in time, rather than as measures of fundamental exposure. We use the broad set of climate change bigrams $\mathbb{C}$ to illustrate how we construct these measures. The topic measures are constructed analogously; we simply replace $\mathbb{C}$ with the bigrams that relate to the corresponding topic.

We construct an overall exposure measure, $CCExposure$, based on how frequently the specified bigrams appear in a transcript. This involves taking the set of climate bigrams $\mathbb{C}$ to the transcript of firm i in quarter t and counting the frequency of these bigrams. To account for the call length, we scale the count by the number of bigrams in the transcript,

$$CCExposure_{i,t} = \frac{1}{B_{i,t}} \sum_{b}^{B_{i,t}} (1[b \in \mathbb{C}]), \tag{1}$$

where $b = 0, 1, ..., B_{i,t}$ are the bigrams in the earnings call transcripts of firm i in quarter t

and $1[\cdot]$ is the indicator function. We create an annual measure for each firm by averaging the quarterly measures. We produce exposure measures from $\mathbb{C}^{Opp}$, $\mathbb{C}^{Reg}$, and $\mathbb{C}^{Phy}$, respectively, by scoring each transcript using the same method. We label the topic-based measures as $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$.

Some of our tests employ two refinements. In the first refinement, we create two sentiment measures by counting the number of climate change bigrams after conditioning on the presence of the positive or negative tone words in Loughran and McDonald (2011),

$$CCSentiment_{i,t}^{Pos/Neg} = \frac{1}{B_{i,t}} \sum_{b}^{B_{i,t}} \{(1[b \in \mathbb{C}]) \times \sum_{b}^{b \in S} \mathcal{T}^{Pos/Neg}(b)\}, \tag{2}$$

where S represents the sentence containing bigrams $b = 0, 1, ..., B_{i,t}$ and $\mathcal{T}^{Pos/Neg}(b)$ assigns sentiment to each bigram b:[21]

$$\mathcal{T}^{Pos}(b) = \begin{cases} 1 & \text{if } b \text{ has a positive tone} \\ 0 & \text{if otherwise} \end{cases}$$

$$\mathcal{T}^{Neg}(b) = \begin{cases} 1 & \text{if } b \text{ has a negative tone} \\ 0 & \text{if otherwise.} \end{cases}$$

In the second refinement, we construct a measure of risk by counting the relative frequency of the climate change bigrams mentioned in the same sentence with the words

[21] Though not used in this paper, we also combine both sentiment measures into an overall measure by counting the climate change bigrams after conditioning on the presence of positive and negative tone words,

$$CCSentiment_{i,t} = \frac{1}{B_{i,t}} \sum_{b}^{B_{i,t}} \{(1[b \in \mathbb{C}]) \times \sum_{b}^{b \in S} \mathcal{T}(b)\},$$

where $\mathcal{T}(b) = 1$ (-1) if b has positive (negative) tone, and zero otherwise.

"risk," "uncertainty," or their synonyms,

$$CCRisk_{i,t} = \frac{1}{B_{i,t}} \sum_{b}^{B_{i,t}} (1[b \in \mathbb{C}] \times 1[b, r \in S]), \tag{3}$$

where r contains the words "risk," "uncertainty," or a synonym.

The exposure measures do not adjust for the differences in the importance or typical frequencies of individual bigrams. For robustness, we account for such differences by constructing measures that weigh each bigram with a score reflecting the bigram's representativeness for climate discussions. We do this so that common terms that appear in most transcripts receive low scores, as these terms are less informative about a call's content, as do rare terms in a given transcript, as these terms have low text frequency. This approach follows Hassan et al. (2019), Gentzkow, Kelly, and Taddy (2019), and EGKLS and is commonly referred as "term frequency–inverse document frequency" (TFIDF). Formally,

$$CCExposure_{i,t}^{TFIDF} = \frac{1}{B_{i,t}} \sum_{b}^{B_{i,t}} \left(1[b \in \mathbb{C}] \times log(\frac{N_{\mathbb{T}}}{f_{b,\mathbb{T}}}) \right), \tag{4}$$

where $N_{\mathbb{T}}$ refers to the number of transcripts and $f_{b,\mathbb{T}}$ to the number of transcripts in which bigram b appears. A bigram appearing in many transcripts therefore has low weight when calculating the TFIDF score, and—in the extreme case—if a given bigram appears in every transcript, it receives zero weight ($log(\frac{N_{\mathbb{T}}}{f_{b,\mathbb{T}}}) = 0$).

Table I reports summary statistics for the exposure measures (for purposes of exposition, the measures are multiplied by 10^3).[22] Table IA.V reports the correlations across the exposure measures. A few correlations deserve further comment. The correlation

[22]The magnitudes of $CCExposure^{TFIDF}$ are larger than those of $CCExposure$ as the inverse document frequency of the climate change bigrams can be much larger than one (the document frequencies of the climate change bigrams are much smaller than the total number of transcripts).

between $CCExposure^{Reg}$ and $CCExposure^{Opp}$ is positive at 33%, and $CCExposure^{Phy}$ is largely unrelated to $CCExposure^{Reg}$ and $CCExposure^{Opp}$. In addition, the correlation between $CCExposure$ and $CCExposure^{TFIDF}$ is 99.7%.

Insert Table I about here.

Tables IA.VI to IA.VIII report the sample distribution at the earnings-call (transcript) level across countries, years, and industries. We report the distributions for all sampled earnings calls and for those calls with nonzero climate change exposure. The tables show meaningful proportions of calls with nonzero climate change exposure across all three sample cuts; transcripts with $CCExposure > 0$ are not concentrated in certain countries, years, or industries. Our analysis does not make use of a binary indicator for whether $CCExposure$ is nonzero, but instead uses a continuous measure.

III. Validation

A. *Validation at the Bigram Level*

A.1. *Face Validity of Climate Change Bigrams*

We validate our exposure measures using a multi-pronged approach. First, we consider the bigrams' face validity. Table II lists the 100 highest-frequency bigrams in $\mathbb{C}$. The top bigrams associated with $CCExposure$ capture aspects of the opportunities and risks associated with climate change. The top bigrams include both opportunity-related word pairs (e.g., "battery power," "new energy") and risk-related terms (e.g., "environmental concern," "extreme weather").

Insert Table II about here.

Table IA.IX considers the three topic-based measures. When we construct $CCExposure^{Opp}$ using initial bigrams such as "wind power" or "solar energy," we find several new bigrams that refer to new (green) technologies (e.g., “solar farm,” “carbon free”) (Panel A). Several word combinations are linked to developments in “electric vehicles,” including “charge infrastructure” and “battery electric.” With respect to $CCExposure^{Reg}$ (Panel B), when we use initial bigrams “carbon tax,” “air pollution,” or “air quality”, that is, terms related to regulatory interventions, we discover bigrams that explicitly include the word “regulation” or its synonyms (e.g., “control regulation,” “environmental legislation”). Turning to the top bigrams for $CCExposure^{Phy}$ (Panel C), we use initial bigrams such as “natural hazard” or “sea level” to identify word pairs intuitively linked to physical climate change (e.g., “area florida,” “ice control,” “wind speed”).

For the 10 highest-scoring firms on $CCExposure$, Table IA.X provides “snippets.” These snippets are text fragments taken from the point in the transcript that the algorithm identifies as the moment when the participants discuss climate issues. Consider Ocean Power Technologies, a U.S. firm that turns ocean wave power into electricity for offshore applications. In its 2008Q4 call, bigrams such as “energy requirement,” “powerbuoy wave,” “wave condition,” and “wave power” were heavily featured. In the top snippet, participants discuss the increased demand for the firm’s trademark technology (the PowerBuoy®) due to heightened attention to renewable energy. Not surprisingly, high-scoring firms are involved in energy production or the broader energy infrastructure. Indeed, when ECOtality call participants use climate change bigrams, they discuss how charging infrastructures are central to advancing zero-emissions transportation.

A.2. Audit Study Based on Human Reading

We developed a two-stage snippet-based audit to evaluate the scoring of our algorithm (Baker, Bloom, and Davis (2016), Hassan et al. (2019)). While our algorithm should be judged in the context of the entire transcript, a snippet-based audit improves our ability to sample across a large number of transcripts. In the first stage, we define a snippet as the 10 sentences around the climate change bigram with the highest text frequency in a transcript. For transcripts with *CCExposure*=0, we randomly choose a snippet of 10 consecutive sentences for the audit. In our pilot study, each of the authors independently coded 250 identical and randomly selected snippets using a binary coding scheme. The coding used the variable *CCAudit*, which equals one if the rater classifies the text as providing evidence of climate change exposure, and zero otherwise. In addition, for each snippet we record *Coding Confidence*, which ranges from three (the rater is highly confident that their coding is correct) to one ("hard calls"). We identified some slight coding differences between the authors and resolved discrepancies. Based on this iterative procedure, we developed a detailed guide with definitions of what text should be coded as climate change exposure and which snippets should not qualify as such. The audit guide describes examples of snippets and offers interpretations and suggested coding to help the raters solve complex cases in the audit process. We then instructed two graduate students based on the audit guide and asked them to audit the same 250 snippets that the author team coded to assess any remaining inconsistencies.

In the second stage, we recruited 19 graduate students to each independently code 250 new snippets from the audit universe. Together they assessed 2,090 unique snippets.[23] Auditors received training based on the audit guide. The snippets were partially overlapping to allow us to conduct some inter-rater correspondence tests. Our goal is to verify the information content of *CCExposure* at various points of its distribution. Following Hassan et al. (2019), we create portfolios with the same number of transcripts based on their percentile of the *CCExposure* distribution. We then count the number of transcripts at that percentile that the auditors rated as *CCAudit*=1 (i.e., the snippet is classified as containing a clear discussion of a firm's climate change exposure). We count

[23]We first sorted all transcripts with nonzero *CCExposure* into deciles. We then randomly selected 10 snippets from each decile and another 10 from *CCExposure*=0 transcripts for each sample year.

310 true positives out of 339 snippets (91% correct positives) in the top-decile portfolio (transcripts with the highest value of *CCExposure*). The rate of correct positives declines almost linearly as we move to the median and bottom portfolios. This is displayed in Figure 1, which plots the relationship between (the predicted probability of) true positives (as judged by the human reading) at each decile and the median percentile score of *CCExposure* at that percentile. The association is positive and nearly linear, as would be expected if our algorithm reliably identifies climate change discussions.[24]

Insert Figure 1 about here.

A.3. Comparison with Approach Using Pre-Specified Keywords

We construct alternative exposure measures from a list of pre-specified climate change keywords to compare these measures with those produced by our algorithm. To obtain such a list, we use the set of unique stemmed unigrams and bigrams $\mathbb{C}^{EGKLS}$ used by EGKLS to build their time-varying, news-based index of climate change attention. These keywords originate from 74 authoritative texts . To create $CCExposure^{EGKLS}$, we replace $\mathbb{C}$ with $\mathbb{C}^{EGKLS}$ and recompute the relative frequency with which the alternative terms appear in the transcripts. We construct a frequency-unweighted and a TFIDF version, denoted by $CCExposure^{EGKLS-EW}$ or $CCExposure^{EGKLS-TFIDF}$, respectively.

Table IA.XI illustrates that the unigrams and bigrams in $\mathbb{C}^{EGKLS}$ appear more frequently in earnings calls than the bigrams in $\mathbb{C}$. This finding is unsurprising as $\mathbb{C}^{EGKLS}$ includes more unigrams and more general terms (the top-3 bigrams are "market," "increase," and "time"). Using unigrams rather than bigrams trades off the higher likelihood

[24]These findings suggest that our algorithm correctly identifies climate change text, even at relatively low *CCExposure* scores. A benchmark is provided in Hassan et al. (2019), where the number of correct positives reduces to below five out of 20 at the 90th percentile of their text-based political risk score. Weighting observations by *Coding Confidence* does not materially change our findings.

of a given term occurring in a text against the higher probability of a false positive, that is, wrongly classifying a fragment as climate change text (van Zaanen and Kanters (2010)). Several of the unigrams in $\mathbb{C}^{EGKLS}$ are part of the top-100 bigrams in $\mathbb{C}$ (e.g., "carbon," "energy," or "water"). As would be expected from Table IA.XI, the mean values of both alternative exposure measures (Table IA.XII, Panel A) are larger than those of $CCExposure$. Thus, larger parts of the earnings calls are classified as discussing climate topics if we use $\mathbb{C}^{EGKLS}$ instead of $\mathbb{C}$. At the same time, Table IA XII, Panel B, indicates that the measures correlate positively with $CCExposure$. The correlation table illustrates that our main measure and the alternative measures yield more similar assessments when the public pays close attention to climate change (*WSJ CC News Index* is in the top quartile). One possible explanation is that at times when the WSJ devotes a lot of space to climate topics, terms from a more general climate library (on which the index and pre-specified keyword measures build) become more commonly used in earnings calls. Intuitively, media attention might homogenize the language used to talk about climate change. When the media pivots to other events, the vocabulary likely used to discuss climate change in earnings calls becomes more idiosyncratic again. Such instances are plausibly better reflected in our keyword-search-based approach.

A question that remains is how our measure and a measure using pre-defined keywords differ economically. Our measure is well suited to capture context-specific jargon used in specialized environments with experts and allows us to construct topic-based measures. The pre-specified keyword approach better captures broader discussions by the public, as reflected in articles published in the WSJ, while identifying specific or emerging topics with a pre-specified keyword approach is more challenging. A further difference is that our approach is "evolutionary," that is, it will reflect changes in the vocabulary used in transcripts over time, while an approach using pre-specified keywords fixes this vocabulary ex ante. Time-series variation in true (unobservable) climate change exposure, especially over long horizons, is more likely to be picked up by such an "evolutionary" approach. Any selection of pre-specified keywords is due to become obsolete the further out one moves in time.

A.4. Perturbation Tests for Individual Initial Bigrams

We evaluate the extent to which our overall exposure measure depends on *individual* bigrams in the initial bigram list (Table IA.III) by performing a perturbation test. We successively exclude one initial bigram at a time, recomputing the modified set of bigrams $\mathbb{C}^{Pert}$ as well as the modified $CCExposure^{Pert}$. Given that our initial short list contains 50 bigrams, we construct 50 new versions of $CCExposure^{Pert}$. After aggregating the measure to the firm-year level, we calculate the correlation of each of these exposure measures with $CCExposure$. These correlations are above 85%, which means that $CCExposure$ does not depend much on specific initial seed bigrams.

A.5. Comparison with Approach Using Initial Bigrams Only

Table II shows that the initial keywords dominate the top-100 bigrams used in the construction of $CCExposure$. This raises the question of how big the performance gain of the keyword discovery approach is relative to the alternative that only uses the initial seed bigrams. To address this question, we construct the new exposure measure $CCExposure^{Initial}$ from the initial bigrams only. Figure 2, Panel A, shows how frequently the new measure signals zero exposure, while $CCExposure$ instead reveals that climate topics are discussed. Results are reported by $CCExposure$ decile. In the top decile, $CCExposure^{Initial}$ indicates no exposure in 27% of transcripts. Hence, even among the most exposed firms, there is a performance gain when applying our approach. This gain increases when we consider other deciles—already in the second decile, $CCExposure^{Initial}$ deviates from $CCExposure$, indicating the absence of exposure in more than 62% of transcripts. The effects increase monotonically as we move to lower exposure deciles.

Panel B reports results of the topic-based exposure measures, with the alternative measures using only the topic-based initial bigrams (Table IA.IV). For all three measures and deciles, significant fractions of the transcripts are incorrectly classified as having zero exposure. Even in the three respective top deciles, the alternative approach misses positive exposure in 10% to 40% of the transcripts. Across all deciles, the gain from the keyword discovery approach is largest for $CCExposure^{Opp}$ (especially in the lower deciles).

Beyond these statistics, identifying exposure using bigrams beyond the initial seed words is economically important. Below we show that, among the set of firms for which $CCExposure^{Initial}$=0, our exposure measures keep predicting green outcomes. These effects are identified purely from the bigrams obtained through the keyword search algorithm.

B. Validation at the Climate Change Exposure Level

B.1. Climate Change Exposure: Industry Variation

We now move away from the bigram level to examine the properties of the exposure measures. This involves several steps. In the first step, we compute averages by industry sector (two-digit SIC code level) and present a ranking of these means in Table III. In Panel A, using $CCExposure$, the sectors with the highest overall exposure include Electric, Gas, & Sanitary (SIC49). Top-ranked firms within this sector include China Longyuan Power Group, China's largest producer of wind power, and the U.S. utility Allete. This sector is followed by Heavy Construction (SIC16) and Construction (SIC17). High-ranking firms in these sectors include A-Power Energy Generation Systems, a Chinese firm providing on-site power generation systems, ReneSola, a U.S. firm developing and operating solar projects, and Quanta Services, a U.S. infrastructure solutions provider for firms in the energy and pipeline business. Top-ranked firms in the Transportation Equipment sector (SIC37), ranked next, include alternative fuel and zero-emission vehicle firms.

Insert Table III about here.

A few sectors are worth commenting on in Panels B to D, which report the topic-based measures. Utilities top the list for $CCExposure^{Opp}$ (Panel B) and $CCExposure^{Reg}$ (Panel C). While the latter ranking position is expected, given the sector's exposure to carbon taxes or related regulations, the earlier position is more surprising. Notwithstanding, it is consistent with Cohen, Gurun, and Nguyen (2021), who find that this sector is a key innovator in the energy transition space. Coal Mining (SIC12) displays high exposure to regulatory and physical shocks (Panels C and D). While high regulatory exposure is expected given the large emissions associated with burning coal, high physical exposure is less obvious. This may reflect mining firms' exposure to heavy precipitation, or heat, which pose physical challenges to their operations. Stone, Clay & Glass Products (SIC32), in the top-5 for $CCExposure^{Reg}$, includes mostly cement producers among its top-ranked firms (they belong to the largest CO_2 emitters). A sector in the top-10 of $CCExposure^{Phy}$ (Panel D) is the insurance industry, which, unsurprisingly, is highly exposed to the costs of storms or flooding.

The large variation in exposure *between* sectors masks important heterogeneity *within* each sector (apparent from the large within-sector standard deviations). To illustrate this heterogeneity, we compare TotalEnergies and ExxonMobil. Both firms operate in Petroleum Refining (SIC29), a sector ranking among the top 10 for $CCExposure^{Opp}$ and $CCExposure^{Reg}$. In terms of the average regulatory exposure since 2010, TotalEnergies' score is similar to that of ExxonMobil ($CCExposure^{Reg}_{TotalEnergies}$=0.21 versus $CCExposure^{Reg}_{ExxonMobil}$=0.18), but the French oil major exhibits much higher average opportunity exposure ($CCExposure^{Opp}_{TotalEnergies}$=1.13 versus $CCExposure^{Opp}_{ExxonMobil}$=0.15). This divergence reflects a broader perception in the market about the extent to which these firms embrace renewable energy and the net-zero transition in their business models (see Pickl (2019)). More generally, the large within-industry variation indicates that sectors have "winners" and "losers." Investors may therefore be able to address climate risks and opportunities by maintaining a broad industry diversification (rather than banning some industries) and then performing negative screening of climate change "losers." This observation echoes arguments by both academics (Andersson, Bolton, and Samama (2016)) and providers of low-carbon index solutions.

B.2. Climate Change Exposure: Times-Series Variation

In Figure 3, Panels A to D, we compute the cross-sectional means for $CCExposure$ and the topic-based measures and plot them over time (for each measure, we focus on top-10 sectors). This figure also highlights key moments in public awareness of climate change, covering climate policy events relevant to regulatory and opportunity shocks (Panels B and C), select physical shocks (Panel D), or both (Panel A). In Panel A, $CCExposure$ generally increases over the sample period, especially since the mid-2000s. The increase in the early years indicates that earnings calls discussed climate issues earlier than we might have expected. A plateau is reached around 2009 (the year of the unsuccessful Copenhagen Climate Summit). We then observe a slight decline in the years leading up to the 2012 Doha Climate Summit. We note a renewed increase in $CCExposure$ since around 2013. At the end of the sample, $CCExposure$ peaks with earnings calls exhibiting about four climate change bigrams per 1,000 bigrams; this compares to about 0.1 political bigrams per 1,000 bigrams in Hassan et al. (2019).

Insert Figure 3 about here.

In Panel B, the time series for $CCExposure^{Opp}$ is similar to that of the overall measure: $CCExposure^{Opp}$ trends upward, especially at the beginning of the sample. In Panel C,

$CCExposure^{Reg}$ increases between 2002 and 2008, varies around a markedly lower level between 2011 and 2013, spikes in 2015 (Paris Agreement), and follows an increasing trend since 2017. This is consistent with intensified policy discussions about how to achieve the Paris goals. In Panel D, $CCExposure^{Phy}$ displays more swings than the other measures, albeit also around an upwards trend. It appears that $CCExposure^{Phy}$ does not strongly reflect highly salient climate events. For example, while there is a jump after major U.S. hurricanes (i.e., Katrina, Sandy, and Harvey), the jumps occur with a considerable lag. This pattern indicates that $CCExposure^{Phy}$ primarily reflects firm-specific exposures to physical climate events, (e.g., local heat waves or droughts).

B.3. Climate Change Exposure and Carbon Emissions

We explore how well the exposure measures correlate with firms' carbon emissions. Carbon emissions constitute an essential variable to measure firm-level exposure to climate change, especially for regulatory shocks (Bolton and Kacperczyk (2021, 2023)). The analysis of carbon emissions is also the most frequently used climate risk management tool of institutional investors (Krueger, Sautner, and Starks (2020)). A benefit of using carbon emissions is that they are easy to understand and compute, readily available for subscribers of environmental, social, and governance (ESG) databases, and genuinely related to changes in the global climate.

We expect that regulatory climate topics arise more frequently in earnings calls of large carbon emitters, as they are more strongly affected by carbon taxes or related regulations. At the same time, regulatory threats related to emissions may also spur technological innovation that provides firms with opportunities in the marketplace.[25] Furthermore,

[25] For example, utilities with a large carbon footprint may have strong incentives to develop low-carbon alternatives (e.g., wind farms, solar farms), which provide future opportunities. Indeed, as mentioned above, Cohen, Gurun, and Nguyen (2021) demonstrate that some of the largest carbon emitters produce more and better green innovation than other firms.

some firms' emissions may be "good" in supporting the transition to a greener economy; these firms, called "climate enablers," include, for example, manufacturers of building materials that help houses become more energy-efficient. Finally, carbon emissions should be unrelated to the exposure to physical shocks at the firm level.

We test these predictions by regressing the exposure measures on lagged emission values (we use lagged values because emissions covering year $t-1$ are reported in year t). Table IV, Panel A, reports the results. In column (1), we observe a strong positive association between $Total\ Emissions$ and $CCExposure$. As predicted, this association originates from positive correlations between emissions and both $CCExposure^{Opp}$ (column (2)) and $CCExposure^{Reg}$ (column (3)). A one-standard-deviation increase in the emissions variable is associated with an increase in $CCExposure^{Reg}$ equal to 23% of its standard deviation (using values for the regression sample). In column (4), we find no association between emissions and physical exposure.

Insert Table IV about here.

B.4. Climate Change Exposure and Public Attention to Climate Change

Time-series variation in public attention to climate change, as proxied by *WSJ CC News Index*, has been shown to affect financial market participants (e.g., Choi, Gao, and Jiang (2020) or Ilhan, Sautner, and Vilkov (2021)). Accordingly, we expect earnings call discussions to react to the salience of climate topics in the public arena. Indeed, Table IV, Panel B, shows that measured climate change exposure is higher at times when public climate attention rises. In column (1), a one-standard-deviation increase in *WSJ CC News Index* is associated with an increase in $CCExposure$ of 0.05 (5% of the mean within the regression sample). This effect reflects a positive association between *WSJ CC News Index* and both $CCExposure^{Opp}$ and $CCExposure^{Reg}$. Hence, when public climate attention is high, earnings calls discuss regulatory shocks and climate opportunities more extensively. Higher values of *WSJ CC News Index* do not translate into more discussions of physical shocks. This suggests that $CCExposure^{Phy}$ mostly captures firm-specific physical shocks, rather than economy-wide shocks that make it to the WSJ (this conclusion is consistent with the time-series evidence in Figure 3).

IV. Variance Decomposition and Role of Measurement Error

A. Variance Decomposition

We conduct a variance analysis to examine the extent to which $CCExposure$ and its components quantify firm-level variation in climate change exposure. Table V reports the incremental explanatory power from conditioning the exposure measures on fixed effects that plausibly drive the variation. Time fixed effects (i.e., economy-wide changes in aggregate exposure) explain little variation, yielding an incremental R^2 below 1% for each measure. For industry fixed effects, the same observation holds only for $CCExposure^{Phy}$. In contrast, exposures to opportunity or regulatory shocks have a sizeable industry component (17% and 8%, respectively), which might stem from regulation targeting specific industries or technological developments affecting entire sectors. The interaction between industry and time fixed effects accounts for, at most, an additional 2.6% of the variation (in the case of $CCExposure^{Opp}$). Country fixed effects provide little additional explanatory power, which mitigates concerns that our measures are strongly affected by the native language in a country or how distant this language is from English. Depending on the measure, between 70% and 96% of the variation is *unexplained* by these sets of fixed effects. Thus, variation plays out at the firm level, rather than at the level of the country, industry, or over time. (The high unexplained variation for $CCExposure^{Phy}$ is unsurprising given that exposure to physical shocks depends highly on the location of a firm's production sites or insurance policies.) Adding firm fixed effects, permanent differences across firms in an industry and country account for 52%, 56%, 45%, and 45% of the variation of $CCExposure$, $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$, respectively. The remaining 48%, 44%, 55%, and 55%, respectively, come from variation over time in the identity of firms in industries and countries most affected by the respective climate change variables.

Insert Table V about here.

B. Assessing Measurement Error

We interpret the large share of variance within the firm-year as capturing economically meaningful heterogeneity. Under this view, a firm's idiosyncratic exposure to climate change is the key determinant of the measured variation. A plausible alternative explanation is that part of the firm-level variation reflects idiosyncratic measurement error. We conduct several tests that dispel this alternative. First, we note that we find robust associations between $CCExposure$ and important real and financial outcomes (as do other papers). These findings suggest that the variation reflected in firm-level $CCExposure$ is not simply noise.

Second, following Hassan et al. (2019), we quantify the amount of measurement error contained in the firm-level variation by assuming that a firm's "true" exposure follows a first-order auto-regressive (AR) process. We then assume that $CCExposure$ measures this true exposure with classical (i.i.d.) measurement error.[26] Suppose a valid instrument for (lagged) $CCExposure_{i,t-1}$ were available. In this case, one could back out the share of its variation consisting of measurement error by comparing the OLS and instrumental

[26] Under these assumptions, if the correlation between two different lags of the firm-year data is known, the AR(1) parameter and the estimated measurement error can be backed out. For example, if the first lag has a correlation of 0.45 (=0.5*0.9) and the second lag a correlation of 0.41 (=0.5*0.9*0.9), that would imply measurement error of 50% of the variation and an AR coefficient of 0.9. If the first lag has a correlation of 0.9 and the second 0.8, this implies no measurement error and an AR coefficient of 0.9.

variable (IV) coefficients. Intuitively, the idea is that candidate IVs measure true climate change exposure with error. Under the i.i.d. assumption, the measurement error in the IV is uncorrelated with that in $CCExposure_{i,t}$ and thus can be used to "purge" the latter's measurement error. For this procedure to work, we do not assume that the IV has lower measurement error—indeed, it is likely to have higher measurement error. We assume only that the measurement error in the IV and in measured climate change exposure are statistically independent.

Table VI shows three implementations of this idea. One implementation uses an alternative exposure measure constructed by applying our algorithm to the "Management Discussion and Analysis" (MD&A) section in firms' annual 10K filings. The two other implementations use lags of this alternative measure and $CCExposure$ itself as instruments. While the estimates of the share of measurement error in $CCExposure$ vary somewhat across the three approaches, approximately 5% to 10% of the variation in measured $CCExposure$ is due to measurement error.[27] The implied measurement error at the firm level (in Panel B) is about 2 percentage points higher than in the overall variation (Panel A). Although we interpret these results with due caution, they suggest that measurement error in the firm-level dimension is higher than in the overall panel, but only modestly. Thus, concerns that the variation displayed at the firm level is subject to more measurement error than the overall climate change exposure measure (before any fixed effects) are not substantiated.

Insert Table VI about here.

[27]These estimates compare favorably to the amount of measurement error found using similar assumptions in firm-level variables measured using accounting data (e.g., measures of total factor productivity constructed by Bloom et al. (2018) and Collard-Wexler (2011)).

V. Economic Applications

A. *Real Outcomes: Green-Tech Jobs and Green Patents*

Significant climate-related innovation is required to reach net-zero emissions by 2050 (Stern and Valero (2021)), implying huge investments by firms in human capital and R&D. According to some estimates, incremental investments of $50 trillion are needed in solar technology, decarbonization, energy efficiency, or carbon capture by 2050 (World Economic Forum (2021)). To illustrate that our exposure measures help predict real outcomes related to the net-zero transition, we relate next year's creation of disruptive green-tech jobs and green patents to this year's values of climate change exposure. Among the sampled U.S. firms, for firm i and year t we estimate

$$Green\,Outcome_{i,t+1} = exp(\alpha_i + \beta\ \log(1 + CCExposure_{i,t}) + \gamma \mathbf{X}_{i,t} + \delta_j \times \delta_t + \epsilon_{i,t+1}), \tag{5}$$

where $Green\,Outcome_{i,t+1}$ is $\#Green\text{-}Tech\,Jobs_{i,t+1}$ or $\#Green\,Patents_{i,t+1}$ in year $t+1$ and $CCExposure$ is the climate change exposure measure in year t (we include the overall and topic-based measures). The vector $\mathbf{X}_{i,t}$ includes *Log(Assets)*, *Debt/Assets*, *Cash/Assets*, *PP&E/Assets*, *EBIT/Assets*, *CAPEX/Assets*, and *R&D/Assets*. The variables $\delta_j \times \delta_t$ represent industry-year fixed effects. We account for industry shocks that vary over time, as firm-level innovation-related activity contains a large time-varying industry component (Aghion et al. (2005)). As demonstrated in Table V, such variation is also an important determinant of climate change exposure, making it important to identify effects *within* industry-year pairs. We cluster standard errors at the industry-year group level.

We estimate equation (5) using Poisson regressions, which have two advantages (Cohn, Liu, and Wardlaw (2022)). First, Poisson regressions account for the distributional characteristics of our count-based outcomes (they provide unbiased estimates for dependent variables with a large mass of values at zero combined with severe skewness). Second, Poisson regressions allow use to include industry-year fixed effects without biasing the estimation. They thus address the issue of separable group fixed effects (in our case at the industry-year level) by basing the estimation only on observations with at least one nonzero value within a group. This is desirable, as it restricts the usable sample to those groups that are informative about the effects of $CCExposure$.[28] For robustness, we also estimate linear and log1plus-linear models (with and without industry-year fixed effects) on the unrestricted sample (we interpret these models' estimates with caution).

The estimation results for $\#GreenTechJobs$ are reported in Table VII. In column (1), the estimates show that firms with higher overall exposure post more vacancies for jobs in disruptive green technologies over the subsequent year. A one-standard-deviation increase in $CCExposure$ is associated with a 109% increase in the number of green-tech jobs over the next year.[29] Columns (2) to (4) consider the topic-based measures. As expected, the overall exposure effect is due in large part to high-opportunity firms (column (2)). Firms with higher regulatory exposure also plan to hire more green-tech workers than firms with lower exposure (column (3)). We do not find that firms with larger physical exposure post more green-tech jobs (column (4)). In column (5), we continue to find that $CCExposure$ positively predicts green-tech hiring if we replace *#Green-Tech Jobs*

[28] Cohn, Liu, and Wardlaw (2022) show that log1plus-linear models may be biased in our context. The admission of separable group fixed effects in Poisson regressions differs from that in other nonlinear count-data models. These alternative models are subject to the incidental parameter problem, which leads to biased and inconsistent estimates (Lancaster (2000)).

[29] In a Poisson model, for a regression coefficient β, the magnitude of a one-standard-deviation change in the independent variable is calculated as $e^{\beta \times STD} - 1$. This effect size (when multiplied by 100%) represents the percentage change in the dependent variable. We use the within-fixed-effects (rather than overall-panel) standard deviation to capture plausible variation. The large magnitude of the effect also indicates that the average number of disruptive green-tech jobs is relatively low.

with $I(Green\text{-}Tech\ Jobs)$, an indicator for whether a firm posts a green-tech job (we estimate a linear model with the same observations as in columns (1) to (4)). Similarly, in column (6) estimates are robust to using the ratio of green-tech jobs to all tech jobs (*Green-Tech Ratio*). Column (7) addresses the concern that high-exposure firms may simply recruit more personnel in disruptive technologies across the board, without a specific focus on green jobs per se (for example, because these firms happen to be more innovative). Specifically, we replace *#Green-Tech Jobs* with *#Nongreen-Tech Jobs* and reestimate the regression in column (1). We do not find positive predictive effects of the exposure measure, which mitigates concerns of spurious relationships. In fact, firms with higher climate change exposure hire less, not more, nongreen-tech jobs. Overall, the data are more consistent with a recruiting shift from nongreen-tech jobs to green-tech jobs, rather than a general expansion of tech-related hiring at high-exposure firms.

Insert Table VII about here.

The results for green-tech jobs broadly extend to green patents in Table VIII. In columns (1) to (3), firms with greater climate change exposure show more green patenting in the next year. A one-standard-deviation increase in $CCExposure$ is associated with a 72% increase in the number of green patents over the next year. The effect for $CCExposure^{Opp}$ is intuitive, as green innovation provides business opportunities during the net-zero transition. To illustrate the intuition behind the effects for $CCExposure^{Reg}$, the case of Caterpillar is insightful. This firm is not only the top green patent producer in our sample (see Section I.E), but it also exhibits high measured regulatory exposure. This latter feature stems from its legacy business related to mining and diesel engines (sample mean of $CCExposure^{Reg}_{Caterpillar}$=0.09, in the top decile of $CCExposure^{Reg}$). We do not find that firms with larger physical exposure generate more green patents (column (4)). In columns (5) and (6), we continue to find that $CCExposure$ predicts green patenting if we replace *#Green Patents* with an indicator for whether a firm created green patents (column (5)) or with the green patents ratio as in Cohen, Gurun, and Nguyen (2021) (column (6)). Column (7) shows that high-exposure firms are not simply more innovative in general; the estimates indicate fewer, not more, nongreen patents by firms with high values of $CCExposure$.

Insert Table VIII about here.

Table IA.XIII shows that the results in Tables VII and VIII are robust to controlling for carbon emissions. This finding demonstrates that our measures contain additional information beyond what is reflected in emissions (the sample size is reduced in the panel due to the lower number of observations on carbon emissions).

In Table IA.XIV, a series of alternative specifications continue to show that $CCExposure$ predicts green-tech job creation. In column (1), we dispel concerns related to strategic disclosure in earnings calls (Mayew (2008), Hassan et al. (2019)). One specific potential concern is that managers may want to distract attention from poor performance and strategically "cheap talk" about climate change (Hail, Kim, and Zhang (2021)). Following Hassan et al. (2019), we test for this possibility by adding a control for the firm's overall sentiment (share of positive and negative tone words across the earnings call transcript) and two proxies for recent performance.[30] The estimates show that our results are robust to adding these controls. In column (2), we restrict the sample to firm-years within the BG database to ensure that the results are unaffected by how we classify the firms missing in BG; recall that we assume no green-tech job creation for these firms (BG may systematically miss scraping some firms' postings). In column (3), exposure is based

[30]We measure performance as the pre-call stock return accumulated over the seven days prior to the earnings call and the earnings surprise. Earnings surprise is defined as earnings per share before extraordinary items minus earnings per share in the same quarter of the prior year, divided by the price per share at the beginning of the quarter (Ball and Bartov (1996)). We average the two variables across the earnings calls of a firm-year to obtain an annual measure.

on a count of bigrams in the Q&A session, that is, the part of the call that is less under management control and in turn less subject to concerns of strategic (non-)disclosure and greenwashing. In column (4), $CCSentiment^{Pos}$ strongly predicts next-year green-tech job creation, while $CCSentiment^{Neg}$ is insignificant (albeit marginally). In column (5), $CCRisk$ is positively associated with green-tech job creation. In column (6) to (9), results hold if we estimate OLS specifications to address potential concerns with the Poisson specification. We estimate models with and without industry-year fixed effects, and with *#Green-Tech Jobs* or $Log(1 + \#Green\text{-}Tech\ Jobs)$. We also provide estimates that replace the log1plus version of $CCExposure$ with an unlogged version. Table IA.XV applies the same alternative specifications to green patenting. The estimates show that our results continue to hold.

Table IA.XVI reports regressions for the subsamples in which the exposure measures that rely exclusively on the initial bigrams indicate zero exposure. In these estimations, our exposure measures continue to predict green outcomes. This finding corroborates the performance gain from using more subtle and less visible climate change bigrams, as the estimation is identified from the bigrams obtained through the keyword search algorithm.

Finally, Table IA.XVII documents the covariate balance of observations that are either included or excluded from the estimations in Tables VII and VIII. Excluded firm-years exhibit lower climate change exposure, implying that our estimates are obtained within the set of firms for which climate change issues are most pressing.

B. Financial Market Outcomes

B.1. Options Market Risks and Risk Premiums

Firms with higher regulatory exposure are more strongly affected by future regulations to combat global warming, and uncertainty over such regulations should be priced in the options market (Kelly, Pastor, and Veronesi (2016)). Likewise, climate opportunities are risky, with plenty of uncertainty surrounding investments in green technologies or renewable energy. We therefore test whether climate change exposure is related to option-implied risks and risk premiums. We consider three sets of risk variables. First, to quantify general risks, we use three implied central moments, namely, variance ($IVar$), skewness ($ISkew$), and kurtosis ($IKurt$). Second, we calculate two heuristic measures quantifying the relative expensiveness of protection against left ($SlopeD$) and right ($SlopeU$) tail risks.[31] Third, we use the variance risk premium (VRP) to measure the premiums that investors are willing to pay to hedge against general climate-related variance risk (or uncertainty, as suggested in Bali and Zhou (2016)). Using each of these variables, we run the regression:

$$OI\ Outcome_{i,t+1} = \alpha_i + \beta\ Log(1 + CCExposure)_{i,t} + \gamma \mathbf{X}_{i,t} + \delta_j \times \delta_t + \epsilon_{i,t+1}, \tag{6}$$

where $OI\ Outcome_{i,t+1}$ is an option-implied measure for firm i measured at the end of quarter t (i.e., a conditional expectation of some quantity over the period $t + 1$), and $CCExposure$ is firm i's climate change exposure in quarter t.[32] The vector $\mathbf{X}_{i,t}$ includes the same controls as before (delayed to be available in the third quarter after the annual close of the fiscal period). The variables δ_j and δ_t represent industry and year fixed

[31] The variable $SlopeD$ increases when the cost of left-tail protection goes up (relative to the cost of at-the-money (ATM) options), and $SlopeU$ decreases (becomes more negative) when the relative cost of obtaining upside growth increases. Note that Sautner et al. (2022) define their measure of $SlopeU$ as minus one times $SlopeU$.

[32] When computing quarterly versions of our measures, we encounter the issue that any specific earnings call in a year might not discuss climate change, even though the conversation turns to the issue in surrounding calls. These incidental gaps in the quarterly data (where the measured $CCExposure$=0) do not reflect business realities. Therefore, we pre-process the quarterly climate change exposure following a method outlined in Sautner et al. (2022), which exponentially smooths each metric for each firm with a half-life of three quarters.

effects, respectively. We cluster standard errors at the industry-year level.

Table IX, Panel A, documents that $CCExposure$ is strongly linked to forward-looking risks and risk premiums. In columns (2) and (3), $CCExposure$ predicts a more negatively skewed return distribution ($ISkew$) and fatter tails ($IKurt$). Furthermore, tail exposure is more costly for firms with higher climate change exposure. More specifically, downside protection in column (4) (positive and significant coefficient on $SlopeD$) and upside potential in column (5) (negative and significant coefficient on $SlopeU$) become more expensive when $CCExposure$ is higher. In terms of magnitudes, the effects are strongest in column (3) for $IKurt$. A one-standard-deviation change in $CCExposure$ is associated with a change in $IKurt$ equivalent to 7% of its standard deviation. The effects for $SlopeD$ and $SlopeU$ are 4.5% and 4.1%, respectively.

Insert Table IX about here.

The three remaining panels consider the topic-based measures. Earnings calls should contain more discussions of climate-related opportunities if a firm is well positioned for the growth potential arising from climate change. The realization of these opportunities could lead to large gains if successful and to large losses if unsuccessful. Investors may in turn trade in the options market to reflect the two-sided effects of climate opportunities. Panel B confirms this intuition: the tail effects for $CCExposure^{Opp}$ in columns (4) and (5) are similar compared to the corresponding estimates in Panel A. The magnitude of a one-standard-deviation increase in $CCExposure^{Opp}$ is 4.3% for $SlopeD$ and 3.9% for $SlopeU$, respectively. Thus, it is not only the case that options are more expensive on both tails if climate opportunities are higher, but also that the cost of upside potential grows faster than the cost of downside crash protection. The link between $CCExposure^{Opp}$ and VRP in column (6) demonstrates that the wedge between the implied and "historically fair" price of out-of-the-money (OTM) calls increases with opportunity exposure. Thus, investors are ready to pay an extra (volatility) premium when buying options on stocks with climate-related upside potential. However, the effect is small in magnitude and only marginally significant.

In Panel C, the pattern for $CCExposure^{Reg}$ is similar to that for $CCExposure^{Opp}$, though the magnitudes are smaller. While the right-tail option expensiveness increases by 2.6% of its standard deviation (i.e., $SlopeU$ diminishes) for a one-standard-deviation change in $CCExposure^{Reg}$, the crash protection grows by 2.3%. This confirms our earlier evidence that some firms with high regulatory exposure face downside risks *and* upside potential due to their green innovation activity. In Panel D, the effects for $CCExposure^{Phy}$ are similar to those of the other measures.[33]

Overall, climate change exposure is priced in the options market. Considering all the evidence, stocks with higher exposure have probability mass shifted to the tails of the distribution, making crash protection and upside potential relatively more expensive. Obtaining protection and upside growth potential comes at a premium, which increases more strongly for firms facing higher opportunities. We acknowledge that the effect magnitudes are modest and hardly tradeable after transaction costs.

B.2. Cross-Section of Stock Returns

Climate change exposure is related to risks and risk premiums in the options market. Consequently, systematic risk related to $CCExposure$ may be associated with a risk premium in the cross-section of returns. That said, testing for the pricing effects of a climate change exposure factor, labelled $CCEXPOSURE$, is challenging for several reasons.

[33] Our inference for the pricing of physical exposure is different from the link between hurricane uncertainty and variance pricing in Kruttli, Roth Tran, and Watugala (2021). For example, while we concentrate on the unconditional pricing using the expected variance risk premium, Kruttli, Roth Tran, and Watugala (2021) study dynamics of the realized variance risk premium. However, these authors also conclude that, especially in the early sample years, investors underprice variance in options of firms strongly exposed to extreme weather events.

A conceptual challenge arises because return effects are theoretically more ambiguous to predict compared to the risk measures. On the one hand, firms with high betas for *CCEXPOSURE* should be more risky and—in expectation—earn a risk premium.[34] On the other hand, the relations may actually be the opposite, with risks gradually getting priced in during the sample period; as risks emerge, stock prices decline, implying lower realized returns. Pastor, Stambaugh, and Taylor (2021) illustrate this difference between ex-ante and ex-post returns. An estimation challenge arises because $CCExposure$ reflects the attention devoted to climate topics at a point in time. This implies that the pricing of *CCEXPOSURE* should vary over time, requiring the estimation of conditional risk premiums. Another challenge arises because the number of assets for such tests is large relative to the time points available for the estimation—less than 20 years of data.

With these challenges in mind, we investigate the conditional pricing of *CCEXPOSURE* in the cross-section of stocks. We follow Jamilov, Rey, and Tahoun (2021) and construct the factor as an unexpected shock to the aggregate value of $CCExposure$. This involves three primary steps. First, we convert quarterly transcript-level values of $CCExposure_{i,t}$ for U.S-traded firms to a monthly frequency by propagating the last exposure values for up to three months forward (i.e., we match the month-year of each climate change exposure to the month-year of the respective quarterly transcript). Second, we compute cross-sectional monthly averages of $CCExposure_m$. Third, we take the first differences in these monthly averages as a proxy for unexpected monthly shocks to the aggregate exposure level, and use them as the *CCEXPOSURE* factor.[35]

[34] For example, such firms face higher uncertainty related to future developments in climate-related areas, that is, their valuation should include real option value depending on the path of climate-related technologies, regulations, or physical climate shifts.

[35] The factor is standardized to have zero mean and annual volatility of 10%. Results are robust to using the residuals from an AR(1) process fitted to the monthly exposure series, as implemented in Jamilov, Rey, and Tahoun (2021) (the resulting factors are almost perfectly correlated). However, fitting

To examine the conditional pricing of *CCEXPOSURE* among S&P500 firms, we follow Gagliardini, Ossola, and Scaillet (2016, GOS), who provide a conditional extension of the two-pass regression approach (Fama and MacBeth (1973)). We use this approach as it delivers good small-sample performance when—as in our case—the cross-section is large relative to the time series. GOS assume a linear conditional factor model for excess returns with time-varying factor exposures and risk premiums. They model the parameters as linear functions of lagged instruments. The factor loadings $\beta_{i,m}$ depend on stock-specific instruments ($Z_{i,m-1}$) as well as common instruments (Z_{m-1}), and the factor expectations only on common instruments. Under this framework, the conditional expected return on stock i in month m is

$$E[R_{i,m}|Z_{i,m-1}, Z_{m-1}] = \beta_{i,m}^{\intercal}\lambda_m, \tag{7}$$

where the risk premium λ_m is the sum of the conditional factor expectation $E[F_m|Z_{m-1}]$ and the process ν_m, estimated from the cross-section of stocks. The process ν_m allows the estimated risk premium to deviate from the conditional expectation of a factor due to market imperfections for tradeable factors (Cremers, Petajisto, and Zitzewitz (2013), GOS) and it also reveals an "implicit cost" of projecting a nontradeable factor (like ours) on returns. A similar framework is used, for example, in Barras and Malkhozov (2016). As in GOS, we use as common instruments the term spread and the default spread and as the stock-specific instrument the log of the book-to-market ratio (see Section I.G for definitions). We estimate the time-varying components of the risk premiums with the four-factor model by Carhart (1997) that is augmented with the *CCEXPOSURE* factor.[36]

When performing the estimation, we obtain average conditional risk premiums in line

an AR(1) process may introduce look-ahead bias.

[36]The factor is essentially orthogonal to the other factors, with all unconditional correlations being smaller than 0.05. The results are robust to using three- and five-factor models.

with expectations (risk premiums for the market, size, value, and momentum factors are 11.4%, 5.0%, -5.8%, and 8.5% per annum (p.a.), respectively). The *CCEXPOSURE* premium is positive, on average (3.7% p.a.), and we obtain positive point estimates for most months. More importantly, the risk premium is not constant over time, and we reject the hypotheses that its two components are constant (p-values of 0.0137 and 0.0001, respectively).

In Table X, we report the estimated annualized components of the risk premium λ_m, that is, the estimates of F and ν. Similar to the results in GOS, most of the action for the risk premiums comes through the cross-sectional component ν. For *CCEXPOSURE*, ν has a positive unconditional mean (constant of 3.73%) and a positive link to the default spread (3.13%)—both are highly significant. This indicates that stocks with high exposure to the *CCEXPOSURE* factor are expected to earn higher returns, especially when market-wide default risk increases.

Insert Table X about here.

The time series of the estimated risk premium on *CCEXPOSURE* is depicted in Figure 4. The series illustrates significant variability over time, with a large spike around the financial crisis. Further tentative interpretations indicate a temporary spike around the time of Hurricane Sandy (October 2012) and the Doha Climate Summit (November 2012). Another temporary spike occurs just after the Paris Agreement (December 2015). Considering the most recent five years, the risk premium was lowest around the time President Trump took office (January 2017); it gradually increased thereafter with a drop around the onset of the Covid pandemic.[37]

[37] As in the previous applications, we estimate the risk premiums separately by topic. The topic-based premiums are on average positive, but demonstrate distinct time-series patterns. For example, when the physical risk premium goes up, the opportunity risk premium tends to go down. We do not want to overemphasize the topic-based differences here, as our framework uses the same set of instruments for

Insert Figure 4 about here.

We emphasize that our objective is not to create an ultimate climate factor to be added to the factor zoo (Feng, Giglio, and Xiu (2020)), but instead to show that attention to climate topics in earnings calls is linked to systematic risk, with shocks to such attention potentially being priced in the cross-section (following a narrative as in Shiller (2017)).

VI. Conclusion

In this paper, we introduce a new method that identifies firm-level climate change exposure from word combinations signaling climate change conversation in earnings conference calls. As these calls reflect the demand side (analysts) and the supply side (management) of a "market for information," our measures reflect the combined views of key stakeholders about a firm's climate change exposure. Furthermore, earnings calls are largely forward-looking; while analysts review past results, they also spend much of their time probing management about future plans (Huang et al. (2018)).

Our measures build on recent work that identifies earnings calls as a source for identifying the various risks and opportunities that firms face over time. We adjust the approach of this prior work along several critical dimensions, allowing us to capture aspects of both the opportunities and the (physical and regulatory) risks associated with climate change. For this purpose, we adapt the machine-learning keyword discovery algorithm proposed by King, Lam, and Roberts (2017) to produce several sets of climate change bigrams. Rather than choosing a training library, we start with a short list of initial bigrams that most experts would agree are related to climate change. Our exposure measures capture the proportion of the earnings call related to climate change topics. These measures are available for a global sample of more than 10,000 firms covering the period 2002 to 2020. We demonstrate that our measures are helpful in predicting important real outcomes

all topic-based factors.

related to the net-zero transition, notably, green-tech growth and green patenting. We also document that the measures contain information that is priced in the options and equity markets.

Initial submission: May 10, 2021; Accepted: June 23, 2022

Editors: Stefan Nagel, Philip Bond, Amit Seru, and Wei Xiong

Appendix. Variable Definitions

Variable	Years	Definition
$CCExposure$	2002-2020	Relative frequency with which bigrams related to climate change occur in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCExposure^{Opp}$	2002-2020	Relative frequency with which bigrams that capture opportunities related to climate change occur in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCExposure^{Reg}$	2002-2020	Relative frequency with which bigrams that capture regulatory shocks related to climate change occur in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCExposure^{Phy}$	2002-2020	Relative frequency with which bigrams that capture physical shocks related to climate change occur in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCExposure^{Q\&A}$	2002-2020	Relative frequency with which bigrams related to climate change occur in the Q&A session part of transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the Q&A session. Source: Self-constructed.
$CCSentiment^{Pos}$	2002-2020	Relative frequency with which bigrams related to climate change are mentioned together with positive tone words that are summarized by Loughran and McDonald (2011) in one sentence in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCSentiment^{Neg}$	2002-2020	Relative frequency with which bigrams related to climate change are mentioned together with the negative tone words that are summarized by Loughran and McDonald (2011) in one sentence in the transcripts of earnings conference calls. Source: Self-constructed.
$CCRisk$	2002-2020	Relative frequency with which bigrams related to climate change are mentioned together with the words "risk" or "uncertainty" (or synonyms thereof) in one sentence in the transcripts of earnings conference calls. We count the number of such bigrams and divide by the total number of bigrams in the transcripts. Source: Self-constructed.
$CCExposure^{10K}$	2002-2020	Climate change exposure constructed by applying our algorithm to the "Management Discussion and Analysis" (MD&A) section in firms' annual 10K filings. Source: Self-constructed.
$Total\ Emissions$	2004-2020	Sum of annual Scope 1 and Scope 2 carbon emissions (metric tons of CO2) at the end of the year. Scope 1 emissions are caused by the combustion of fossil fuels or releases during manufacturing. Scope 2 emissions originate from the purchase of electricity, heating, or cooling. Source: Trucost.
$WSJ\ CC\ News\ Index$	2002-2017	Time-series index of the fraction of the *Wall Street Journal* dedicated to the topic of climate change. Source: Engle et al. (2020).

Variable	Years	Definition
#Green-Tech Jobs	2007, 2010-2020	Number of job postings for disruptive green-tech jobs in a year according to the Burning Glass (BG) database. Disruptive green-tech job postings relate to jobs in one of four climate-related technology areas identified by Bloom et al. (2021) as having been disruptive ("hybrid vehicle electric car," "lithium battery," "solar power," and "fracking"). We assume that no disruptive green-tech job has been posted if a firm-year is not included in the BG database. Source: Bloom et al. (2021) and Burning Glass.
I(Green-Tech Jobs)	2007, 2010-2020	Indicator equal to one if $\#Green - Tech\ Jobs$ is positive, and zero otherwise. Source: Bloom et al. (2021) and Burning Glass.
Green-Tech Ratio	2007, 2010-2020	Number of job postings for disruptive green-tech jobs relative to the total number of all disruptive job postings. Set to zero if the number of disruptive job postings is zero. Source: Bloom et al. (2021) and Burning Glass.
#Nongreen-Tech Jobs	2007, 2010-2020	Number of job postings for nongreen disruptive tech jobs in a year according to the Burning Glass (BG) database. Nongreen disruptive tech job postings relate to jobs in one of 25 climate-related technology areas identified by Bloom et al. (2021) as having been disruptive and are unrelated to climate change. We assume that no nongreen disruptive tech job has been posted if a firm-year is not included in the BG database. Source: Bloom et al. (2021) and Burning Glass.
#Green Patents	2002-2019	Number of green patents obtained in a year according to the Google Patents (GP) database. To identify "green" patents, we follow Cohen, Gurun, and Nguyen (2021) and apply the OECD classification to identify what constitutes a patent with the potential to address environmental problems. We assume that no green patenting has occurred if we are unable to identify a green patent in the GP database for a firm-year. Source: Google Patents.
I(Green Patents)	2002-2019	Indicator equal to one if *#Green Patents* is positive, and zero otherwise. Source: Google Patents.
Green Patents Ratio	2002-2019	Number of green patents (*#Green Patents*) relative to the total number of patents. Set to zero if the number of total patents is zero. Source: Google Patents.
#Nongreen Patents	2002-2019	Number of nongreen patents obtained in a year according to the Google Patents (GP) database. We assume that no patenting has occurred if we are unable to identify a nongreen patent in the GP database for a firm-year. Source: Google Patents.
Assets	2002-2020	Total assets (in $ millions) at the end of the year (Compustat item AT). Winsorized at the 1% level. Source: Compustat NA/Global
Debt/Assets	2002-2020	Sum of the book value of long-term debt (Compustat data item DLTT) and the book value of current liabilities (DLC) divided by total assets (Compustat data item AT). Winsorized at the 1% level. Source: Compustat NA/Global.
Cash/Assets	2002-2020	Cash and short-term investments (Compustat data item CHE) divided by total assets (Compustat data item AT). Winsorized at the 1% level. Source: Compustat NA/Global.
PPE/Assets	2002-2020	Property, plant, and equipment (Compustat data item PPENT) divided by total assets (Compustat data item AT). Winsorized at the 1% level. Source: Compustat NA/Global.

Variable	Years	Definition
$EBIT/Assets$	2002-2020	Earnings before interest and taxes (Compustat data item EBIT) divided by total assets (Compustat data item AT). Winsorized at the 1% level. Source: Compustat NA/Global
$R\&D/Assets$	2002-2019	R&D expenditures (Compustat data item XRD) divided by total assets (Compustat data item AT). Missing values set to zero. Winsorized at the 1% level. Source: Compustat NA/Global.
$CAPEX/Assets$	2002-2020	Capital expenditures (Compustat data item CAPX) divided by total assets (Compustat data item AT). Winsorized at the 1% level. Source: Compustat NA/Global.
$IVar$	2002-2020	Implied variance of log returns computed from 30-day out-of-the-money options following Bakshi, Kapadia, and Madan (2003). Winsorized at the 1% level. Source: Ivy DB Option-Metrics Volatility Surface File.
$ISkew$	2002-2020	Implied skewness of log returns computed from 30-day out-of-the-money options following Bakshi, Kapadia, and Madan (2003). Winsorized at the 1% level. Source: Ivy DB Option-Metrics Volatility Surface File.
$IKurt$	2002-2020	Implied kurtosis of log returns computed from 30-day out-of-the-money options following Bakshi, Kapadia, and Madan (2003). Winsorized at the 1% level. Source: Ivy DB OptionMetrics Volatility Surface File.
$SlopeD$	2002-2020	Slope of the implied volatility smile on the left side from the at-the-money level (i.e., for negative returns relative to ATM), computed as the slope coefficient from regressing implied volatilities of out-of-the-money puts on the respective option deltas (and a constant). The variable is computed from 30-day options. Winsorized at the 1% level. Source: Ivy DB OptionMetrics Volatility Surface File.
$SlopeU$	2002-2020	Slope of the implied volatility smile on the right side from the at-the-money level (i.e., for positive returns relative to ATM), computed as the slope coefficient from regressing implied volatilities of out-of-the-money calls on the respective option deltas (and a constant). The variable is computed from 30-day options. Winsorized at the 1% level. Source: Ivy DB OptionMetrics Volatility Surface File.
VRP	2002-2020	Variance risk premium computed as the difference between the implied variance of log returns ($IVar$) and the realized variance of daily log returns over a historical monthly window. Winsorized at the 1% level. Source: Ivy DB OptionMetrics Volatility Surface File for options data and CRSP for daily returns.

REFERENCES

Acemoglu, Daron, Philippe Aghion, Lint Barrage, and David Hemous, 2019, Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution, Working paper, MIT.

Addoum, Jawad, David Ng, and Ariel Ortiz-Bobea, 2020, Temperature shocks and establishment sales, *Review of Financial Studies* 33, 1331–1366.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt, 2005, Competition and innovation: An inverted-u relationship, *Quarterly Journal of Economics* 120, 701–728.

Andersson, Mats, Patric Bolton, and Frédéric Samama, 2016, Hedging climate risk, *Financial Analysts Journal* 72, 13–32.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.

Bakshi, Gurdip S., Nikunj Kapadia, and Dilip B. Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bali, Turan G., and Hao Zhou, 2016, Risk, uncertainty, and expected returns, *Journal of Financial and Quantitative Analysis* 51, 707–735.

Ball, Ray, and Eli Bartov, 1996, How naive is the stock market's use of earnings information?, *Journal of Accounting and Economics* 21, 319–337.

Barras, Laurent, and Aytek Malkhozov, 2016, Does variance risk have two prices? Evidence from the equity and option markets, *Journal of Financial Economics* 121, 79–92.

Baz, Salim, Lara Cathcart, Alexander Michaelides, and Yi Zhang, 2022, Firm-level climate regulatory exposure, Working paper, Imperial College London.

Berkman, Henk, Jonathan Jona, and Naomi S. Soderstrom, 2019, Firm-specific climate risk and market valuation, Working paper, University of Auckland Business School.

Bingler, Julia Anna, Mathias Kraus, and Markus Leippold, 2022, Cheap talk and cherry-picking: What ClimateBert has to say on corporate climate risk disclosures, *Finance Research Letters* 47, 102776.

Bloom, Nicholas, Max Floetotto, Nir Jaimovich, Itay Saporta-Eksten, and Stephen Terry, 2018, Really uncertain business cycles, *Econometrica* 86, 1031–1065.

Bloom, Nicholas, Tarek A. Hassan, Aakash Kalyani, Josh Lerner, and Ahmed Tahoun, 2021, The diffusion of disruptive technologies, Working paper, Stanford University.

Bolton, Patrick, and Marcin Kacperczyk, 2021, Do investors care about carbon risk?, *Journal of Financial Economics* 142, 517–549.

Bolton, Patrick, and Marcin Kacperczyk, 2022, Carbon disclosure and the cost of capital, Working paper, Imperial College London.

Bolton, Patrick, and Marcin Kacperczyk, 2023, Global pricing of carbon-transition risk, *Journal of Finance* forthcoming.

Bolton, Patrick, Marcin Kacperczyk, and Frédéric Samama, 2021, Carbon disclosure and the cost of capital, Working paper, Imperial College London.

Campello, Murillo, Janet Gao, and Qiping Xu, 2021, Personal income taxes and labor downskilling: Evidence from 27 million job postings, Working paper, Kelley School of Business.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Chen, Shuping, Stephan Hollander, and Kelvin Law, 2014, The price of silence: When no one asks questions during conference calls, Working paper, University of Texas at Austin.

Choi, Garwin, Zhenyu Gao, and Wenxi Jiang, 2020, Attention to global warming, *Review of Financial Studies* 33, 1112–1145.

Cohen, Lauren, Umit Gurun, and Quoc Nguyen, 2021, The ESG-innovation disconnect: Evidence from green patenting, Working paper, Harvard Business School.

Cohn, Jonathan B, Zack Liu, and Malcolm Wardlaw, 2022, Count (and count-like) data in finance, Working paper, University of Texas at Austin.

Collard-Wexler, Allan, 2011, Productivity dispersion and plant selection in the ready-mix concrete industry, Working paper, NYU Stern.

Cremers, Martijn, Antti Petajisto, and Eric Zitzewitz, 2013, Should benchmark indices have alpha? Revisiting performance evaluation, *Critical Finance Review* 2, 1–48.

Darendeli, Alper, Kelvin Law, and Michael Shen, 2022, Green new hiring, *Review of Accounting Studies* 27, 986–1037.

De Haas, Ralph, and Alexander Popov, 2022, Finance and carbon emissions, *The Economic Journal* forthcoming.

Di Giuli, Alberta, Alexandre Garel, Roni Michaely, and Arthur Petit-Romec, 2022, Climate change and mutual fund voting on environmental proposals, Working paper, European Corporate Governance Institute.

Duong, Huu Nhan, Petko S. Kalev, Madhu Kalimipalli, and Saurabh Trivedi, 2021, Do firms benefit from carbon risk management? Evidence from the credit default swaps market, Working paper, Monash University.

Engle, Robert F., Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2020, Hedging climate change news, *Review of Financial Studies* 33, 1184–1216.

Fama, Eugene F., and Kenneth R. French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653–1678.

Fama, Eugene F., and James D. MacBeth, 1973, Risk return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Feng, Guanhao, Stefano Giglio, and Dacheng Xiu, 2020, Taming the factor zoo: A test of new factors, *Journal of Finance* 75, 1327–1370.

Gagliardini, Patrick, Elisa Ossola, and Olivier Scaillet, 2016, Time-varying risk premium in large cross-sectional equity data sets, *Econometrica* 84, 985–1046.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, 2019, Text as data, *Journal of Economic Literature* 57, 535–574.

Giglio, Stefano, Bryan Kelly, and Johannes Stroebel, 2021, Climate finance, *Annual Review of Financial Economics* 13, 15–36.

Ginglinger, Edith, and Quentin Moreau, 2022, Climate risk and capital structure, Working paper, Université Paris–Dauphine.

Gostlow, Glen, 2021, The materiality and measurement of physical climate risk: Evidence from Form 8-K, Working paper, London School of Economics.

Hail, Luzi, Shawn Kim, and Rachel Xi Zhang, 2021, How do managers greenwash? Evidence from earnings conference calls, Working paper, University of Pennsylvania Wharton School.

Haščič, Ivan, and Mauro Migotto, 2015, Measuring environmental innovation using patent data, Working paper, OECD.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *Quarterly Journal of Economics* 134, 2135–2202.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2021, Firm-level exposure to epidemic diseases: Covid-19, SARS, and H1N1, NBER Working paper 26971.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2023a, The global impact of Brexit uncertainty, *Journal of Finance* forthcoming.

Hassan, Tarek A., Jesse Schreger, Markus Schwedeler, and Ahmed Tahoun, 2023b, Sources and transmission of country risk, *Review of Economic Studies* forthcoming.

Heath, Davidson, Daniele Macciocchi, Roni Michaely, and Matthew C. Ringgenberg, 2022, Does socially responsible investing change firm behavior?, Working paper, University of Utah.

Heo, Yuna, 2021, Climate change exposure and firm cash holdings, Working paper, Rutgers Business School.

Hollander, Stephan, Maarten Pronk, and Erik Roelofsen, 2010, Does silence speak? An empirical analysis of disclosure choices during conference calls, *Journal of Accounting Research* 48, 531–563.

Hong, Harrison, Frank Weikai Li, and Jiangmin Xu, 2019, Climate risks and market efficiency, *Journal of Econometrics* 208, 265–281.

Huang, Allen H., Reuven Lehavy, Amy Y. Zang, and Rong Zheng, 2018, Analyst information discovery and interpretation roles: A topic modeling approach, *Management Science* 64, 2833–2855.

Ilhan, Emirhan, Zacharias Sautner, and Grigory Vilkov, 2021, Carbon tail risk, *Review of Financial Studies* 34, 1540–1571.

Jamilov, Rustam, Hélène Rey, and Ahmed Tahoun, 2021, The anatomy of cyber risk, NBER Working paper 28906.

Kelly, Bryan, Dimitris Papanikolaou, Amit Seru, and Matt Taddy, 2021, Measuring technological innovation over the long run, *American Economic Review: Insights* 3, 303–320.

Kelly, Bryan, Lubos Pastor, and Pietro Veronesi, 2016, The price of political uncertainty: Theory and evidence from the option market, *Journal of Finance* 71, 2417–2480.

King, Gary, Patrick Lam, and Margaret E. Roberts, 2017, Computer-assisted keyword and document set discovery from unstructured text, *American Journal of Political Science* 61, 971–988.

Kogan, Leonid, Dimitris Papanikolaou, Amit Seru, and Noah Stoffman, 2017, Technological innovation, resource allocation, and growth, *Quarterly Journal of Economics* 132, 665–712.

Kölbel, Julian F., Markus Leippold, Jordy Rillaerts, and Qian Wang, 2022, Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure, *Journal of Financial Econometrics* forthcoming.

Krueger, Philipp, Zacharias Sautner, and Laura T. Starks, 2020, The importance of climate risks for institutional investors, *Review of Financial Studies* 33, 1067–1111.

Kruttli, Mathias S., Brigitte Roth Tran, and Sumudu W. Watugala, 2021, Pricing Poseidon: Extreme weather uncertainty and firm return dynamics, Working paper, Federal Reserve Bank of San Francisco.

Lancaster, Tony, 2000, The incidental parameter problem since 1948, *Journal of Econometrics* 95, 391–413.

Lewellen, Jonathan, and Stefan Nagel, 2006, The conditional CAPM does not explain asset pricing anomalies, *Journal of Financial Economics* 82, 289–314.

Li, Fengfei, Chen Lin, and Tse-Chun Lin, 2022, A one-two punch to the economy: Climate vulnerability and corporate innovation strategies, Working paper, The University of Hong Kong.

Li, Qing, Hongyu Shan, Yuehua Tang, and Vincent Yao, 2021, Corporate climate risk: Measurement and response, Working paper, University of Florida.

Loughran, Tim, and Bill McDonald, 2011, When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks, *Journal of Finance* 66, 35–65.

Mayew, William J., 2008, Evidence of management discrimination among analysts during earnings conference calls, *Journal of Accounting Research* 46, 627–659.

Mueller, Isabella, and Eleonora Sfrappini, 2022, Climate change-related regulatory risks and bank lending, Working paper, IWH-Halle.

Pankratz, Nora, and Christoph Schiller, 2021, Climate change and adaptation in global supply-chain networks, Working paper, Board of Governors of the Federal Reserve System.

Pastor, Lubos, Robert Stambaugh, and Lucian Taylor, 2021, Sustainable investing in equilibrium, *Journal of Financial Economics* 142, 550–571.

Pickl, Matthias, 2019, The renewable energy strategies of oil majors – From oil to energy?, *Energy Strategy Reviews* 29, 1–7.

Sautner, Zacharias, Laurence van Lent, Grigory Vilkov, and Ruishen Zhang, 2022, Pricing climate change exposure, *Management Science* forthcoming.

Shiller, Robert J., 2017, Narrative economics, *American Economic Review* 107, 967–1004.

Stern, Nicholas, and Anna Valero, 2021, Innovation, growth and the transition to net-zero emissions, *Research Policy* 50, 104293.

Tomar, Sorabh, 2021, Greenhouse gas disclosure and emissions benchmarking, Working paper, Southern Methodist University, Cox School of Business.

van Zaanen, Menno, and Pieter Kanters, 2010, Automatic mood classification using tf*idf based on lyrics, *Proceedings of the 11th International Society for Music Information Retrieval Conference* 75–80.

Varini, Francesco S., Jordan Boyd-Graber, Massimiliano Ciaramita, and Markus Leippold, 2020, Climatext: A dataset for climate change topic detection, Working paper, ETH Zurich.

von Schickfus, Marie-Theres, 2021, Institutional investors, climate policy risk, and directed innovation, Working paper, ifo Institute Munich.

Webersinke, Nicolas, Mathias Kraus, Julia Anna Bingler, and Markus Leippold, 2021, ClimateBert: A pretrained language model for climate-related text, Working paper, University of Zurich.

World Economic Forum, 2021, Financing the transition to a net-zero future, Report.

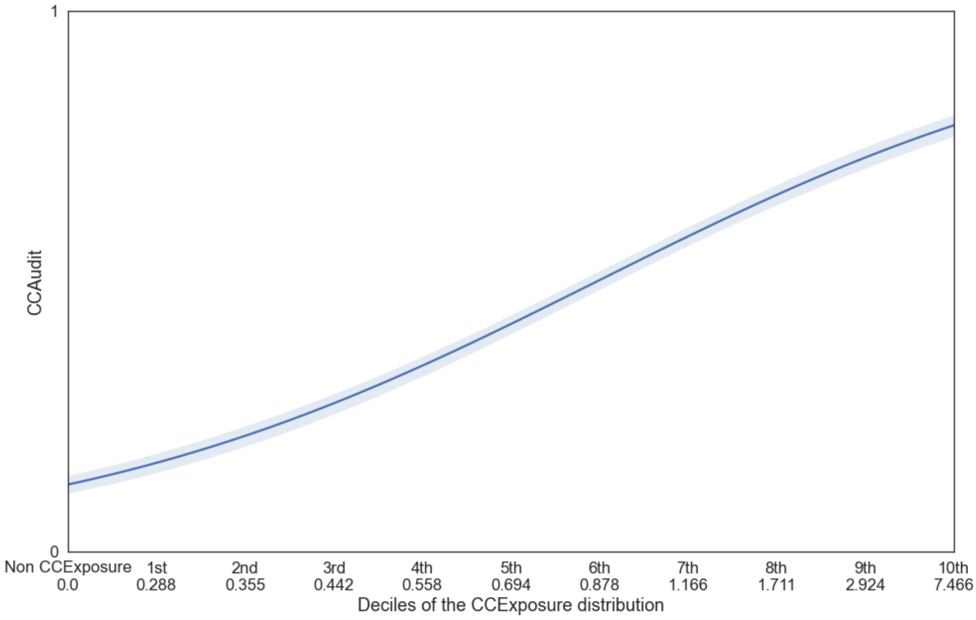


Figure 1. Probability of correctly identified positives by decile. This figure plots on the vertical axis the predicted probability of having a correctly identified positive (i.e., the audit study of the snippet confirms climate change-related text) against deciles of the $CCExposure$ distribution. The median score of $CCExposure$ in a given decile is shown on the axis. Predicted probabilities are computed by estimating a logit model on the sample of 2,090 audited snippets.

Panel A. Initial keywords Panel B. Initial keywords with topics

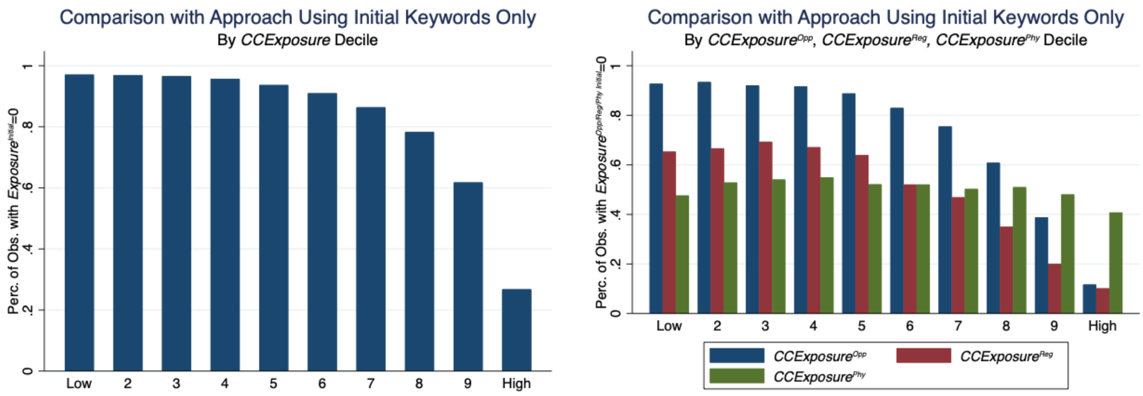
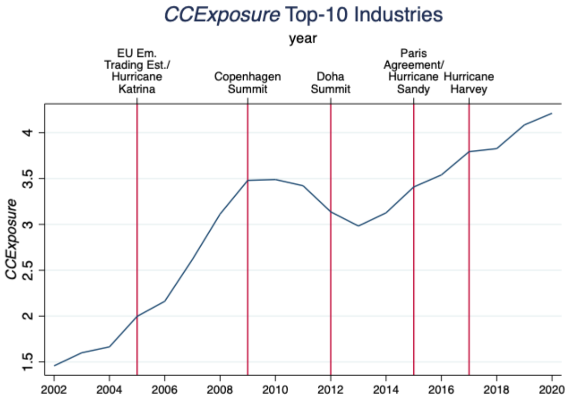
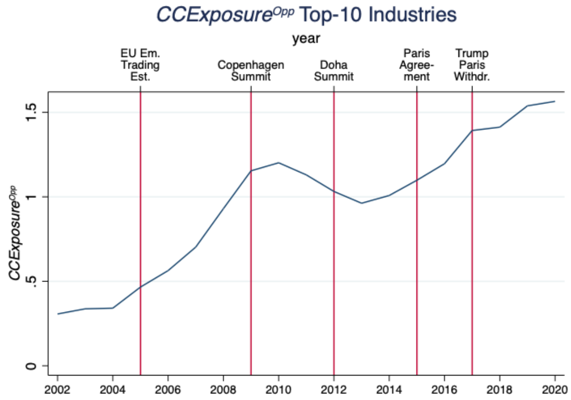


Figure 2. Climate change exposure calculated with initial bigrams. This figure shows how frequently $CCExposure^{Initial}$ signals zero climate change exposure, while $CCExposure$ instead reveals that such exposure exists. Results are reported by $CCExposure$ decile. $CCExposure^{Initial}$ is a measure of climate change exposure based on the initial seed bigrams only. Panel A reports results for the overall climate change exposure measure, and Panel B for the topic-based measures. In the figure, the exposure measures are calculated at the quarterly (transcript) level.

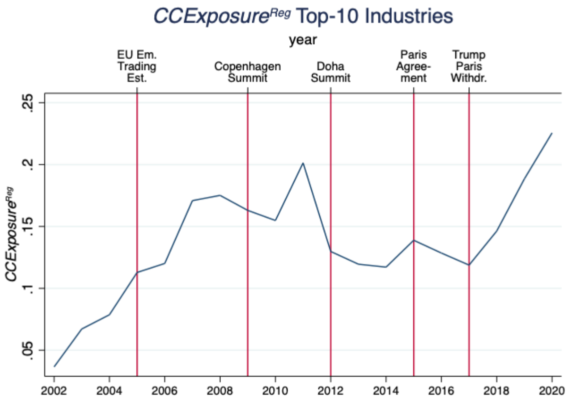
Panel A. *CCExposure*



Panel B. $CCExposure^{Opp}$



Panel C. $CCExposure^{Reg}$



Panel D. $CCExposure^{Phy}$

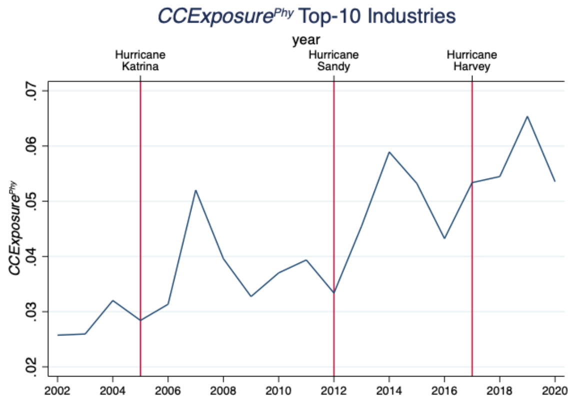


Figure 3. Climate change exposure over time. This figure shows firms' average climate change exposures over time. *CCExposure* measures the relative frequency with which climate change bigrams occur in earnings calls. $CCExposure^{Opp}$ measures the relative frequency with which bigrams that capture opportunities related to climate change occur in earnings calls. $CCExposure^{Reg}$ measures the relative frequency with which bigrams that capture regulatory shocks related to climate change occur in earnings calls. $CCExposure^{Phy}$ measures the relative frequency with which bigrams that capture physical shocks related to climate change occur in earnings calls. For each exposure measure, we construct the time series for firms in the top-10 industries (see Table III). The Appendix provides detailed variable definitions.

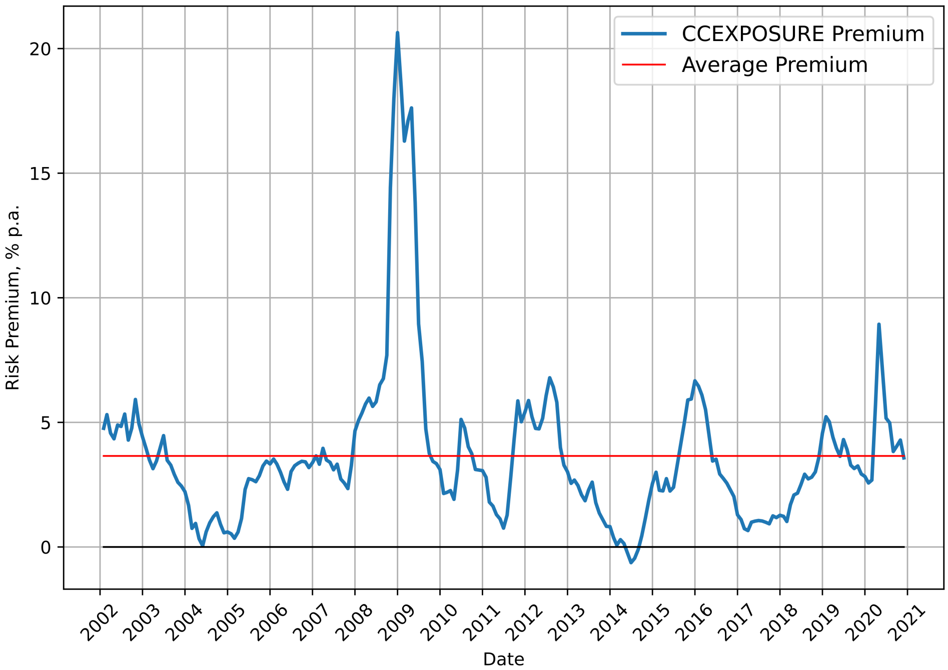


Figure 4. Risk premium on the climate change exposure factor. This figure shows the time series of the risk premium on the *CCEXPOSURE* factor, estimated together with the four-factor Carhart (1997) model using the conditional framework of Gagliardini, Ossola, and Scaillet (2016). The factor is constructed as the monthly change in the cross-sectional average of *CCExposure* across U.S.-traded sample firms. The factor is standardized to have zero mean and an annual volatility of 10%.

Table I
Climate Change Exposure Variables: Summary Statistics

This table reports summary statistics for different measures of climate change exposure, carbon emissions, and public attention to climate change at the firm-year level. For the climate change exposure measures, we average values of the four earnings calls during the year. The sample includes 10,673 unique firms from 34 countries over the period 2002 to 2020. The Appendix provides detailed variable definitions.

	Mean	STD	25%	Median	75%	N
CC Measures ($\times 10^3$)						
$CCExposure_{i,t}$	1.01	2.53	0.10	0.30	0.78	86,152
$CCExposure_{i,t}^{Opp}$	0.31	1.23	0.00	0.00	0.15	86,152
$CCExposure_{i,t}^{Reg}$	0.04	0.23	0.00	0.00	0.00	86,152
$CCExposure_{i,t}^{Phy}$	0.01	0.11	0.00	0.00	0.00	86,152
CC Measures (TFIDF Version) ($\times 10^3$)						
$CCExposure_{i,t}$	7.99	19.69	0.77	2.44	6.26	86,152
$CCExposure_{i,t}^{Opp}$	2.35	9.08	0.00	0.00	1.18	86,152
$CCExposure_{i,t}^{Reg}$	0.32	1.68	0.00	0.00	0.00	86,152
$CCExposure_{i,t}^{Phy}$	0.10	0.81	0.00	0.00	0.00	86,152
CC Q&A Measure ($\times 10^3$)						
$CCExposure_{i,t}^{Q\&A}$	0.67	1.95	0.00	0.12	0.54	86,152
CC Sentiment and Risk Measures ($\times 10^3$)						
$CCSentiment_{i,t}^{Pos}$	0.38	1.10	0.00	0.07	0.32	86,152
$CCSentiment_{i,t}^{Neg}$	0.19	0.55	0.00	0.00	0.16	86,152
$CCRisk_{i,t}$	0.04	0.17	0.00	0.00	0.00	86,152
Carbon Emissions and Climate Change Attention						
$Total\ Emissions_{i,t}$	2,961,549	13,608,989	27,472	133,847	751,772	33,789
$WSJ\ CC\ News\ Index_t$	0.007	0.001	0.006	0.006	0.008	68,794

Table II
Top-100 Bigrams Captured by Climate Change Exposure (*CCExposure*)

This table reports the top-100 bigrams associated with *CCExposure*, which measures the relative frequency with which bigrams related to climate change occur in earnings call transcripts. The Appendix defines all variables in detail.

Bigram	Frequency	Bigram	Frequency	Bigram	Frequency
renewable energy	15,605	onshore wind	878	carbon intensity	641
electric vehicle	9,508	electric motor	869	energy application	615
clean energy	6,430	provide energy	851	produce electricity	604
new energy	4,544	efficient solution	839	help state	604
climate change	4,374	global warm	837	environmental standard	593
wind power	4,253	power generator	828	power agreement	586
wind energy	4,035	solar pv	827	supply energy	585
energy efficient	3,899	scale solar	827	electric hybrid	585
greenhouse gas	3,416	need clean	821	source power	575
solar energy	2,511	coastal area	816	sustainability goal	572
air quality	2,409	energy star	793	energy reform	571
clean air	2,301	environmental footprint	792	plant power	564
carbon emission	2,088	design use	777	compare conventional	560
gas emission	1,910	area energy	777	gas vehicle	560
extreme weather	1,773	charge station	762	effort energy	560
carbon dioxide	1,583	clean water	759	pass house	559
water resource	1,423	major design	747	carbon free	558
autonomous vehicle	1,394	vehicle manufacturer	740	driver assistance	545
energy environment	1,279	future energy	737	electrical energy	543
wind resource	1,245	motor control	726	solar installation	541
government india	1,201	combine heat	718	snow ice	538
battery power	1,147	electric bus	709	renewable natural	536
air pollution	1,127	distribute power	703	promote use	536
battery electric	1,121	environmental benefit	695	farm project	531
integrate resource	1,052	eco friendly	695	laser diode	528
clean power	1,008	electrical vehicle	695	deliver energy	526
carbon price	999	carbon neutral	690	protect environment	525
world population	977	fast charge	675	sustainable energy	523
solar farm	971	cell power	657	manage energy	522
energy regulatory	967	energy team	650	invest energy	521
obama administration	957	cycle gas	646	electric energy	519
heat power	941	coal gasification	643	forest land	512
carbon tax	928	environmental concern	643	capacity energy	512
unite nation	925				

Table III
Industry Distribution of Climate Change Exposure Measures

This table reports firms' climate change exposure measures for the top-10 industries. Statistics are reported at the firm-year level across different SIC2 industries. We rank sectors by the average values of the climate change exposure measures. $CCExposure$ measures the relative frequency with which climate change bigrams occur in earnings calls. $CCExposure^{Opp}$ measures the relative frequency with which bigrams that capture opportunities related to climate change occur in earnings calls. $CCExposure^{Reg}$ measures the relative frequency with which bigrams that capture regulatory shocks related to climate change occur in earnings calls. $CCExposure^{Phy}$ measures the relative frequency with which bigrams that capture physical shocks related to climate change occur in earnings calls. For all measures, we average values of the four earnings calls during the year. We report results only those industries for which we have more than 20 firm-year observations. The Appendix defines all variables in detail.

Panel A: $CCExposure$ ($\times 10^3$)

Industry (SIC2)	Mean	Std.Dev.	Median	N
49 Electric, Gas, & Sanitary Services	6.95	6.23	5.34	3,259
16 Heavy Construction, Except Building	3.04	4.35	1.53	537
17 Construction	2.26	2.95	1.16	131
37 Transportation Equipment	2.12	3.17	1.07	2,021
36 Electronic & Other Electric Equipment	2.07	4.20	0.57	5,812
12 Coal Mining	2.05	1.48	1.70	253
29 Petroleum Refining	1.72	2.14	1.06	730
41 Local & Suburban Transit	1.69	2.06	0.84	94
55 Automative Dealers & Service Stations	1.63	3.90	0.69	484
33 Primary Metal	1.56	1.54	1.14	1,149

Panel B: $CCExposure^{Opp}$ ($\times 10^3$)

Industry (SIC2)	Mean	Std.Dev.	Median	N
49 Electric, Gas, & Sanitary Services	2.50	3.30	1.26	3,259
16 Heavy Construction, Except Building	1.37	2.78	0.30	537
17 Construction	0.91	1.71	0.34	131
36 Electronic & Other Electric Equipment	0.90	2.38	0.09	5,812
37 Transportation Equipment	0.81	1.70	0.23	2,021
55 Automative Dealers & Service Stations	0.54	1.34	0.16	484
29 Petroleum Refining	0.47	0.93	0.16	730
35 Industrial Machinery & Equipment	0.46	1.85	0.07	4,056
75 Auto Repair, Services, & Parking	0.42	1.04	0.11	171
87 Engineering & Accounting & Research	0.38	0.94	0.00	1,443

Table III (continued)

Panel C: $CCExposure^{Reg}$ ($\times 10^3$)				
Industry (SIC2)	Mean	Std.Dev.	Median	N
49 Electric Gas & Sanitary Services	0.34	0.61	0.10	3,259
12 Coal Mining	0.14	0.24	0.00	253
29 Petroleum Refining	0.14	0.32	0.00	730
32 Stone Clay Glass Products	0.12	0.35	0.00	622
10 Metal Mining	0.08	0.32	0.00	1,465
33 Primary Metal	0.08	0.22	0.00	1,149
37 Transportation Equipment	0.08	0.27	0.00	2,021
35 Industrial Machinery & Equipment	0.08	0.47	0.00	4,056
24 Lumber & Wood	0.07	0.43	0.00	471
16 Heavy Construction	0.07	0.21	0.00	537
Panel D: $CCExposure^{Phy}$ ($\times 10^3$)				
Industry (SIC2)	Mean	Std.Dev.	Median	N
41 Local and Suburban Transit	0.17	0.47	0.00	94
26 Paper & Allied Products	0.08	0.35	0.00	852
24 Lumber & Wood	0.07	0.26	0.00	471
49 Electric, Gas, & Sanitary Services	0.06	0.24	0.00	3,259
14 Mining & Quarrying	0.05	0.14	0.00	208
12 Coal Mining	0.04	0.19	0.00	253
64 Insurance Agents, Brokers, & Service	0.03	0.15	0.00	297
10 Metal Mining	0.03	0.12	0.00	1,465
15 Building Construction	0.03	0.09	0.00	600
35 Industrial Machinery & Equipment	0.03	0.25	0.00	4,056

Table IV
Climate Change Exposure Measures: Effects of Carbon Emissions and Climate Change News

This table reports regressions that relate carbon emissions and climate change news to the climate change exposure measures. Regressions are estimated at the firm-year level. $CCExposure$ measures the relative frequency with which bigrams related to climate change occur in earnings calls. $CCExposure^{Opp}$ measures the relative frequency with which bigrams that capture opportunities related to climate change occur in earnings call transcripts. $CCExposure^{Reg}$ measures the relative frequency with which bigrams that capture regulatory shocks related to climate change occur in earnings calls. $CCExposure^{Phy}$ measures the relative frequency with which bigrams that capture physical shocks related to climate change occur in earnings calls. For all measure, we average values of the four earnings calls during the year. $Total\ Emissions$ is the sum of a firm's Scope 1 and Scope 2 carbon emissions. $WSJ\ CC\ News\ Index$ is a time-series index developed in Engle et al. (2020) that captures climate change news in the *Wall Street Journal*. We divide the coefficient on *WSJ Climate Change News Index* by 100. The regressions control for *Log(Assets)*, *Debt/Assets*, *Cash/Assets*, *PP&E/Assets*, *EBIT/Assets*, *CAPEX/Assets*, and *R&D/Assets* (all in $t-1$). In Panel B, we do not include time-varying industry fixed effects, as *WSJ CC News Index* varies only in the time series. Standard errors, clustered at the industry-year level, are in parentheses. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Carbon Emissions

	$CCExposure_{i,t}$	$CCExposure^{Opp}_{i,t}$	$CCExposure^{Reg}_{i,t}$	$CCExposure^{Phy}_{i,t}$
	(1)	(2)	(3)	(4)
$Log(1 + Total\ Emissions_{i,t-1})$	0.169***	0.036***	0.023***	-0.000
	(0.023)	(0.009)	(0.003)	(0.001)
Model	OLS	OLS	OLS	OLS
Sample	All	All	All	All
Controls	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes
Industry Fixed Effects	No	No	No	No
Country Fixed Effects	Yes	Yes	Yes	Yes
N	30,905	30,905	30,905	30,905
Adj. R^2	0.390	0.267	0.145	0.035

Panel B: Public Attention to Climate Change

	$CCExposure_{i,t}$	$CCExposure^{Opp}_{i,t}$	$CCExposure^{Reg}_{i,t}$	$CCExposure^{Phy}_{i,t}$
	(1)	(2)	(3)	(4)
$WSJ\ CC\ News\ Index_t$	0.427**	0.154*	0.034***	0.002
	(0.168)	(0.089)	(0.010)	(0.004)
Model	OLS	OLS	OLS	OLS
Sample	All	All	All	All
Controls	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	No	No	No	No
Industry Fixed Effects	Yes	Yes	Yes	Yes
Country Fixed Effects	Yes	Yes	Yes	Yes
N	54,824	54,824	54,824	54,824
Adj. R^2	0.298	0.185	0.090	0.024

Table V
Variance Decomposition of Climate Change Exposure Measures

This table provides a variance decomposition of the climate change exposure measures. Regressions are estimated at the firm-year 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. $CCExposure$ measures the relative frequency with which climate change bigrams occur in earnings calls. $CCExposure^{Opp}$ measures the relative frequency with which bigrams that capture opportunities related to climate change occur in earnings calls. $CCExposure^{Reg}$ measures the relative frequency with which bigrams that capture regulatory shocks related to climate change occur in earnings calls. $CCExposure^{Phy}$ measures the relative frequency with which bigrams that capture physical shocks related to climate change occur in earnings calls. For all measures, we average values of the four earnings calls during the year. The Appendix defines all variables in detail.

	$CCExposure_{i,t}$	$CCExposure^{Opp}_{i,t}$	$CCExposure^{Reg}_{i,t}$	$CCExposure^{Phy}_{i,t}$
	(1)	(2)	(3)	(4)
Panel A: Incremental R^2				
Year Fixed Effect	0.7%	0.7%	0.5%	0.05%
Industry Fixed Effect	27.1%	16.9%	7.8%	2.0%
Industry x Year Fixed Effect	1.9%	2.6%	1.4%	1.5%
Country Fixed Effect	0.6%	0.7%	0.4%	0.3%
"Firm Level"	69.7%	79.1%	89.9%	96.2%
Sum	100.0%	100.0%	100.0%	100.0%
Panel B: Fraction of Variation				
Firm Fixed Effect: Permanent differences across firms within sector and countries	51.6%	56.4%	44.7%	45.1%
Residual: Variation over time in the identity of firms within industries and countries most affected by exposure variable	48.4%	43.7%	55.3%	54.9%
Sum	100.0%	100.0%	100.0%	100.0%

Table VI
Quantifying Measurement Error in Climate Change Exposure Measures

This table shows AR(1) regressions of climate change exposure. Regressions are estimated at the firm-year level. *CCExposure* measures the relative frequency with which climate change bigrams occur in earnings calls. We average values of the four earnings calls during the year. $CCExposure^{10K}$ measures climate change exposure by applying our algorithm to the "Management Discussion and Analysis" (MD&A) section in firms' annual 10K filings. Following Hassan et al. (2019), *CCExposure* and $CCExposure^{10K}$ in this table are standardized by subtracting the sample mean and dividing by the sample standard deviation. *Implied Share Measurement Error* is calculated as $1 - (\hat{\beta}_{OLS}/\hat{\beta}_{IV})$, where $\hat{\beta}_{OLS}$ is the estimated coefficient in $CCExposure_{i,t} = \alpha + \beta CCExposure_{i,t-1} + \epsilon$ and $\hat{\beta}_{IV}$ is the coefficient on the instrumented $CCExposure_{i,t}$ in the same specification. To obtain bootstrapped standard errors for *Implied Share Measurement Error*, we repeat the following procedure 500 times: draw a random sample of the same sample size (with replacement and clustered by firm) from our regression sample, run the two regressions, and obtain the implied share of measurement error. These standard errors are clustered at the firm level. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Overall Variation

	$CCExposure_{i,t}$			
	(1)	(2)	(3)	(4)
$CCExposure_{i,t-1}$	0.922***	1.008***	0.991***	0.958***
	(0.002)	(0.003)	(0.003)	(0.002)
Model	OLS	IV	IV	IV
Instrument		$CCExposure^{10K}_{i,t-1}$	$CCExposure^{10K}_{i,t-2}$	$CCExposure_{i,t-2}$
Sample	US	US	US	US
Industry x Year Fixed Effects	No	No	No	No
N	47,589	47,589	41,794	41,794
Implied Share Measurement Error		0.085	0.069	0.037
		(0.007)	(0.007)	(0.005)

Panel B: Firm-Level Variation

	$CCExposure_{i,t}$			
	(1)	(2)	(3)	(4)
$CCExposure_{i,t-1}$	0.886***	0.992***	0.966***	0.932***
	(0.002)	(0.004)	(0.002)	(0.003)
Model	OLS	IV	IV	IV
Instrument		$CCExposure^{10K}_{i,t-1}$	$CCExposure^{10K}_{i,t-2}$	$CCExposure_{i,t-2}$
Sample	US	US	US	US
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes
N	47,502	47,502	41,712	41,712
Implied Share Measurement Error		0.107	0.083	0.050
		(0.002)	(0.012)	(0.007)

Table VII
Green-Tech Jobs and Climate Change Exposure Measures

This table reports regressions that relate green-tech jobs to the climate change exposure measures. Regressions are estimated at the firm-year level. *#Green-Tech Jobs* is the number of job postings for disruptive green-tech jobs. *I(Green-Tech Jobs)* is an indicator that equals one if *#Green-Tech Jobs* is positive, and zero otherwise. *#Nongreen-Tech Jobs* is the number of job postings for nongreen disruptive tech jobs. $Green\text{-}Tech\ Ratio_{i,t+1}$ is the number of job postings for disruptive green jobs relative to the total number of all disruptive job postings. $CCExposure$, $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$ are defined as in previous tables. The regressions control for *Log(Assets)*, *Debt/Assets*, *Cash/Assets*, *PP&E/Assets*, *EBIT/Assets*, *CAPEX/Assets*, and *R&D/Assets* (all in t). In columns (5) to (7), we use the same observations as in columns (1) to (4). In columns (1) to (4) and (7), the economic effects are computed as the percentage change in the dependent variable for a one-standard-deviation change in the exposure variable of interest. In columns (5) and (6), the economic effect is computed as the effect of a one-standard-deviation change in the exposure variable relative to the standard deviation of the dependent variable. We use the within-fixed-effect standard deviations. Standard errors, clustered at the industry-year level, are in parentheses. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\text{-}Tech\ Jobs_{i,t+1}$				$I(\#Green\text{-}Tech\ Jobs)_{i,t+1}$	$Green\text{-}Tech\ Ratio_{i,t+1}$	$\#Non\text{-}green\text{-}Tech\ Jobs_{i,t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Log(1 + CCExposure_{i,t})$	1.564***				0.077***	0.015***	-0.204***
	(0.199)				(0.006)	(0.003)	(0.060)
$Log(1 + CCExposure^{Opp}_{i,t})$		1.833***					
		(0.229)					
$Log(1 + CCExposure^{Reg}_{i,t})$			1.458***				
			(0.445)				
$Log(1 + CCExposure^{Phy}_{i,t})$				1.079			
				(1.217)			
Model	Poisson	Poisson	Poisson	Poisson	OLS	OLS	Poisson
Sample	US	US	US	US	US	US	US
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	23,870	23,870	23,870	23,870	23,870	23,870	23,870
Adj./ps. R^2	0.754	0.767	0.687	0.684	0.116	0.049	0.526
Dep. Variable: Mean	2.82	2.82	2.82	2.82	0.07	0.003	845.09
Dep. Variable: STD	89.56	89.56	89.56	89.56	0.26	0.042	3613.42
Economic Effect, %	108.7	79.5	20.0	6.8	14.0	16.9	-9.1

Table VIII
Green Patents and Climate Change Exposure Measures

This table reports regressions that relate green patents to the climate change exposure measures. Regressions are estimated at the firm-year level. *#Green Patents* is the number of green patents. *I(Green Patents)* is an indicator that equals one if *#Green Patents* is positive, and zero otherwise. *Green Patents* $Ratio_{i,t+1}$ is the number of green patents relative to the total number of all patents. *#Nongreen Patents* is the number of nongreen patents. $CCExposure$, $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$ are defined as in previous tables. The regressions control for *Log(Assets)*, *Debt/Assets*, *Cash/Assets*, *PP&E/Assets*, *EBIT/Assets*, *CAPEX/Assets*, and *R&D/Assets* (all in t). In columns (5) to (7), we use the same observations as in columns (1) to (4) (the Poisson estimation in Column (7) drops some observations). In columns (1) to (4) and (7), the economic effects are computed as the percentage change in the dependent variable for a one-standard-deviation change in the exposure variable of interest. In columns (5) and (6), the economic effect is computed as the effect of a one-standard-deviation change in the exposure variable relative to the standard deviation of the dependent variable. We use the within-fixed-effect standard deviations. Standard errors, clustered at the industry-year level, are in parentheses. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\ Patents_{i,t+1}$				$I(Green\ Patents)_{i,t+1}$	$Green\ Patents\ Ratio_{i,t+1}$	$\#Non\text{-}green\ Patents_{i,t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$Log(1 + CCExposure_{i,t})$	1.102***				0.025***	0.006***	-0.436***
	(0.231)				(0.003)	(0.001)	(0.118)
$Log(1 + CCExposure^{Opp}_{i,t})$		0.854***					
		(0.312)					
$Log(1 + CCExposure^{Reg}_{i,t})$			3.061***				
			(0.272)				
$Log(1 + CCExposure^{Phy}_{i,t})$				-1.155			
				(2.865)			
Model	Poisson	Poisson	Poisson	Poisson	OLS	OLS	Poisson
Sample	US	US	US	US	US	US	US
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	21,914	21,914	21,914	21,914	21,914	21,914	21,776
Adj./ps. R^2	0.617	0.603	0.614	0.598	0.078	0.023	0.752
Dep. Variable: Mean	0.28	0.28	0.28	0.28	0.03	0.003	22.10
Dep. Variable: STD	4.07	4.07	4.07	4.07	0.18	0.040	224.23
Economic Effect, %	71.7	32.0	47.3	-6.9	7.0	7.4	-19.3

Table IX
Forward-Looking Risk Measures and Climate Change Exposure Measures

This table reports regressions that relate forward-looking risk measures to the climate change exposure measures. Regressions are estimated at the firm-quarter level. $IVar$ is implied variance, $ISkew$ is implied skewness, $IKurt$ is implied kurtosis, $SlopeD$ and $SlopeU$ are implied volatility slopes on the left and right of the distribution, and VRP is the variance risk premium. Construction of the option-implied measures is detailed in Section II of the Internet Appendix. $CCExposure$, $CCExposure^{Opp}$, $CCExposure^{Reg}$, and $CCExposure^{Phy}$ are defined as in previous tables. The regressions control for *Log(Assets)*, *Debt/Assets*, *Cash/Assets*, *PP&E/Assets*, *EBIT/Assets*, *CAPEX/Assets*, and *R&D/Assets* (all in t). The economic effect is computed as the effect of a one-standard-deviation change in the exposure variable of interest relative to the standard deviation of the dependent variable (in %). We use the within-fixed-effect standard deviation. Standard errors, clustered at the industry-year level, are in parentheses. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$IVar_{i,t+1}$	$ISkew_{i,t+1}$	$IKurt_{i,t+1}$	$SlopeD_{i,t+1}$	$SlopeU_{i,t+1}$	$VRP_{i,t+1}$
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: $CCExposure$						
$Log(1+CCExposure_{i,t})$	-0.002	-0.049***	0.303***	0.033***	-0.026***	0.003
	(0.005)	(0.009)	(0.049)	(0.007)	(0.006)	(0.002)
N	42,093	42,093	42,093	42,093	42,093	42,089
Adj. R^2	0.424	0.140	0.349	0.231	0.236	0.094
Economic Effect, %	-0.42	-4.57	7.01	4.46	-4.14	0.89
Panel B: $CCExposure^{Opp}$						
$Log(1+CCExposure^{Opp}_{i,t})$	0.004	-0.053***	0.403***	0.048***	-0.037***	0.006*
	(0.009)	(0.012)	(0.067)	(0.011)	(0.010)	(0.003)
N	42,093	42,093	42,093	42,093	42,093	42,089
Adj. R^2	0.424	0.140	0.348	0.231	0.236	0.094
Economic Effect, %	0.56	-3.27	6.18	4.30	-3.91	1.19
Panel C: $CCExposure^{Reg}$						
$Log(1+CCExposure^{Reg}_{i,t})$	-0.007	-0.075***	0.453***	0.054**	-0.053***	0.005
	(0.014)	(0.024)	(0.146)	(0.027)	(0.019)	(0.008)
N	42,093	42,093	42,093	42,093	42,093	42,089
Adj. R^2	0.424	0.139	0.346	0.230	0.235	0.094
Economic Effect, %	-0.46	-2.19	3.28	2.28	-2.64	0.47
Panel D: $CCExposure^{Phy}$						
$Log(1+CCExposure^{Phy}_{i,t})$	-0.033	-0.083	1.336***	0.145***	-0.175***	-0.012
	(0.020)	(0.059)	(0.319)	(0.048)	(0.048)	(0.011)
N	42,093	42,093	42,093	42,093	42,093	42,089
Adj. R^2	0.424	0.139	0.347	0.230	0.236	0.094
Economic Effect, %	-1.02	-1.13	4.51	2.85	-4.06	-0.52
Model	OLS	OLS	OLS	OLS	OLS	OLS
Sample	S&P500	S&P500	S&P500	S&P500	S&P500	S&P500
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Variable: Mean	0.176	-0.571	4.678	0.317	-0.101	0.042
Dep. Variable: STD	0.199	0.453	1.823	0.312	0.265	0.142

Table X
Climate Change Exposure Factor: Components of F and ν.

This table reports the estimated annualized components of F and ν for the four-factor Carhart (1997) model augmented by a *CCEXPOSURE* factor. The estimation is based on the conditional framework by Gagliardini, Ossola, and Scaillet (2016). The factor is constructed as the monthly change in the cross-sectional average of *CCExposure* across U.S.-traded sample firms. The factor is standardized to have zero mean and an annual volatility of 10%. All instruments are centered and standardized in the time series. The common instruments are the default spread and the term spread, and the firm-specific instrument is the log of the book-to-market ratio. *p<0.1; **p<0.05; ***p<0.01.

Factors	*Instruments*	F	$SE(F)$	ν	$SE(\nu)$
		(1)	(2)	(3)	(4)
Market	Constant	8.9838***	3.4981	2.3908***	0.7110
	Default Spread	-1.0201	5.4550	2.4676***	0.8715
	Term Spread	-1.9715	3.3962	1.4489**	0.6705
SMB	Constant	2.3669	1.9164	2.6523*	1.3459
	Default Spread	2.5406	2.0404	-1.3983	1.0227
	Term Spread	2.1356	1.8985	-4.6391***	0.9302
HML	Constant	-2.1553	2.0893	-3.5959***	1.0965
	Default Spread	-3.6834	3.9437	3.7360***	0.8545
	Term Spread	4.8748**	2.2504	-0.0444	0.8434
MOM	Constant	1.3199	3.5668	7.2011***	1.6444
	Default Spread	-14.359*	8.2567	7.8356***	1.7552
	Term Spread	2.4766	2.9728	-0.6825	1.2843
CCEXPOSURE	Constant	-0.0032	2.3008	3.7273***	1.1654
	Default Spread	0.0805	2.7644	3.1262***	1.0855
	Term Spread	-0.2941	2.6282	-0.1834	0.9978

Internet Appendix

for

"Firm-Level Climate Change Exposure"

ZACHARIAS SAUTNER, LAURENCE VAN LENT, GRIGORY VILKOV,
and RUISHEN ZHANG*

This Internet Appendix provides additional tables and figures supporting the main text. Section I presents the climate change bigrams search algorithm. Section II explains the construction of option-implied measures. Section III provides additional tables.

*Citation format: Zacharias Sautner, Laurence van Lent, Grigory Vilkov, and Ruishen Zhang, Internet Appendix for "Firm-Level Climate Change Exposure," *Journal of Finance* [DOI STRING]. Please note: Wiley is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the authors of the article.

I. Climate Change Bigrams Search Algorithm

We create $\mathbb{C}$ from the union of two separate sets of bigrams: i) a set containing 50 very general and ex-ante specified climate change bigrams, and ii) a set created with machine learning algorithms that construct bigrams directly from analyst conference call transcripts.

Defining the search set. To enable an algorithm to self-discover climate change bigrams from conference call transcripts, we start by compiling a set of conference call transcripts that potentially discuss climate change topics. As a "rough" climate-change training library $\mathbb{C}^R$, we use climate change bigrams in a comprehensive set (288 MB) of research reports issued by the Intergovernmental Panel on Climate Change (IPCC). We lemmatize and stem the textual IPCC data, removing digits, punctuation, and stop words, and we drop bigrams with a text frequency lower than 10.

We also construct a nonclimate-change training library $\mathbb{N}$, which consists of English-language novels taken from Project Gutenberg; news articles on technology, business, and politics from BBC and Thomas Reuters; IMF research reports; and accounting and econometrics textbooks. We then apply the method in Hassan et al. (2019) and compute a "rough" climate change exposure score for each transcript as follows:

$$RoughCCExposure_{it} = \frac{1}{B_{it}} \sum_{b}^{B_{it}} \left(1[b \in \mathbb{C}^{\mathbb{R}} \setminus \mathbb{N}]\right), \tag{IA.1}$$

Although the nonclimate-change training library $\mathbb{N}$ includes extensive sources of textual data, we find that the set of bigrams $\mathbb{C}^R \setminus \mathbb{N}$ is still contaminated by a considerable number of nonclimate change bigrams. The reason is that many climate change bigrams often inherently relate to a broad domain of other topics that conference call participants are likely to discuss in contexts unrelated to climate change, such as economic growth, commercial feasibility, and technology development. Moreover, conference call participants tend to view climate change from different perspectives compared to the scientists that write the IPCC reports.

To address these problems, we construct a new set $\mathbb{M}$, which consists of sentences in transcripts with positive "rough" climate change bigrams (i.e., those reports in which bigrams $\mathbb{C}^R \setminus \mathbb{N}$ occurred). The goal of constructing this new set is to find the sentences that actually discuss climate change topics and then to extract climate change bigrams from these sentences.

Defining the reference set. We next partition $\mathbb{M}$ into reference and search sets. To do so, we define a set of 50 very general climate change bigrams, $\mathbb{C}^0$, which includes terms such as "climate change," "global warming," and "carbon emission." We then partition $\mathbb{M}$ based on these initial bigrams into the reference set $\mathbb{R}$ (6.8 MB), which contains about 60,000 sentences containing bigrams in $\mathbb{C}^0$, and the search set $\mathbb{S}$ (3.56 GB), which contains about 70 million sentences not containing any bigrams in $\mathbb{C}^0$. The key difference between the two sets is that the reference set contains sentences almost certainly related to discussions of climate change. In contrast, the search set may mention climate change topics not captured by the bigrams specified in $\mathbb{C}^0$, but it may also contain pure noise.

Partitioning the search set. To partition the search set, we construct a training set consisting of the reference set $\mathbb{R}$ and a random sample of the search set $\mathbb{S}$ (100,000 sentences). Next, we fit three machine-learning classifiers—Multinomial Naive Bayes, Support Vector Classification, and Random Forest—to the training set. These classifiers use the content of each sentence to predict whether a sentence belongs to $\mathbb{R}$. For each classifier, we use grid-search cross-validation to select hyper-parameters that optimizes their performance. We then use the optimized parameters from each classifier to fit the search set and estimate for each sentence in $\mathbb{S}$ the predicted probability of belonging to $\mathbb{R}$. Once we have these predicted probabilities, we group sentences into a target set $\mathbb{T}$ if any of the three classifiers that we use predicts a probability of $\mathbb{R}$ membership that is higher than 0.8 for that sentence. The resulting target set contains about 700,000 sentences that do not contain any "obvious" climate change bigrams but are likely to mention climate change contents not captured by $\mathbb{C}^0$.

Finding climate change bigrams. In a last step, we identify bigrams that best discriminate the target set $\mathbb{T}$ from the nontarget set $\mathbb{S} \setminus \mathbb{T}$. We first mine all bigrams $\mathbb{T}$ and $\mathbb{S} \setminus \mathbb{T}$. We find that about 3,800 bigrams appears only in $\mathbb{T}$ and not $\mathbb{S} \setminus \mathbb{T}$. We call this set of bigrams $\mathbb{C}^S$.

For the bigrams that appear in both $\mathbb{T}$ and $\mathbb{S} \setminus \mathbb{T}$, we calculate the document frequencies of each bigram in each of the two sets and keep those bigrams that appear more frequently in the target set than in the nontarget set. For example, if a bigram appears in two out of 10 $\mathbb{T}$ sentences and in 10 out of 100 $\mathbb{S} \setminus \mathbb{T}$ sentences, this bigram appear more frequent in $\mathbb{T}$ (frequency of 0.2 versus 0.1). We then rank the bigrams that we keep based on how well they discriminate the two sets. Specifically, we compute a modified version of the likelihood metric suggested in King, Lam, and Roberts (2017) for each bigram and then add the bigrams with a top 5% likelihood into set $\mathbb{C}^S$ (about 5,000 bigrams). We use a log-gamma function instead of a gamma function because the size of the search set is so large that the gamma function cannot return a numeric value. The 5% threshold significantly reduces false positives.

Creating a final climate change bigrams library. We define the final climate change bigrams library $\mathbb{C}$ as $\mathbb{C} = \mathbb{C}^0 \cup \mathbb{C}^S$. The benefit of our approach is that the algorithms generate various meaningful climate change bigrams based on the initial bigram set $\mathbb{C}^0$.

II. Construction of Option-Implied Measures

A. Data

Data on option-implied variables come from the Volatility Surface File of Ivy DB OptionMetrics. These tests focus on S&P500 firms, for which data on liquid options are available. We match options data through the historical CUSIP link of OptionMetrics. The high frequency of the option-implied measures allows us to use quarterly values of $CCExposure$. To prepare the Volatility Surface, we select out-of-the-money (OTM) options with absolute deltas strictly (weakly) smaller than 0.5 for puts (calls) for maturities of 30 days. We interpolate the implied volatilities available as a function of moneyness between the available moneyness points. We then extrapolate the data by filling in the missing extreme data using the implied volatility values from the left and right boundaries. This method enables us to fill in the moneyness range of [1/3,3] with a total of 1,001 points. For the interpolations, we use a piece-wise cubic Hermite interpolating polynomial.

B. Measures

Implied variance, skewness, kurtosis. To measure implied variance (IV) of log returns, we take the Bakshi, Kapadia, and Madan (2003) variance swap rate $IVar_{t,t+\Delta t}$ for a given maturity $t + \Delta t$, constructed from the prices of OTM calls $C(t, t + \Delta, K)$ and puts $P(t, t + \Delta, K)$ with strike prices K observed at t:

$$IVar_{t,t+\Delta t} = R_{t,t+\Delta t} \int_0^{S_t} \frac{2(1 - \ln K/S_t)}{K^2} P(t, t + \Delta, K)dK + R_{t,t+\Delta t} \int_{S_t}^{\infty} \frac{2(1 - \ln K/S_t)}{K^2} C(t, t + \Delta, K)dK, \tag{IA.2}$$

where $R_{t,t+\Delta t}$ is the gross risk-free rate of return and S_t is the spot price of the underlying stock. We use a similar approach for implied skewness, $ISkew$, and for implied kurtosis, $IKurt$, applying the formulas for the log returns provided in Bakshi, Kapadia, and Madan (2003). We approximate each integral in equation (IA.2) for IV using a finite sum of 1,001 option prices (we do likewise for integrals in the formulas for $ISkew$ and $IKurt$).

Implied volatility slope. We measure the steepness of the implied volatility slope on the left (*SlopeD*) and right (*SlopeU*) from the at-the-money (ATM) point. As in Kelly, Pastor, and Veronesi (2016), the measures are the slopes of functions relating implied volatilities of OTM options to their deltas. We estimate *SlopeD* by regressing implied volatilities of puts with deltas between −0.1 and −0.5 on their deltas (and a constant). For *SlopeU*, we regress implied volatilities of calls with deltas between 0.1 and 0.5 on their deltas. Note that for *SlopeD*, the independent variable (delta) is *increasing* for more OTM options, so a positive (and higher) slope coefficient indicates more expensive tail protection, while for *SlopeU* the independent variable is *decreasing* for more OTM options, and hence a more negative slope coefficient indicates higher cost of obtaining right-tail exposure. The variable *SlopeD* is on average positive and *SlopeU* is on average negative as far-OTM options are typically more expensive (in terms of implied volatilities) than ATM options.

Variance risk premium. We calculate risk premiums for particular risks by comparing expected quantities under the physical and risk-neutral probability measures as follows. (The theoretically sound definition of the finite-period risk premium is the expectation under the risk-neutral (Q) measure minus expectation under the physical (P) measure; for convenience, we follow an informal tradition of computing the finite-period risk premium as the Q minus P expectation.) The variance risk premium (VRP) allows us to evaluate the cost of protection against general variance risk (or uncertainty, as suggested in Bali and Zhou (2016)). We compute VRP as the difference between the risk-neutral expected and the past realized variances (the latter acting as a proxy for expected variance under the physical measure):

$$VRP_{t,t+\Delta t} = IVar_{t,t+\Delta t} - RVar_{t-\Delta t,t}, \tag{IA.3}$$

where $RVar_{t-\Delta t,t}$ is computed from daily simple returns over the rolling window $[t - \Delta t, t]$.

C. Costs and Benefits of Measures

While these "risk quantities" do not directly reflect expectations of risk in the real (physical) world, they efficiently aggregate the forward-looking consensus of market participants with respect to the future return distribution. A key benefit is their forward-looking character. For example, $IVar$ is a strong predictor of the future realized variance (Poon and Granger (2003)), $ISkew$ allows for the quantification of the asymmetry of the risk-neutral distribution, and $SlopeD/SlopeU$ represents a heuristic proxy for the relative price of protection against tail risk (Kelly, Pastor, and Veronesi (2016)). A cost is potential bias stemming from the risk premium effect (see Vanden (2008), Chang et al. (2012), Cremers, Halling, and Weinbaum (2015), and DeMiguel et al. (2013)).

III. Additional Tables

Table IA. I
Firm-Years Across Countries

This table reports the distribution of firm-year observations across countries.

Country/Region	N	Percent
Australia	1,460	1.69
Austria	193	0.22
Belgium	262	0.3
Bermuda	727	0.84
Brazil	1,049	1.22
Canada	5,924	6.88
Chile	227	0.26
China	1,459	1.69
Denmark	428	0.5
Finland	472	0.55
France	1,314	1.53
Germany	1,320	1.53
Greece	234	0.27
Hong Kong	450	0.52
India	1,227	1.42
Ireland; Republic of	646	0.75
Israel	738	0.86
Italy	553	0.64
Japan	1,675	1.94
Korea; Republic (S. Korea)	296	0.34
Luxembourg	271	0.31
Mexico	542	0.63
Netherlands	798	0.93
New Zealand	206	0.24
Norway	450	0.52
Russia	335	0.39
Singapore	256	0.3
South Africa	480	0.56
Spain	504	0.59
Sweden	930	1.08
Switzerland	975	1.13
Taiwan	344	0.4
United Kingdom	3,300	3.83
United States of America	56,107	65.13
Total	86,152	100

Table IA. II
Firm-Years with/without Trucost Emissions Data

This table reports summary statistics of climate change exposure measures and firm characteristics depending on whether a firm-year is included in the Trucost database on carbon emissions. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	Firm-Year Observations with Trucost Emissions Data (N=33,789)			Firm-Year Observations without Trucost Emissions Data (N=52,363)			
	Mean	Std.Dev.	Median	Mean	Std.Dev.	Median	Difference-in-Means
$CCExposure_{i,t}$	1.190	2.831	0.350	0.899	2.312	0.277	0.291***
$CCExposure_{i,t}^{Opp}$	0.379	1.362	0.000	0.265	1.129	0.000	0.114***
$CCExposure_{i,t}^{Reg}$	0.054	0.256	0.000	0.038	0.218	0.000	0.016***
$CCExposure_{i,t}^{Phy}$	0.016	0.128	0.000	0.012	0.095	0.000	0.004***
$Assets_{i,t}$	23616	57774	4798	4976	22603	707	18640***
$Debt/Assets_{i,t}$	0.260	0.193	0.239	0.252	0.241	0.204	0.008***
$Cash/Assets_{i,t}$	0.139	0.160	0.083	0.203	0.225	0.110	-0.064***
$PPE/Assets_{i,t}$	0.265	0.242	0.189	0.236	0.247	0.136	0.028***
$EBIT/Assets_{i,t}$	0.069	0.118	0.070	0.003	0.203	0.049	0.066***
$CAPEX/Assets_{i,t}$	0.044	0.046	0.031	0.042	0.053	0.024	0.002***
$R\&D/Assets_{i,t}$	0.026	0.064	0.000	0.051	0.106	0.000	-0.025***
$USfirm_{i,t}$	0.488	0.500	0.000	0.756	0.429	1.000	-0.268***

Table IA. III
Initial Bigrams for Searching Climate Change Bigrams

air pollution	electric vehicle	new energy
air quality	energy climate	ozone layer
air temperature	energy conversion	renewable energy
biomass energy	energy efficient	sea level
carbon dioxide	energy environment	sea water
carbon emission	environmental sustainability	snow ice
carbon energy	exterme weather	solar energy
carbon neutral	flue gas	solar thermal
carbon price	forest land	sustainable energy
carbon sink	gas emission	water resource
carbon tax	ghg emission	water resources
clean air	global decarbonization	wave energy
clean energy	global warm	weather climate
clean water	greenhouse gas	wind energy
climate change	heat power	wind power
coastal area	Kyoto protocol	wind resource
costal region	natural hazard	

Table IA. IV
Initial Bigrams for Opportunity, Regulatory, and Physical Climate Change Exposure Measures

Panel A: Initial Opportunity Bigrams				
heat power	new energy	plug hybrid	rooftop solar	renewable electricity
renewable energy	wind power	renewable resource	sustainable energy	wave power
electric vehicle	wind energy	solar farm	hybrid car	geothermal power
clean energy	solar energy	electric hybrid		
Panel B: Initial Regulatory Bigrams				
greenhouse gas	gas emission	carbon tax	emission trade	carbon reduction
reduce emission	air pollution	carbon price	dioxide emission	carbon market
carbon emission	reduce carbon	environmental standard	epa regulation	mercury emission
carbon dioxide	energy regulatory	nox emission	energy independence	
Panel C: Initial Physical Bigrams				
coastal area	forest land	storm water	natural hazard	water discharge
global warm	sea level	heavy snow	sea water	ice product
snow ice	nickel metal	air water	warm climate	

Table IA. V

Climate Change Exposure Measures: Correlations

This table shows correlations across different climate change exposure measures. We report correlations for frequency-unweighted ("EW") and TFIDF-adjusted ("TFIDF") versions of climate change exposure. The Appendix defines all variables in detail.

		Frequency-Unweighted Measures (EW Measures)				TFIDF-Adjusted Measures (TFIDF Measures)			
		$CCExpo_{i,t}$	$CCExpo^{Opp}_{i,t}$	$CCExpo^{Reg}_{i,t}$	$CCExpo^{Phy}_{i,t}$	$CCExpo_{i,t}$	$CCExpo^{Opp}_{i,t}$	$CCExpo^{Reg}_{i,t}$	$CCExpo^{Phy}_{i,t}$
EW Measures	$CCExposure_{i,t}$	1.000							
	$CCExposure^{Opp}_{i,t}$	0.897	1.000						
	$CCExposure^{Reg}_{i,t}$	0.523	0.301	1.000					
	$CCExposure^{Phy}_{i,t}$	0.224	0.157	0.092	1.000				
TFIDF Measures	$CCExposure_{i,t}$	0.997	0.882	0.521	0.222	1.000			
	$CCExposure^{Opp}_{i,t}$	0.900	0.994	0.306	0.156	0.892	1.000		
	$CCExposure^{Reg}_{i,t}$	0.519	0.295	0.992	0.092	0.520	0.300	1.000	
	$CCExposure^{Phy}_{i,t}$	0.219	0.153	0.088	0.998	0.217	0.152	0.088	1.000

Table IA. VI
Earnings Call Observations across Countries

This table reports the distribution of earnings calls (earnings calls with $CCExposure > 0$) across sample countries. The sampling criteria are specified in Section I.A of the paper.

Country/Region	# Calls	# Calls with $CCExposure > 0$	Percentage of Calls with $CCExposure > 0$
Argentina	468	199	42.52
Australia	3,881	2,319	59.75
Austria	938	538	57.36
Belgium	1,047	548	52.34
Bermuda	2,855	1,433	50.19
Brazil	4,619	2,396	51.87
Canada	20,995	11,524	54.89
Chile	831	513	61.73
China	5,024	2,516	50.08
Denmark	1,845	879	47.64
Finland	2,024	1,068	52.77
France	3,931	2,525	64.23
Germany	5,539	3,169	57.21
Greece	987	445	45.09
Hong Kong	1,325	664	50.11
India	4,921	2,892	58.77
Ireland; Republic of	2,386	1,228	51.47
Israel	2,759	972	35.23
Italy	2,772	1,525	55.01
Japan	7,688	2,463	32.04
Korea; Republic (S. Korea)	1,304	625	47.93
Luxembourg	1,102	660	59.89
Mexico	2,301	1,225	53.24
Netherlands	2,959	1,611	54.44
New Zealand	477	274	57.44
Norway	2,088	1,116	53.45
Poland	673	372	55.27
Portugal	486	255	52.47
Russia	1,193	683	57.25
Singapore	1,086	561	51.66
South Africa	1,445	960	66.44
Spain	2,240	1,389	62.01
Sweden	4,250	2,065	48.59
Switzerland	3,197	1,759	55.02
Taiwan	1,377	531	38.56
Turkey	586	244	41.64
United Kingdom	10,116	6,109	60.39
United States of America	217,191	109,531	50.43
Total	330,906	169,786	51.31

Table IA. VII
Earnings Call Observations across Years

This table reports the distribution of earnings calls (earnings calls with $CCExposure > 0$) across sample years. The sampling criteria are specified in Section I.A of the paper.

Year	# Calls	# Calls with $CCExposure > 0$	Percentage of Calls with $CCExposure > 0$
2002	6,188	2,739	44.26
2003	11,908	5,377	45.15
2004	14,339	6,668	46.50
2005	15,431	7,391	47.90
2006	16,388	7,990	48.76
2007	17,405	8,487	48.76
2008	18,737	9,597	51.22
2009	18,247	9,439	51.73
2010	18,291	9,378	51.27
2011	18,642	9,796	52.55
2012	18,736	9,777	52.18
2013	16,737	8,606	51.42
2014	17,752	9,136	51.46
2015	17,785	9,220	51.84
2016	17,234	8,996	52.20
2017	19,580	10,107	51.62
2018	22,073	11,587	52.49
2019	22,757	12,157	53.42
2020	22,676	13,338	58.82
Total	330,906	169,786	51.31

Table IA. VIII
Earnings Call Observations across Industries

This table reports the distribution of earnings calls (earnings calls with $CCExposure > 0$) across industries. The sampling criteria are specified in Section I.A of the paper.

Industry (SIC2)	# Calls	# Calls with $CCExposure > 0$	Percentage of Calls with $CCExposure > 0$
01 Agricultural Production – Crops	371	234	63.07
07 Agricultural Services	129	38	29.46
09 Fishing, Hunting, & Trapping	27	23	85.19
10 Metal, Mining	4,891	3,403	69.58
12 Coal Mining	834	751	90.05
13 Oil & Gas Extraction	11,941	7,335	61.43
14 Nonmetallic Minerals, Except Fuels	742	577	77.76
15 General Building Contractors	2,018	1,117	55.35
16 Heavy Construction, Except Building	1,874	1,615	86.18
17 Construction	471	361	76.65
20 Food & Kindred Products	7,614	3,894	51.14
21 Tobacco Products	678	239	35.25
22 Textile Mill Products	569	245	43.06
23 Apparel & Other Textile Products	2,338	859	36.74
24 Lumber & Wood	1,735	918	52.91
25 Furniture & Fixtures	1,428	595	41.67
26 Paper & Allied Products	3,263	1,987	60.89
27 Printing & Publishing	2,643	879	33.26
28 Chemical & Allied Products	30,174	13,134	43.53
29 Petroleum Refining	3,062	2,329	76.06
30 Rubber & Miscellaneous Plastics Products	2,041	1,221	59.82
31 Leather & Leather Products	941	384	40.81
32 Stone, Clay, & Glass Products	2,058	1,494	72.59
33 Primary Metal	3,998	3,097	77.46
34 Fabricated Metal Products	2,996	1,882	62.82
35 Industrial Machinery & Equipment	15,292	9,588	62.70
36 Electronic & Other Electric Equipment	22,426	14,200	63.32
37 Transportation	7,796	6,043	77.51
38 Instruments & Related Products	15,524	7,721	49.74
39 Miscellaneous Manufacturing Industries	1,831	738	40.31
40 Railroad Transportation	723	601	83.13
41 Local & Suburban Transit	241	190	78.84
42 Trucking & Warehousing	1,599	853	53.35
44 Water Transportation	2,656	1,579	59.45
45 Transportation by Air	3,063	1,827	59.65
46 Pipelines, Except Natural Gas	767	423	55.15
47 Transportation Services	1,686	819	48.58
48 Communications	13,528	5,734	42.39
49 Electric, Gas, & Sanitary Services	11,798	11,122	94.27
50 Wholesale Trade – Durable Goods	5,353	2,674	49.95

Table IA. VIII (continued)

51 Wholesale Trade – Nondurable Goods	3,449	1,827	52.97
52 Building Materials & Gardening Supplies	531	334	62.90
53 General Merchandise Stores	2,316	918	39.64
54 Food Stores	1,817	800	44.03
55 Automative Dealers & Service Stations	1,747	1,256	71.89
56 Apparel & Accessory Stores	3,173	976	30.76
57 Furniture & Homefurnishings Stores	1,095	395	36.07
58 Eating & Drinking Places	3,655	1,359	37.18
59 Miscellaneous Retail	5,147	1,833	35.61
60 Depository Institutions	17,204	6,168	35.85
61 Nondepository Institutions	3,405	1,325	38.91
62 Security & Commodity Brokers	6,553	2,899	44.24
63 Insurance Carriers	10,060	4,226	42.01
64 Insurance Agents, Brokers, & Service	1,071	519	48.46
65 Real Estate	3,573	1,394	39.01
67 Holding & Other Investment Offices	18,046	7,526	41.70
70 Hotels & Other Lodging Places	1,246	388	31.14
72 Personal Services	820	299	36.46
73 Business Services	36,858	15,186	41.20
75 Auto Repair, Services, & Parking	626	435	69.49
78 Motion Pictures	982	281	28.62
79 Amusement & Recreation Services	2,636	1,067	40.48
80 Health Services	4,674	1,943	41.57
81 Legal Services	134	73	54.48
82 Educational Services	1,728	656	37.96
83 Social Services	281	93	33.10
87 Engineering & Management Services	4,953	2,882	58.19
89 Services, Not Elsewhere Classified	7	5	71.43
Total	330,906	169,786	51.31

Table IA. IX
Top Bigrams for Topic-Based Climate Change Exposure Measures

Panel A: Top-100 Opportunity Climate Change Bigrams ($CCExposure^{Opp}$)

Bigrams	Frequency	Bigrams	Frequency	Bigrams	Frequency
renewable energy	15,605	electrify vehicle	352	opportunity clean	272
electric vehicle	9,508	hybrid technology	339	safe clean	272
clean energy	6,430	energy vehicle	338	solar storage	272
new energy	4,544	vehicle lot	337	solar program	272
wind power	4,253	gigawatt install	337	geothermal power	270
wind energy	4,035	metal hydride	335	vehicle good	269
solar energy	2,511	gas clean	332	supply industrial	268
battery electric	1,121	focus renewable	331	cost renewable	267
solar farm	971	vehicle type	327	grid technology	265
heat power	941	renewable electricity	326	solar battery	263
combine heat	718	bus truck	326	ton carbon	262
carbon neutral	690	energy commitment	325	vehicle electric	260
cell power	657	support renewable	325	vehicle small	260
electric hybrid	585	battery charge	324	vehicle hybrid	259
carbon free	558	vehicle place	319	demand wind	259
sustainable energy	523	reduction carbon	310	power world	258
rooftop solar	498	vehicle space	309	construction wind	258
grid power	493	expand energy	308	term electric	257
vehicle charge	476	vehicle future	308	project solar	254
issue rfp	475	pure electric	305	carbon energy	254
charge infrastructure	469	fully electric	303	target gigawatt	252
construction megawatt	468	gas reduction	302	energy target	252
guangdong province	431	additional renewable	301	energy landscape	249
cell vehicle	413	invest renewable	298	affordable reliable	248
energy standard	406	cell electric	297	customer clean	248
energy renewable	403	community solar	288	conventional energy	247
hybrid car	403	emission reduce	288	efficient sustainable	245
include renewable	381	ton waste	287	vehicle talk	243
grid connect	376	type energy	282	charge network	243
solar capacity	375	energy goal	281	medical electronic	242
vehicle battery	374	vehicle development	280	efficiency renewable	239
micro grid	370	energy important	279	vehicle offer	238
build transmission	366	energy bring	277	vehicle opportunity	237
energy wind	352				

Table IA. IX (continued)

Panel B: Top-100 Regulatory Climate Change Bigrams ($CCExposure^{Reg}$)

Bigrams	Frequency	Bigrams	Frequency	Bigrams	Frequency
greenhouse gas	3,416	save technology	222	control upgrade	163
carbon emission	2,088	place energy	219	issue air	162
gas emission	1,910	carbon economy	217	gas regulation	162
carbon dioxide	1,583	talk clean	216	emission profile	162
air pollution	1,127	energy alternative	214	nitrous oxide	160
carbon price	999	meet renewable	208	receive air	159
energy regulatory	967	address environmental	207	air clean	158
carbon tax	928	change climate	207	produce carbon	156
combine heat	718	power initiative	204	reduce sulfur	156
environmental standard	593	climate action	204	national renewable	156
emission trade	480	produce renewable	199	require utility	156
dioxide emission	478	transition clean	198	market carbon	155
nox emission	475	produce clean	197	effective energy	154
energy renewable	403	reduce nox	194	impact clean	152
energy independence	399	carbon disclosure	194	product carbon	152
epa regulation	381	emission year	192	emission rate	150
development renewable	344	target energy	191	recovery pollution	150
support renewable	325	investment clean	189	emission compare	147
deliver clean	322	state renewable	188	emission increase	147
market clean	311	air resource	186	emission low	145
reduction carbon	310	address climate	184	water efficiency	145
gas reduction	302	environmental legislation	183	achieve carbon	144
carbon market	298	control regulation	180	economy emission	144
emission reduce	288	energy clean	179	capture sequestration	139
trade scheme	283	global climate	179	technology clean	138
cross state	279	use clean	177	clean job	137
emission intensity	268	gas initiative	177	emission improve	137
energy help	266	energy carbon	171	talk carbon	137
impact climate	265	efficient natural	170	emission energy	136
reduce air	254	promote energy	169	generate renewable	136
efficiency renewable	239	source electricity	167	nation energy	135
carbon offset	230	energy smart	166	emission come	135
disclosure project	229	efficiency environmental	163	ghg emission	133
emission free	223				

Table IA. IX (continued)

Panel C: Top-50 Physical Climate Change Bigrams ($CCExposure^{Phy}$)

Bigrams	Frequency	Bigrams	Frequency	Bigrams	Frequency
global warm	837	heavy snow	260	battery hybrid	96
coastal area	816	security energy	238	fight global	86
electric bus	709	water discharge	233	land forest	84
snow ice	538	sea water	232	product landscape	84
forest land	512	ice product	202	partially unfavorable	78
wind speed	489	management district	193	particularly coastal	70
provide water	429	water act	187	especially coastal	68
sea level	421	management water	172	strong preseason	68
area florida	402	hydride battery	168	shipment battery	67
coastal region	389	weather snow	165	use lithium	65
nickel metal	375	air clean	158	area coastal	63
supply water	352	water food	148	performance lithium	62
metal hydride	335	ice control	142	separator film	60
natural hazard	295	value forest	130	solution act	57
storm water	292	non coastal	117	residential utility	56
air water	290	sale forest	110	fluorine product	52
quality water	277	contractor product	96		

Table IA. X
Snippets of Top Climate Change Exposure Firms

Firm	HQ	SIC	Time	Assets ($ millions)	Bigrams	Top Snippet
China Longyuan Power Group Corp Ltd	China	4991	2013Q4	18,136	good wind; wind speed	in general the experience is the good wind results year probably will follow a low wind speed year and if the wind speed in the northern part is good probably in the southern parts the wind speed would be lower.
Xinjiang Goldwind Science & Technology Co Ltd	China	3510	2019Q4	14,750	grid tariff; potential wind; renewable energy; tariff wind; wind power	the last slide here presents research results of china renewable energy engineering institute, showing that on-grid tariff of wind power in majority of regions in china has reached the same level with benchmark coal-fired power tariff, demonstrating the potential of wind power marketed transaction.
ECOtality Inc	U.S.	3620	2009Q2	20	charge infrastructure; development electric; electric vehicle; promote development; vehicle charge	we also achieved significant operational milestones with our partnerships with nissan; the maricopa county association of governments, which represents the phoenix metropolitan area, as well as the pima county association of government, which represents the tucson metropolitan area; in order to advance zero emission mobility by promoting the development of electric vehicle and charging infrastructure.
China Ming Yang Wind Power Group Ltd	China	3510	2015Q2	2567	biomass energy; clean energy; consumption energy; development wind; energy china; solar biomass	it was stated clearly in the government's 2015 work report that the development of wind, solar and biomass energy should be strongly promoted, and we should accelerate the consumption of clean energy to boost our revolution in the consumption of energy in china.
Advanced Battery Technologies Inc	U.S.	3690	2009Q3	158	advance battery; electric vehicle; focus electric; vehicle china	as we move through 2009, our key initiatives include aligning ourselves to benefit from the increasing focus on electric vehicles in china and worldwide, especially in china, where government initiatives will provide meaningful incentives; leveraging our current leadership to secure new contracts, especially large-scale rechargeable polymer lithium-ion battery sales, and ultimately driving revenue mix shift to reflect higher-margin sales; ensuring an improving operational efficiency at both advanced battery and wuxi zq entities.
Ocean Power Technologies Inc	U.S.	3511	2008Q4	97	energy requirement; increase renewable; population center; powerbuoy wave; renewable energy; wave condition; wave power	these areas represent strong potential markets for our powerbuoy wave power stations because they combine favorable wave conditions, political and economic stability, large population centers, high levels of industrialization, and significant and increasing renewable energy requirements.

Table IA. X (continued)

Firm	HQ	SIC	Time	Assets ($ millions)	Bigrams	Top Snippet
Otter Tail Corp	U.S.	4911	2015Q3	1,821	dioxide mercury; emission nitrogen; oxide sulfur; regional haze; sulfur dioxide	remember, the state-of-the-art control system will reduce emissions of nitrogen oxide, sulfur dioxide, and mercury by 80% to 90% to meet the epa's regional haze and mats requirements.
FuelCell Energy Inc	U.S.	3690	2010Q4	151	cell power; coal derive; emission coal; gas emission; greenhouse gas; solid oxide	another deal we contract, we are working toward a long range goal of developing megawatt class, solid oxide, fuel cell power plants filled with coal-derived synthesis gas, thereby reducing greenhouse gas emissions from coal up to 90%.
ALLETE Inc	U.S.	4931	2019Q3	5,483	carbon free; clean energy; double wind; wind energy; wind facility	with approximately 555 megawatts of carbon-free wind generation already in operation, allete clean energy is on schedule with its planned construction of several new wind facilities that upon completion will roughly double its wind energy fleet, adding almost 500 megawatts in total generation capability.
Clean Energy Fuels Corp	U.S.	5500	2017Q3	792	air quality; clean air; fuel renewable; nox engine; renewable natural	moving on to the clean air action plan that is being drafted by the ports of la and long beach, we believe ultimately that any final plan must immediately address the horrendous air quality by requiring the thousands of trucks that operate on dirty diesel at the ports to be replaced with new zero nox engines fueled by renewable natural gas.

Table IA. XI
Top-100 Unigrams and Bigrams Captured by $CCExposure^{EGKLS}$

This table reports the top-100 unigrams and bigrams associated with $CCExposure^{EGKLS}$, which measures the relative frequency with which the pre-specified list of unigrams and bigrams from Engle et al. (2020, EGKLS) appear in earnings call transcripts.

Uni/Bigrams	Frequency	Uni/Bigrams	Frequency	Uni/Bigrams	Frequency
market	7,271,737	reduction	837,331	land	242,464
increase	6,125,831	unit	809,445	party	239,304
time	4,859,969	potential	794,827	national	234,016
term	4,527,681	effect	779,984	weather	229,339
cost	4,508,616	set	633,825	natural	228,976
result	4,422,260	world	613,867	develop	227,216
high	3,834,354	gas	597,801	response	219,786
impact	2,759,159	global	593,695	establish	208,680
net	2,539,728	international	583,436	water	199,055
include	2,407,273	measure	575,794	define	151,324
level	2,403,504	event	549,727	implementation	148,991
base	2,290,149	country	541,835	wind	138,807
project	1,658,992	region	495,348	air	129,668
area	1,465,194	plant	488,014	scenario	129,584
balance	1,415,498	pressure	479,484	chemical	100,768
report	1,348,573	power	471,925	feedback	92,600
future	1,335,014	energy	461,752	assessment	86,797
development	1,309,010	condition	451,798	social	83,111
range	1,271,805	organic	436,245	solar	75,186
benefit	1,266,657	economic	433,424	environmental	74,964
current	1,200,773	relative	427,686	mass	72,671
activity	1,193,239	cycle	383,043	human	60,708
process	1,188,852	action	373,261	mechanism	57,164
average	1,120,217	produce	368,668	layer	53,525
production	1,081,807	form	368,562	framework	52,067
group	1,014,513	refer	357,518	sea	48,936
technology	1,001,416	resource	341,490	concentration	48,560
reduce	986,884	fuel	303,120	carbon	44,655
place	933,857	source	293,965	surface	44,483
number	917,925	industrial	291,517	protocol	41,692
state	880,937	occur	273,066	ecosystem	35,182
environment	877,466	live	270,116	emission	34,987
capacity	859,368	policy	259,081	warm	34,698
model	857,032				

Table IA. XII
Climate Change Exposure Measures: Comparison with Measure using EGLKS's Keywords

This table reports summary statistics for a climate change exposure measure constructed from the list of *pre-specified* climate change keywords in listed in EGKLS's Figure 1. To create the alternative measure, $CCExposure^{EGKLS}$, we replace our bigrams set $\mathbb{C}$ with $\mathbb{C}^{EGKLS}$ and then recompute the relative frequency with which the alternative terms appear in the transcripts. We construct a frequency-unweighted and a TFIDF version, denoted by $CCExposure^{EGKLS-EW}$ or $CCExposure^{EGKLS-TFIDF}$, respectively. Panel A reports summary statistics, and Panel B correlations. In Panel B, we report overall sample correlations and correlations depending on whether the time-series index by EGKLS of public climate change attention indicates that such attention is high (*WSJ CC News Index* is in the top quartile) or low. The Appendix defines all variables in detail.

Panel A: Summary Statistics (x10^3)

	Mean	STD	25%	Median	75%	N
$CCExposure_{i,t}^{EGLKS-EW}$	54.0	11.0	47.4	52.9	59.8	86,152
$CCExposure_{i,t}^{EGLKS-TFIDF}$	17.3	8.5	12.1	15.0	20.0	86,152

Panel B: Correlations

	$CCExposure_{i,t}$	$CCExposure_{i,t}^{EGLKS-EW}$
$CCExposure_{i,t}^{EGLKS-EW}$	0.35	1.00
$CCExposure_{i,t}^{EGLKS-TFIDF}$	0.59	0.73
	WSJ Climate Change Index High (Top 25%)	
$CCExposure_{i,t}^{EGLKS-EW}$	0.38	1.00
$CCExposure_{i,t}^{EGLKS-TFIDF}$	0.62	0.73
	WSJ Climate Change Index Low (Bottom 75%)	
$CCExposure_{i,t}^{EGLKS-EW}$	0.33	1.00
$CCExposure_{i,t}^{EGLKS-TFIDF}$	0.55	0.73

Table IA. XIII
Green-Tech Jobs and Green Patents Results: Controlling for Emissions

This table reports regressions that relate green-tech jobs and green patents to the climate change exposure measures. Regressions are estimated at the firm-year level. The regressions complement those in Tables VII and VIII by additionally controlling for a firm's carbon emissions ($Total\,Emissions$). The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\text{-}Tech\ Jobs_{i,t+1}$				$\#Green\ Patents_{i,t+1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(1 + CCExposure_{i,t})$	1.390***				1.505***			
	(0.192)				(0.241)			
$Log(1 + CCExposure^{Opp}_{i,t})$		1.799***				1.311***		
		(0.275)				(0.231)		
$Log(1 + CCExposure^{Reg}_{i,t})$			1.102***				3.104***	
			(0.382)				(0.344)	
$Log(1 + CCExposure^{Phy}_{i,t})$				1.316				-7.494
				(0.829)				(5.201)
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Sample	US	US	US	US	US	US	US	US
Carbon Emissions	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	8,767	8,767	8,767	8,767	5,417	5,417	5,417	5417
Ps. R^2	0.778	0.791	0.730	0.728	0.597	0.568	0.573	0.554

Table IA. XIV
Green-Tech Jobs Results: Alternative Model Specifications

This table reports regressions that relate green-tech jobs to the climate change exposure measures. Regressions are estimated at the firm-year level. The table estimates different alternative model specifications of the regressions in Tables VII. Column (1) additionally controls for proxies for strategic disclosure, notably, the firm's overall sentiment (the share of positive and negative tone words across the full earnings call transcript) and two proxies for the firm's recent financial performance. We measure performance as the pre-call stock return (accumulated over the seven days before the earnings call) and the earnings surprise. Earnings surprise is defined as earnings per share before extraordinary items minus earnings per share in the same quarter of the prior year, divided by the price per share at the beginning of the quarter (Ball and Bartov (1996)). We average the two performance variables across the earnings calls of a firm-year to obtain an annual measure. Column (2) is estimated within the Burning Glass (BG) sample (i.e., when we do not impute missing *#Green-Tech Jobs* data). The OLS models with industry-year fixed effects permit more observations (the linear model averages out the incidental parameter problem) than the Poisson models. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\text{-}Tech\ Jobs_{i,t+1}$							$Log(1+\#Green\text{-}Tech\ Jobs_{i,t+1})$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Log(1+CCExposure_{i,t})$	1.679***	1.511***				10.006***		0.218***	0.187***
	(0.219)	(0.193)				(2.641)		(0.020)	(0.019)
$Log(1+CCExposure^{Q\&A}_{i,t})$			1.114***						
			(0.175)						
$Log(1+CCSentiment^{Pos}_{i,t})$				1.052***					
				(0.313)					
$Log(1+CCSentiment^{Neg}_{i,t})$				0.568					
				(0.370)					
$Log(1+CCRisk_{i,t})$					3.281***				
					(0.579)				
$CCExposure_{i,t}$							1.642***		
							(0.451)		
Model	Poisson	Poisson	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS
Sample	US	US, BG	US	US	US	US	US	US	US
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes
N	16,892	15,840	23,870	23,870	23,870	28,963	28,963	28,963	28,934
Adj./ps. R^2	0.778	0.766	0.735	0.719	0.712	0.007	0.006	0.085	0.112

Table IA. XV
Green Patents Results: Alternative Model Specifications

This table reports regressions that relate green patents to the climate change exposure measures. Regressions are estimated at the firm-year level. The table estimates different alternative model specifications of the regressions in Tables VIII. Column (1) additionally controls for proxies for strategic disclosure, notably, the firm's overall sentiment (the share of positive and negative tone words across the full earnings call transcript) and two proxies for the firm's recent financial performance. We measure performance as the pre-call stock return (accumulated over the seven days before the earnings call) and the earnings surprise. Earnings surprise is defined as earnings per share before extraordinary items minus earnings per share in the same quarter of the prior year, divided by the price per share at the beginning of the quarter (Ball and Bartov (1996)). We average the two performance variables across the earnings calls of a firm-year to obtain an annual measure. Column (2) is estimated within the Google Patent (GP) sample (i.e., when we do not impute missing *#Green Patents* data). The OLS models with industry-year fixed effects permit more observations (the linear model averages out the incidental parameter problem) than the Poisson models. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\ Patents_{i,t+1}$							$Log(1+\#Green\ Patents_{i,t+1})$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$Log(1+CCExposure)_{i,t}$	1.442***	0.955***				0.147***		0.028***	0.016***
	(0.183)	(0.142)				(0.032)		(0.004)	(0.003)
$Log(1+CCExposure^{Q\&A}_{i,t})$			0.878***						
			(0.201)						
$Log(1+CCSentiment^{Pos}_{i,t})$				0.754***					
				(0.225)					
$Log(1+CCSentiment^{Neg}_{i,t})$				1.028***					
				(0.388)					
$Log(1+CCRisk)_{i,t}$					2.534***				
					(0.723)				
$CCExposure_{i,t}$							0.013***		
							(0.004)		
Model	Poisson	Poisson	Poisson	Poisson	Poisson	OLS	OLS	OLS	OLS
Sample	US	US, GP	US	US	US	US	US	US	US
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes
N	15,020	3,692	21,914	21,914	21,914	43,390	43,390	43,390	43,348
Adj./ps. R^2	0.639	0.687	0.612	0.615	0.601	0.009	0.009	0.029	0.043

Table IA. XVI
Green-Tech Jobs and Green Patents Results: $CCExposure^{Initial} = 0$ Subsample

This table reports regressions that relate green-tech jobs and green patents to the climate change exposure measures. Regressions are estimated at the firm-year level. The table estimates the same model specifications as in Tables VII and VIII, but within the sample in which $CCExposure^{Initial} = 0$ (or the corresponding topic-based measures). $CCExposure^{Initial}$ is the climate change exposure score computed, based on the initial seed bigrams in Table III only. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

	$\#Green\text{-}Tech\ Jobs_{i,t+1}$				$\#Green\ Patents_{i,t+1}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Log(1 + CCExposure_{i,t})$	1.161*				1.966***			
	(0.702)				(0.336)			
$Log(1 + CCExposure^{Opp}_{i,t})$		2.511***				2.023***		
		(0.836)				(0.530)		
$Log(1 + CCExposure^{Reg}_{i,t})$			4.263***				6.184***	
			(1.064)				(0.792)	
$Log(1 + CCExposure^{Phy}_{i,t})$				1.305				-11.862**
				(1.901)				(5.276)
Model	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson	Poisson
Sample	US	US	US	US	US	US	US	US
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry x Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	18,704	20,885	22,036	22,981	16,470	18,708	19,738	21,054
Ps. R^2	0.442	0.440	0.687	0.678	0.686	0.656	0.627	0.611

Table IA. XVII
Comparison of Observations with and without Within-Fixed Effects Variation

This table compares observations with and without variation within industry-year groups. The Poisson regressions in Tables VII and VIII base the estimation only on observations with at least one nonzero value within a given industry-year group. This restriction is desirable to avoid biased estimators. It restricts the usable sample to those groups that are informative about the effects of *CCExposure* (Cohn, Liu, and Wardlaw (2022)). Panel A reports statistics for the green-tech job estimation, and Panel B for the green patent estimation. The Appendix defines all variables in detail. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Green-Tech Jobs Estimation

	Observations Included in the Estimation (N=23,870)			Observations Excluded from the Estimation (N=5,093)			
	Mean	Std.Dev.	Median	Mean	Std.Dev.	Median	Difference-in-Means
$CCExposure_{i,t}$	1.100	2.784	0.306	0.536	1.112	0.258	0.564***
$CCExposure^{Opp}_{i,t}$	0.337	1.279	0.000	0.111	0.402	0.000	0.225***
$CCExposure^{Reg}_{i,t}$	0.046	0.254	0.000	0.021	0.111	0.000	0.026***
$CCExposure^{Phy}_{i,t}$	0.013	0.108	0.000	0.012	0.065	0.000	0.001
$Assets_{i,t}$	8872	30271	1223	10732	36266	1721	-1860***
$Debt/Assets_{i,t}$	0.246	0.232	0.203	0.324	0.249	0.287	-0.077***
$Cash/Assets_{i,t}$	0.199	0.221	0.112	0.115	0.143	0.061	0.084***
$PPE/Assets_{i,t}$	0.216	0.239	0.118	0.282	0.267	0.199	-0.066***
$EBIT/Assets_{i,t}$	0.013	0.202	0.056	0.071	0.111	0.070	-0.058***
$CAPEX/Assets_{i,t}$	0.041	0.051	0.024	0.045	0.051	0.031	-0.004***
$R\&D/Assets_{i,t}$	0.057	0.110	0.001	0.006	0.030	0.000	0.051***

Panel B: Green Patents Estimation

	Observations Included in the Estimation (N=21,914)			Observations Excluded from the Estimation (N=21,476)			
	Mean	Std.Dev.	Median	Mean	Std.Dev.	Median	Difference-in-Means
$CCExposure_{i,t}$	1.107	2.653	0.336	0.781	2.236	0.253	0.327***
$CCExposure^{Opp}_{i,t}$	0.330	1.232	0.000	0.209	0.978	0.000	0.122***
$CCExposure^{Reg}_{i,t}$	0.047	0.271	0.000	0.032	0.167	0.000	0.015***
$CCExposure^{Phy}_{i,t}$	0.012	0.116	0.000	0.013	0.076	0.000	-0.001
$Assets_{i,t}$	5142	19303	653	11384	36924	1730	-6241***
$Debt/Assets_{i,t}$	0.213	0.222	0.168	0.284	0.240	0.245	-0.071***
$Cash/Assets_{i,t}$	0.253	0.239	0.171	0.124	0.159	0.063	0.128***
$PPE/Assets_{i,t}$	0.207	0.227	0.119	0.252	0.253	0.166	-0.045***
$EBIT/Assets_{i,t}$	-0.008	0.228	0.057	0.059	0.127	0.064	-0.067***
$CAPEX/Assets_{i,t}$	0.044	0.053	0.026	0.042	0.048	0.027	0.002***
$R\&D/Assets_{i,t}$	0.085	0.122	0.037	0.011	0.051	0.000	0.074***

REFERENCES

Bakshi, Gurdip S., Nikunj Kapadia, and Dilip B. Madan, 2003, Stock return characteristics, skew laws, and the differential pricing of individual equity options, *Review of Financial Studies* 16, 101–143.

Bali, Turan G., and Hao Zhou, 2016, Risk, uncertainty, and expected returns, *Journal of Financial and Quantitative Analysis* 51, 707–735.

Ball, Ray, and Eli Bartov, 1996, How naive is the stock market's use of earnings information?, *Journal of Accounting and Economics* 21, 319–337.

Chang, Bo-Young, Peter Christoffersen, Kris Jacobs, and Gregory Vainberg, 2012, Option-implied measures of equity risk, *Review of Finance* 16, 385–428.

Cohn, Jonathan B, Zack Liu, and Malcolm Wardlaw, 2022, Count (and count-like) data in finance, Working paper, University of Texas at Austin.

Cremers, K. J. Martijn, Michael Halling, and David Weinbaum, 2015, Aggregate jump and volatility risk in the cross-section of stock returns, *Journal of Finance* 70, 577–614.

DeMiguel, Victor, Yuliya Plyakha, Raman Uppal, and Grigory Vilkov, 2013, Improving portfiolio selection using option-implied volatility and skewness, *Journal of Financial and Quantitative Analysis* 48, 1813–1845.

Engle, Robert F., Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2020, Hedging climate change news, *Review of Financial Studies* 33, 1184–1216.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun, 2019, Firm-level political risk: Measurement and effects, *Quarterly Journal of Economics* 134, 2135–2202.

Kelly, Bryan, Lubos Pastor, and Pietro Veronesi, 2016, The price of political uncertainty: Theory and evidence from the option market, *Journal of Finance* 71, 2417–2480.

King, Gary, Patrick Lam, and Margaret E. Roberts, 2017, Computer-assisted keyword and document set discovery from unstructured text, *American Journal of Political Science* 61, 971–988.

Poon, Ser-Huang, and Clive W.J. Granger, 2003, Forecasting volatility in financial markets: A review, *Journal of Economic Literature* 41, 478–539.

Vanden, Joel M., 2008, Information quality and options, *Review of Financial Studies* 21, 2635–2676.

about ECGI

The European Corporate Governance Institute has been established to improve *corporate governance through fostering independent scientific research and related activities.*

The ECGI will produce and disseminate high quality research while remaining close to the concerns and interests of corporate, financial and public policy makers. It will draw on the expertise of scholars from numerous countries and bring together a critical mass of expertise and interest to bear on this important subject.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

Electronic Access to the Working Paper Series

The full set of ECGI working papers can be accessed through the Institute's Web-site (www.ecgi.global/content/working-papers) or SSRN:

Finance Paper Series	http://www.ssrn.com/link/ECGI-Fin.html
Law Paper Series	http://www.ssrn.com/link/ECGI-Law.html