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: 10.1111/jofi.13219. 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***
$CCExposure^{Opp}_{i,t}$	0.379	1.362	0.000	0.265	1.129	0.000	0.114***
$CCExposure^{Reg}_{i,t}$	0.054	0.256	0.000	0.038	0.218	0.000	0.016***
$CCExposure^{Phy}_{i,t}$	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)			
		$CCExpo_{i,t}$	$CCExpo_{i,t}^{Opp}$	$CCExpo_{i,t}^{Reg}$	$CCExpo_{i,t}^{Phy}$	$CCExpo_{i,t}$	$CCExpo_{i,t}^{Opp}$	$CCExpo_{i,t}^{Reg}$	$CCExpo_{i,t}^{Phy}$
EW Measures	$CCExposure_{i,t}$	1.000							
	$CCExposure_{i,t}^{Opp}$	0.897	1.000						
	$CCExposure_{i,t}^{Reg}$	0.523	0.301	1.000					
	$CCExposure_{i,t}^{Phy}$	0.224	0.157	0.092	1.000				
TFIDF Measures	$CCExposure_{i,t}$	0.997	0.882	0.521	0.222	1.000			
	$CCExposure_{i,t}^{Opp}$	0.900	0.994	0.306	0.156	0.892	1.000		
	$CCExposure_{i,t}^{Reg}$	0.519	0.295	0.992	0.092	0.520	0.300	1.000	
	$CCExposure_{i,t}^{Phy}$	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.