Imperial College Business School
Department of Finance

# Text Mining in Climate Finance

Yi Zhang

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Finance of of Imperial College London and, March 2023

## Declaration of Originality

I herewith declare that this thesis is my own work. All material and sources that do not constitute my own work have been explicitly acknowledged and referenced.

***Yi Zhang***

© The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

***Yi Zhang***

# Abstract

This thesis is to explore the relationship between climate risk and stock returns.

In the first chapter, I develop a quantitative dynamic general equilibrium model to explore how climate regulatory risk is reflected in cross-sectional asset pricing. The representative household has preferences for low carbon. Greener environmental preference makes the market price of climate policy risk more negative conditional on a low elasticity of substitution between green and brown capital. The quantitative implications of the model can rationalize the recent empirical evidence on positive price of carbon risk. Finally, using textual analysis to measure transition risk from 10K filings, the paper shows that lower transition risk-exposed firms carry a 8.4% annualized risk premium. Consistent with the model, the price of transition risk tends to be negative when climate awareness is high.

In the second chapter, we propose a measure of firm-level climate regulatory exposure based on 10-K filings. Using the 2016 Trump election as an exogenous shock to perceived climate regulatory risks, we identify a positive effect on stock returns for firms with higher climate regulatory exposures; they experience economically and statistically significant higher cumulative returns post-election. In the year following the election, firms with higher climate regulatory exposure experience higher carbon emissions and lower investor attention. Both findings indicate that, post-election, investors become less concerned with climate regulatory risks. Results are robust to physical climate, trade, tax and oil price exposures.

In the third chapter, We develop a machine learning framework to extract the text-implied climate risk in the SEC 10K filings for the text-implied price of risk estimation. We exploit both supervised and unsupervised learning algorithms in NLP to accomplish this task. The proposed risk exposure is testable within the Fama-MacBeth framework and has statistical sufficiency. Empirically, this technique could generate a monthly abnormal return on climate risk from 1.2% to 2.1%, depending on the specification of the benchmark model. The pricing finding is robust at the firm level for the Fama-MacBeth regression. Additionally, we find that the sufficient text-implied risk exposure is consistent with recent empirical evidence with sensible external validation.

# Contents

**Declaration** **1**

**Abstract** **2**

**1 Dissecting Carbon Risk: An Asset Pricing View from Climate Policy** **1**

1.1 Introduction . . . . . 1

1.2 Related Literature . . . . . 5

1.3 A Two-Period Model . . . . . 6

1.4 Dynamic General Equilibrium Model . . . . . 10

1.4.1 Household . . . . . 11

1.4.2 Intermediate Good Sector . . . . . 12

1.4.3 Energy Sector . . . . . 15

1.4.4 Final Good Sector . . . . . 16

1.4.5 Carbon Cycle . . . . . 17

1.4.6 Government . . . . . 17

1.4.7 Market Equilibrium . . . . . 18

1.5 Model Solution . . . . . 19

1.5.1 Investment and Equity Return . . . . . 19

1.5.2 Climate Adjusted SDF . . . . . 20
1.6 Quantitative Implication . . . . . 21
1.6.1 Calibration . . . . . 21
1.6.2 Model Prediction . . . . . 23
1.6.3 Endogenous Climate Policy . . . . . 27
1.7 Data . . . . . 29
1.7.1 SEC Filings . . . . . 29
1.7.2 Stock Returns and Firm Fundamentals . . . . . 31
1.8 Empirical Results . . . . . 31
1.8.1 Portfolio Evidence . . . . . 31
1.8.2 Abnormal Return . . . . . 33
1.8.3 Predictive Panel Regression . . . . . 33
1.8.4 Event Study . . . . . 35
1.8.5 Carbon Emission . . . . . 36
1.8.6 Book-to-Market Ratio . . . . . 37
1.9 Conclusion . . . . . 37
1.10 Appendix A. . . . . 66
1.10.1 Figure . . . . . 66
1.10.2 Proof 1: Two-Period Model Solution . . . . . 75
1.10.3 Proof 2: . . . . . 76
1.10.4 Decentralized Economics . . . . . 77
1.10.5 Model Equation . . . . . 77
1.11 Appendix B. . . . . 85

**2 Firm-Level Climate Regulatory Exposure** **86**

2.1 Introduction . . . . . 86

2.2 Empirical Strategy . . . . . 91

2.2.1 Hypothesis Development . . . . . 91

2.2.2 Identification . . . . . 92

2.2.3 Roadmap . . . . . 93

2.3 A Measure of Climate Regulatory Exposure . . . . . 93

2.4 Data . . . . . 95

2.4.1 SEC Filings . . . . . 95

2.4.2 Text Mining Data . . . . . 96

2.4.3 Stock Returns, Firm Fundamentals and Carbon Emissions . . . . . 97

2.4.4 EDGAR Attention Measure . . . . . 98

2.4.5 Bloomberg Readership . . . . . 99

2.5 Climate Regulatory Exposure . . . . . 99

2.5.1 Word Clouds . . . . . 99

2.5.2 Time Series Variation . . . . . 100

2.5.3 Cross-Industry Variation . . . . . 101

2.5.4 CRE External Validation . . . . . 101

2.6 Empirical Results . . . . . 102

2.6.1 Donald Trump Election . . . . . 102

2.6.2 Paris Agreement . . . . . 108

2.6.3 Climate Regulatory Risk vs Alternative Climate Risks . . . . . 109

2.7 Supportive Evidence . . . 110
2.7.1 Carbon Emissions . . . 110
2.7.2 Investor Attention . . . 111
2.8 Conclusion . . . 113
2.9 Figures . . . 130
Appendix A . . . 137
Appendix B . . . 141

**3 Pricing Climate Risk: A Machine Learning Approach 150**

3.1 Introduction . . . 150
3.2 Related Literature . . . 153
3.2.1 Textual Analysis . . . 153
3.2.2 Climate Risk . . . 154
3.3 Data . . . 155
3.4 Methodology . . . 156
3.4.1 Feature Extraction . . . 157
3.4.2 Econometric Specification . . . 158
3.4.3 Multinomial Inverse Regression . . . 159
3.4.4 Three-Pass Estimation . . . 161
3.5 Application in Climate Finance . . . 162
3.5.1 Validation . . . 162
3.5.2 Empirical Analysis . . . 165
3.5.3 Portfolio Analysis: Robustness Tests . . . 168

3.6 Conclusion . . . . . 169

3.7 Derivation . . . . . 190

3.7.1 Proof of Proposition 1 . . . . . 190

3.8 Dictionary Construction Details . . . . . 192

3.8.1 Discovery of Climate Change Unigrams . . . . . 192

3.8.2 Discovery of Climate Change Bigrams . . . . . 192

3.8.3 Integrated Dictionary . . . . . 194

3.9 Extra Experimental Results . . . . . 195

3.9.1 Variable Definitions . . . . . 195

3.9.2 Accumulated Returns of Different Factor-Mimicking Portfolios . . . . . 196

3.9.3 Cross-sectional Distribution of Projections . . . . . 197

3.9.4 Climate Change Vocabulary . . . . . 198

3.9.5 Visualisation of the Aggregate DTM . . . . . 200

**Bibliography** **202**

# List of Tables


1.1 Summary of Statistics for 10K Raw Information . . . . . 39

1.2 Summary of Statistics for Firm Characteristics . . . . . 40

1.3 Calibration Parameter in the Model . . . . . 41

1.4 Macro Fundamentals and Key Moment Simulation . . . . . 42

1.5 One-Way Portfolio Sorted Result . . . . . 43

1.6 Two-Way Portfolio Sorted Result . . . . . 44

1.7 Two-Way Portfolio Sorted Result . . . . . 45

1.8 Portfolio Based Asset Pricing Test . . . . . 46

1.9 Predictive Equity Return Regression . . . . . 47

1.10 Fama-MacBeth Cross Sectional Regression for Asset Pricing Test . . . . . 48

1.11 Carbon Emission Regression and Climate Regulatory Exposure . . . . . 49

1.12 Book-to-Market Ratio and CRE regression . . . . . 50

1.13 Two-Way Portfolio Sorted Result . . . . . 70

1.14 Predictive Equity Return Regression Controlling Direct Carbon Emission . . . . 71

1.15 Predictive Equity Return Regression Controlling Indirect Carbon Emission . . . 72

1.16 Climate Attention and Market Price of Climate Regulation . . . . . 73

1.17 Carbon Emission and Climate Regulatory Exposure without Industry Fixed Effect 74

1.18 Matched Top 30 GPV and CCV . . . . . 85

2.1 Summary Statistics of 10-K Filings . . . . . 114

2.2 Summary Statistics of Firm Characteristics . . . . . 115

2.3 CRE Summary Statistics by Year . . . . . 116

2.4 CRE Summary Statistics by Industry from 2006 to 2018 . . . . . 117

2.5 Determinants of SEC based CRE disclosure . . . . . 118

2.6 Difference-in-Difference around Trump Election with CRE (Cumulative Raw Return) . . . . . 119

2.7 Difference-in-Difference around Trump Election (Cumulative Raw Return) . . . 120

2.8 Difference-in-Differences (Cumulative Abnormal Returns) . . . . . 121

2.9 CAR Regression on CRE for a Different Training Sample (2009-2015) . . . . . 122

2.10 Controlling for Alternative Explanations . . . . . 123

2.11 Placebo Test: 2012 U.S. Presidential Election . . . . . 124

2.12 Difference-in-Difference around Paris Agreement (Cumulative Raw Return) . . . 125

2.13 Difference-in-Difference around Donald Trump Election with Confounding Climate Variation . . . . . 126

2.14 Difference-in-Difference around Paris Agreement with Confounding Climate Variation . . . . . 127

2.15 Carbon Emission Difference-in-Differences around Trump Election . . . . . 128

2.16 Information Attention Difference-in-Differences around Trump Election . . . . . 129

A1 Matched Top 30 GPV and CCV . . . . . 137

A2 Matched Climate Change Exposure (CCE) . . . . . 138

A3 CCE Summary Statistics by Year . . . 139

A4 CCE Summary Statistics by Industry from 2006 to 2018 . . . 140

A5 Difference-in-Differences (Cumulative Abnormal Returns) at Daily Frequency Around Trump Interview, Nov 23rd, 2016 . . . 146

A6 Difference-in-Difference around Paris Agreement (CAPM Cumulative Abnormal Return) . . . 147

A7 Difference-in-Difference around Paris Agreement (CAPM Cumulative Abnormal Return) . . . 148

A8 Difference-in-Difference around Paris Agreement (FF3 Cumulative Abnormal Return) . . . 149

A1 List of the IPCC Reports . . . 170

A2 Top-5 Similar Words of Initial Unigrams . . . 171

A3 Initial Bigrams for Searching Climate Change Bigrams . . . 172

A4 Top-100 Bigrams Captured by Climate Change Exposure . . . 173

A5 Summary Statistics . . . 174

A6 Paris Agreement and Stock Returns: Difference-in-differences Estimation . . . 175

A7 Climate Risk and Carbon Intensity . . . 176

A8 Climate Risk Portfolios . . . 177

A9 Correlations and Auto-correlations of Different Factors . . . 178

A10 Double-Sorted Portfolios . . . 179

A11 Individual-stock-level Fama-MacBeth Cross-sectional Regressions . . . 180

A12 Climate Risk Portfolios - Robustness Tests . . . 181

A13 Variable Definitions . . . 195

# List of Figures

1.1 Cross-Industry Cumulative Raw Return from 2009 to 2015 . . . . . 51

1.2 Media Climate Change Concern from 2000 to 2019 . . . . . 52

1.3 Rolling Average of Carbon Growth from 1990 to 2019 . . . . . 53

1.4 Multi-Sector Dynamic Model . . . . . 54

1.5 Multi-Sector Dynamic Model . . . . . 55

1.6 Impulse Response Function of Pricing Performance to Climate Transition Risk . 56

1.7 Fundamental Response to Climate Policy Shock . . . . . 57

1.8 Impulse Response to Productivity Risk (LRR vs SRR) . . . . . 58

1.9 Impulse Response to Long Run Risk . . . . . 59

1.10 Impulse Response Function and Elasticity of Substitution . . . . . 60

1.11 Book-to-Market Impulse Response to Risks . . . . . 61

1.12 Impulse Response to Productivity Risk with Endogenous Climate Policy . . . . 62

1.13 Cross-Industry Variation of the CRE . . . . . 63

1.14 Stock Return around Climate Regulatory Events: Paris Agreement . . . . . 64

1.15 Stock Return around Climate Deregulatory Events: Trump Election . . . . . 65

1.16 Fundamental Impulse Response to Climate Regulatory Risk . . . . . 66

1.17 Fundamental Impulse Response to Short Run Productivity Risk . . . . . 67

1.18 Numerical Solutions for Determinant State Variables: $log(F_{ss})$ and $\theta_{ss}$ . . . . . 68

1.19 Fundamental and Realized Carbon Price at Steady State . . . . . 69

2.1 Word Cloud of Climate Regulatory Exposure . . . . . 130

2.2 Word Cloud of Physical Climate Exposure . . . . . 130

2.3 Aggregate Climate Regulatory and Physical Exposure . . . . . 131

2.4 Climate Regulatory Exposure (CRE) by Industry . . . . . 132

2.5 Climate Change Exposure (CCE) by Industry . . . . . 132

2.6 Cross Sectional Correlation between CRE and alterntive Climate Measures . . . 133

2.7 Cumulative Raw Returns (October 2016 to January 2017) . . . . . 134

2.8 Cumulative Raw Returns (October 2012 to January 2013) . . . . . 135

2.9 Cumulative Raw Returns (October 2015 to January 2016) . . . . . 136

3.1 Clusters of Similar Words from IPCC Reports (preplexity=20) . . . . . 182

3.2 Term Frequency of Initial Unigrams in IPCC Reports . . . . . 183

3.3 Word Cloud of the Aggregate Document-Term Matrix from 2009 to 2012 . . . . . 184

3.4 Cumulative Recall and Precision . . . . . 185

3.5 Projection Heatmap at Industry-Level . . . . . 186

3.6 Projection Heatmap of Firms . . . . . 187

3.7 Return Volatility of the Factor Mimicking Portfolio . . . . . 188

3.8 Climate Risk Portfolios - Robustness Tests (Industry Excl.) . . . . . 189

3.9 Accumulated Returns of Different Factor-Mimicking Portfolios . . . . . 196

3.10 Cross-sectional Distribution of Projections . . . . . 197

3.11 Word Cloud of the Aggregate Document-Term Matrix from 2005 to 2008 . . . . . 200

3.12 Word Cloud of the Aggregate Document-Term Matrix from 2013 to 2016 . . . . . 201

3.13 Word Cloud of the Aggregate Document-Term Matrix in 2017 . . . . . 202

# 1. Dissecting Carbon Risk: An Asset Pricing View from Climate Policy

Yi Zhang

## 1.1 Introduction

As concerns about climate disasters rise, governments around the world are taking action to target the zero-net carbon emissions, a process that picked up pace after the Paris Agreement. With 195 countries participating, transition risk captures global attention and plays a prominent role in driving cross-sectional pricing performance. For instance, Figure 1.1 displays that carbon-intensive industries exposed to higher regulatory risk in Bolton et al. (2020) , including utility, gas, electricity, and transportation[1], have lower equally-weighted historical realized returns than the others.

In addition, transition risk is becoming more critical as investors increase awareness of climate issues. Figure 1.2 shows the aggregate increasing awareness towards climate issues of the Media Climate Change Concern Index (MCCC) from Ardia et al. (2020). The upward trend implies an increasing demand for higher environmental quality, potentially linking to the environmental demand channel. Acemoglu et al. (2012a) investigates a neoclassical model with household environmental awareness and explores the economic implication from a social welfare view. Bernard et al. (2022) surveys and finds that households are willing to pay to offset the negative impact of carbon emissions on climate change. This evidence implies a demand channel for transition risk propagation to cross-sectional pricing, even though few papers tackle this question in cross-sectional asset pricing. Therefore, the first goal of the paper is to explore

[1]SIC industry classification

and understand the propagation of transition risk over cross-sectional asset pricing through the demand mechanism.

The model is also motivated by recent findings on transition risk and cross-sectional stock returns directly or indirectly with rich firm performance. Bolton and Kacperczyk (2021a) documents that carbon risk exposure generates a positive price of carbon risk. Bolton and Kacperczyk (2021c) shows that firms in the tightening domestic climate policy countries tend to have lower realized stock returns. Moreover, the carbon premiums are related to the energy structure negatively. Tomar (2022) shows that even mandatory climate disclosure policy in the US triggers significant carbon emission reduction. Economically, Fried (2018a) develops a quantitative model and finds that firms in the US reduce carbon emissions with higher climate policy risk.

To my knowledge, no existing literature could reconcile all the facts within one theoretical asset pricing framework. From the view of -based asset pricing, one potential explanation is Hsu et al. (2022) linking the positive price of pollutant emission risk[2] to firm-exposed transition risk. However, the challenge is that the climate policy risk must be proportional to the firm emission for risk-averse investors in this economy unconditionally. Therefore, this assumption cannot capture the endogeneity between climate policy risk and carbon emissions and is contradictory to some aggregate empirical evidence in recent years. For example, Figure 1.3 displays the rolling average (4-year) window of the carbon emission growth in the US. It displays a current negative growth rate with the global tightening of the climate policy. Bansal et al. (2016) describes a consumption-based pricing model but focus on the temperature risk on aggregate pricing performance. Bansal et al. (2021) documents a consumption-based framework from transition risk to carbon emissions on cross-sectional asset pricing, but it focuses on the implications through temperature risk. More importantly, Hsu et al. (2022), Bansal et al. (2016) and Bansal et al. (2021) use a consumption-based model without specifying firm-specific production decisions under environmental regulatory uncertainty, a step that is needed to capture the energy structure implications for cross-sectional asset pricing.

The paper first develops a general equilibrium model to explore the propagation of climate regulatory risk within cross-sectional asset pricing. The innovations of the model rely on several perspectives. First, the model introduces the emission externalities into utility following a general utility assumption in Acemoglu et al. (2012a)[3]. Therefore, the price of transition risk is

[2]Higher pollutant emission indicates higher expected return

[3]The format of the environmental preference is also similar to Campbell and Cochrane (1999)

conditional on the sensitivity of investor awareness toward carbon emission accumulation since investors prefer a low-carbon environment. Second, by introducing a CES production function, the model introduces the energy structure as Fried (2018a) and Hassler et al. (2021) but focuses on its implication on cross-sectional pricing implications. Third, in contrast to the asset pricing model (Hsu et al. (2022)), the model endogenizes firm carbon emissions with climate policy risk and could match with different conditional moments. Therefore, the equilibrium of the energy capital market is determined according to the tightening of climate policy. This optimal decision makes it possible to track the trajectory of carbon emission growth. Finally, the model also quantitatively identifies the difference between long-run productivity risk against short-run productivity risk, compares the implication for cross-sectional pricing and matches the macroeconomic fundamental data as Massimiliano Croce (2014).

Given this structure, the model shows an innovative combined mechanism within cross-sectional asset pricing. Intuitively, the price of transition risk could be determined by demand and cash flow mechanisms through the relationship between marginal utility and cash flow. When exposed to positive transition risk, even though consumption is reduced with tightening climate policy, the gains from a low-carbon environment could offset household consumption loss and reduce marginal utility. Simultaneously, in contrast to green firms with lower climate regulatory exposure ($CRE_{i,t}$), since brown firms with higher climate regulatory exposure ($CRE_{i,t}$) are inflexible in substituting the brown capital with green, they are finally penalized with heavier climate taxation and lower market valuation. This combined effect indicates a positive relationship between cash flow and marginal utility for high transition risk-exposed firms and therefore generates lower expected returns.

In addition, when facing with positive productivity risk, the consumption gains dominate the loss from increasing brown capital demand implying a lower marginal utility. The complementarity between brown and green capita for brown (higher CRE) firms, contributes to higher profitability because of their less proportional demand for expensive green capital. This circumstance documents a negative relationship between marginal utility and realized return. Therefore, carbon-intensive firms with higher CRE tend to have higher expected returns. This prediction implies a positive price of carbon risk.

Besides the price of risk, the model has implications for the substitutability between green capital and brown capital on cross-sectional pricing. A lower elasticity of substitutability in the CES aggregation function indicates that the demand for green and brown capital is more

complementary. This property quantitatively increases the firm exposure to regulatory and long-run productivity risks. Therefore, the model predicts that higher substitutable renewable energy could reduce risk premiums. This premium reduction holds when exposed to either transition or long-run productivity risks.

With the identification between long-run risk and short-run risk, the model also shows that only the long-run risk contributes to brown (high CRE) firm with more exposure and help identify a positive price of carbon risk as Bolton and Kacperczyk (2021a) documents. In contrast, there is no significant finding when exposed to short-run risk. This result still holds when the model is recalibrated under endogenous carbon taxation in the form of Pigovian Tax. This finding indicates that a lower expected stock return for a higher CRE firm is more likely due to independent climate policy shocks.

Finally, the paper investigates the model implication with new text data from 10K filings in Baz et al. (2022a) and tests the model predictions. The measure is internally valid across industries and externally consistent around climate events. With this climate regulatory exposure, the paper documents a significantly negative correlation with scope 1 direct carbon emissions and an insignificantly negative relationship with scope 2 indirect emission. This negative finding indicates that firms persistently endogenously account for the climate policy risks and reduce carbon emissions.

The paper provides evidence of the negative price of transition risk when climate attention is high for equity market performance. A long-short portfolio of high CRE against low CRE equities shows a significantly negative abnormal alpha for cross-sectional equity pricing. The alpha ranges from 8.4% to 12% on average conditional on how to specify the control variables. The portfolio return is robust with different weighting schemes and against other double-sorting variables, including firm sales, size, investment, leverage, and text content complexity. In time series regressions, the paper finds that the robustness of abnormal alpha still holds after controlling for existing factor models. Finally, the paper shows that higher CRE during high climate sensitivity predicts future stock returns negatively. This finding confirms our prediction that higher attention to climate externalities contributes to a negative price of transition risk.

## 1.2 Related Literature

This paper relates to a series of other climate finance theoretical papers but caps implications from different perspectives. Hong et al. (2019) discusses the social welfare view of a multi-sector model with climate tipping events and optimal government intervention via limited investment. The model shows cross-sectional pricing implications with exposure to productivity risk. Hambel et al. (2021) proposes a similar structure with climate tipping event but still focus on the implications of optimal climate taxation endogenized by productivity risk. Barnett (2017) introduces a continuous production-based asset pricing framework, but emission externalities mainly work through the productivity channel. His model and empirical findings are primarily about the pricing implication on temperature risk. Jaccard et al. (2020) focuses on the emission externality through the utility preference channel and discusses the aggregate pricing implication with exposure to productivity risk. Pastor et al. (2021c) documents a preference heterogeneity model and interprets pricing inefficiency as the implication of heterogeneous ESG tastes. My model differentiates from this literature by focusing on exogenous climate policy and discussing emission externalities on cross-sectional implications through the stochastic discount factor.

My model technically relates to some literature incorporating climate inputs or integrated assessment model (IAM) into neoclassical DSGE models. Golosov et al. (2014) provides a DSGE model to analyze the optimal design of fossil fuel taxation and explore its welfare implications. Bornstein et al. (2022) describes a DSGE model with an energy sector that could match the energy supply quantitatively. Cai and Lontzek (2019) follows the long-run risk literature and incorporates EZ preferences with climate tipping events to explore the social cost of carbon.

The paper also relates to several other recent empirical findings but documents a new fact of cross-sectional pricing over climate policy. Engle et al. (2020a) documents that news-implied climate risk predicts future stock returns positively. Pastor et al. (2021a) finds that the realized green stock returns are higher than the realized brown stock returns and the expected return holds reversely. Sautner et al. (2020) constructs a text-implied firm-level climate change exposure with earning conference calls. They investigate its rationality with green innovation and derivative characteristics. My paper is distinct by focusing on the 10K filings information on climate policy concerns. Compared with earning conference calls, the firm-level compulsory disclosed risk from SEC filings is a more canonical unstructured data source in asset pricing

literature (Koijen et al. (2016),Florackis et al. (2022) etc).

The remainder of the paper is organized following. The section 1.3 starts with a simple static model to explore the main analytic properties of the theory. In section 1.4, a generalized model is introduced to match with the real data and is used to explore its prediction. Section 1.5 shows the analytic discussion of the model solution and potential properties. Section 1.6 displays the model quantitative implications. In section 1.7, the paper introduces the primary dataset on firm-level and macro-level information. The section 1.8 conducts empirical regression to test the model prediction on climate risk and equity premiums.

## 1.3 A Two-Period Model

To highlight the main intuition that how the climate regulatory shock affect marginal utility and profitability simultaneously, I include a simple two-period model with brown capital as its only product input. Figure 1.4 displays its agent-flow chart. It is a two-period model with time 0 and time 1.

$$\max_{C_0,w,B} log(C_0 - \Psi \widetilde{S}_0) + \beta E_0(log(C_1 - \Psi \widetilde{S}_1)) \tag{1}$$

$$C_0 + B_0 + w_0 P_0 = W_0 \tag{2}$$

$$C_1 = w_0 D_1 + B_0 R_f + TR_1 + \Pi_1^O \tag{3}$$

$C_t$ denotes the consumption at time t ($t = 0, 1$). $W_0$ denotes the initial wealth realized at time 0. $\omega_0$ refers to the equity holding from time 0 to 1. $\widetilde{S}_t$ refers to the carbon emissions at the atmosphere with $t = 0, 1$. $\Psi$ is the key parameter demining the household sensitivity towards the emission externalities. $B_0$ denotes the government bond holdings from time 0 to 1. $TR_1$ refers to the lump-sum taxation transferred from the government and it's defined as $TR_1 \equiv B_1 - R_f B_0 + Tax_1$ where $B_1 = 0$ and $P_1 = 0$ since no continuation values. $\Pi_1^O$ is the profits earned by energy entrepreneurs.

In this economy, I assume that the representative household initial wealth is observed and is taken as given. The climate regulatory risk is realized only in period 1, contributing to an expected loss for holding brown capital. Consistent with the literature assumption (Acemoglu et al. (2012b)), the main channel highlighting the importance of emission externality on instantaneous utility is through the additive external emission term, $\Psi \widetilde{S}_t$ (t=0, 1).

For tractability, there is only one representative firm in the model issuing equity $\omega_0 = 1$, and the only feasible capital for this aggregate production sector is in the form of brown capital, $F_t$. A potential assumption implies that representative household doesn't intervene the firm brown capital usage is due to the information asymmetry between shareholder and CEO. Implicitly, the inputs of green capital and labor are either predetermined before 1 or supplied inelastically with $L_1 = 1$. Given these assumptions, the aggregate production sector is:

$$Y_t = A_t F_t^{\alpha} \text{ for } t = 0, 1 \tag{4}$$

$A_t$ stands for the aggregate uncertainty of the economy. $F_t$ refers to the brown capital with $\alpha \in (0, 1)$. Additionally, the productivity shocks, $A_t$, are realized at the starting points of $0, 1$. Within this two-period structure, the representative firm is maximizing market valuation given the realized aggregate shocks at 0:

$$V_0 = \max_{F_1} D_0 + E_0(M_{0,1} D_1) \tag{5}$$

where the dividend payoffs are:

$$D_t = \begin{cases} D_0 & \text{if } t = 0 \\ Y_1 - Tax_1 - P_1 F_1 & \text{if } t = 1 \end{cases} \tag{6}$$

where $Tax_t$ denotes the market climate regulatory uncertainty that is equivalently realized in the form of the carbon taxation with $Tax_1 = \tau_1 \phi_1 F_1$. The climate regulatory policy, $\tau_1$, is realized at time $t = 1$ and independent with the productivity risk, $A_1$.

The brown capital is supplied by an energy entrepreneur as a representative oil inventory holder without access to equity finance. The target of the energy firm is to maximize its valuation through brown capital supply at time 1. Consistent with Bornstein et al. (2022), the supply of the brown capital is linear with a convex cost function in G and proportional to the average fossil fuel capital reserve. The convex cost function could be interpreted as an expense of energy extraction. Therefore, the energy sector profits follow:

$$\Pi_t^o = \max_{J_t} P_t J_t - G(J_t; \bar{R}), \text{ where t=1} \tag{7}$$

$\bar{R}$ stands for the energy reserve on average and the paper assumes that it's provided without exhaustive resource constraint binding ($J > 0$); $G(J_t; \bar{R})$ refers to a convex inventory cost function for energy. For simplicity, I assume the cost function as $G(J; \bar{R}) = \frac{\pi_1}{2}(\frac{J}{\bar{R}})^2\bar{R}$. For the carbon block, the carbon emission, $CE_t$, is assumed to be proportional to the use of the brown capital with sensitivity parameter, $\phi_1$.

$$CE_t = \phi_1 F_t \text{ and } \widetilde{S}_t = CE_t, t = 0, 1 \tag{8}$$

Combined with the representative household budget constraint, we get good market clearing condition, $C_1 = Y_1 - G(J_1; \bar{R})$ and the environmental demand adjusted consumption:

$$\tilde{C}_1 = A_1 F_1^{\alpha} - G(F_1; \bar{R}) - \Psi\phi_1 F_1 \tag{9}$$

The adjusted environmental consumption, $\widetilde{C}_t$, illustrates two channels that transition risk propagates into the stochastic discount factor. One is consumption response, and the other is the low-carbon preference. The proposition below first shows the response of economic fundamentals to the aggregate shocks with realized carbon taxation:

**Proposition 1** *Given the settings above, I explore the properties of cash flow and marginal utility at time 1 respectively:*
*1. the brown capital usage is negatively correlated with the realized carbon taxation:* $\frac{\partial F_1}{\partial \tau} < 0$, *where a higher climate policy tax induces a lower usage of the brown capital in the economy and lower carbon emission. A higher productivity risk implies higher carbon emissions since* $\frac{\partial F_1}{\partial a} > 0$[4].
*2. When preference on environmental quality is higher (lower) enough,* $\Psi > \tau$ *(*$\Psi < \tau$*), a higher (lower) carbon price contributes to a higher (lower) environmental adjusted consumption as,* $\frac{\partial \widetilde{C}_1}{\partial \tau_1} > 0$. *The marginal utility is correspondingly lower (higher),* $\frac{\partial M_{0,1}}{\partial \tau_1} < 0$ *where* $M_{o,1} = \left(\frac{\widetilde{C}_1}{\widetilde{C}_0}\right)^{-1}$.

Proposition 1 1) proves realized carbon taxation's effect on firm performance at time 1. When realized carbon taxation at time 1 is higher, the economy reduces its usage of the brown capital for production and generates less carbon emissions due to higher brown penalization. However,

[4]The two-period model also shows the tax-inclusive price of energy. Its tax-inclusive price increases with a higher realized carbon taxation: $\frac{\partial P^*}{\partial \tau_1} > 0$. This is consistent with generalized model in the following section

a higher TFP at time 1 induces active production performance while contributing to higher brown capital with more carbon emission. Proposition 1 1) shows the economic response to realized $\tau_1$ and $a$, and these are helpful for us to identify the signs of premium with fundamentals at steady states in proportion 2.

The second bullet point 2) in Proposition 1 shows how carbon taxation could affect the marginal utility through the low-carbon preference channel. The change of marginal utility can be decomposed into two parts. The lower output drives down the adjusted consumption, but improved environmental quality offsets the reduction with a given preference, $\Psi$. Therefore, when $\Psi$ is at a relatively higher (lower) level, the offset effect could dominate the adjusted consumption and contributes to a lower (higher) marginal utility.

Overall, proposition 1 shows how carbon taxation could jointly affect the marginal utility and cash flow. Conditional on a high (low) low-carbon preference, $\Psi$, the response of marginal utility is significant in determining the signs for the price of transition risk. In an economy with uncertainty at time 0 for equity premiums, follow the Euler equation in this two-period model where $E_0(M_{0,1}R_1) = 1$. The cost of capital is resolved with Taylor expansion analytically in Proposition 2.

**Proposition 2** *Assume that both* $\tilde{a} \equiv log(A_1) - \mu_a \sim \mathcal{N}(0, \sigma_a)$ *and* $\tilde{\tau} \equiv log(\tau_1) - \mu_\tau \sim \mathcal{N}(0, \sigma_\tau)$ *with i.i.d. Given the two-period setting and appropriate support for* $\tilde{C}_1 > 0$*, the expected excess return of the aggregate market can be decomposed into as :*

$$\begin{aligned} E_0(R_1) - R_f &= -R_f Cov(M_1, R_1) \\ &\approx -R_f \beta \kappa_m \kappa_d Cov(\kappa_{a,m}\tilde{a} + \kappa_{\tau,m}\tilde{\tau}, \kappa_{a,d}\tilde{a} + \kappa_{\tau,d}\tilde{\tau}) \\ &= -\underbrace{R_f \beta \kappa_m \kappa_d \kappa_{a,m} \kappa_{a,d} \sigma_a^2}_{\textit{productivity risk premium}} - \underbrace{R_f \beta \kappa_m \kappa_d \kappa_{\tau,m} \kappa_{\tau,d} \sigma_\tau^2}_{\textit{climate regulatory risk premium}} \end{aligned} \tag{10}$$

*where* $\kappa_m = (\frac{C_1^{ss}}{C_0})^{-1} > 0$*,* $\kappa_d = \frac{D_1^{ss}}{P_0} > 0$*,* $\kappa_{a,d} > 0$ *and* $\tilde{a}, \tilde{\tau}$ *refer to the demeaned linear states,* $\kappa_{\tau,d} < 0$*.*

$$\kappa_{\tau,m} = \frac{(\Psi - e^{\mu_\tau})\phi_1(\frac{\partial F_1}{\partial \tau})_{ss}}{\tilde{C}_1^{ss}}, \kappa_{a,m} = -\frac{e^{\mu_a}(F_1^{ss})^\alpha + (e^{\mu_\tau} - \Psi)\phi_1(\frac{\partial F_1}{\partial a})_{ss}}{\tilde{C}_{ss}^1} \tag{11}$$

*since* $(\frac{\partial F_1}{\tau})_{ss} < 0$ *and* $(\frac{\partial F_1}{\partial a})_{ss} > 0$ *from proposition 1.*

*1). As* $\Psi$ *increase, the expected climate regulatory risk premium tends to be negative when the investor environmental sensitivity is larger than the expected carbon price,* $\Psi > e^{\mu_\tau}$*.*

*2). When* $\Psi < e^{\mu_\tau} + \frac{e^{\mu_a}(F_1^{ss})^\alpha}{\phi_1(\frac{\partial F_1}{\partial a})_{ss}}$ *where* $\kappa_{a,m} < 0$*, the systematic risk premium tends to be positive; Otherwise, the systematic risk premium is negative.*

where $R_1 \equiv \frac{D_1}{P_0}$ $(P_1 = 0)$. The correlation between marginal utility and cash flow could be either positive or negative conditional on the environmental sensitivity parameter, $\Psi$. The intuition is that when the representative agents are sensitive enough towards the low-carbon environment ($\Psi$ is higher than $e^{\mu_\tau}$), an intensified climate policy could be compensated for lower expected return since investors require environmental improvement. As for productivity risk, $A_1$, when the economy is exposed to the productivity shock, as long as the investors are not extremely sensitive above some critical value ($e^{\mu_\tau} + \frac{e^{\mu_a}(F_1^{ss})^\alpha}{\phi_1(\frac{\partial F_1}{\partial a})_{ss}}$), the expected return tends to be positive and risk compensation story dominates the loss of low-carbon environment, and marginal utility becomes smaller in response to positive, $A_1$. The appendix details a Taylor expansion for the premium decomposition at the steady states.

Overall, proposition 2 shows that the relationship between cash flow and marginal utility is jointly determined by exogenous shocks and environmental preferences with uncertainty. A positive (negative) risk premium implies that the marginal utility of the representative household is negatively (positively) correlated with the cash flow. In a more general setting below, a multi-sector model is introduced to understand the driving forces for both prices of transition and productivity risk.

## 1.4 Dynamic General Equilibrium Model

The paper constructs a two-sector general equilibrium model to explore the quantitative relationship between climate action and cross-sectional stock returns at the industry level. Figure 1.5 presents the agents in the economy. In the model, the firm operates through green capital, brown capital, and labor management. The carbon emission is introduced with aggregate generalized externalities, which could affect the economics through two channels: the long-run growth and the representative household preference. The representative household has an Epstein-Zin preference for high consumption and low-emission environment. The government introduces carbon emission taxation, independent of the productivity risk first in a decentralized economy.

### 1.4.1 Household

The representative household in our economy holds all the feasible capital resources and labor sources. It has an Epstein-Zin preference in the economy with adjusted consumption, $\widetilde{C}_t$ :

$$U_t = \max_{C_t, w_{i,t}, B_t} \{(1-\beta)\widetilde{C}_t^{1-\frac{1}{\psi}} + \beta(E_t[U_{t+1}^{1-\gamma}])^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\}^{\frac{1}{1-1/\psi}} \tag{12}$$

$$C_t + \int_{i\in\{B,G\}} \omega_{i,t} P_{i,t} di + B_t = B_{t-1} R_{f,t-1} + W_t L_t + \int_{i\in\{B,G\}} \omega_{i,t-1}(P_{i,t} + D_{i,t}) di + TR_t + \Pi_t^o \tag{13}$$

Consistent with the two-period static model,$\widetilde{C}_t := C_t - \Psi S_t$, where $S_t$ denotes the cumulative carbon emission multiplied by economic level at $t-1$, $\widetilde{S}_t \times Z_{t-1}$. Intuitively, the long-run cumulative carbon emissions contributes to an instantaneous utility for representative household given a sensitivity parameter, $\Psi$. This preference is similar to GHH preference as Colacito et al. (2018) where consumption is adjusted with labor supply. However, the preference in this paper is conditional on the dynamics of carbon emission, which is exposed to both productivity and transition risk. $TR_t$ refers to the lump-sum tax realized by the government. $\Pi_t^o$ denotes the profits from the energy sector production. $L_t$ stands for the aggregate labor supply. $\omega_{i,t}$ refers to the asset holdings in asset $i$ at time t. For simplicity, $L_t = 1$, where the labor supply in economics is inelastic. $\gamma$ is the risk aversion of the representative agent and $\psi$ is the household elasticity of intertemporal substitution. The model assumes that $\psi > \frac{1}{\gamma}$, so the agent has an early resolution of uncertainty as Bansal and Yaron (2004a).

The model assumes low-carbon preference formation based on long-run carbon emission. The justification of this structure refers to Acemoglu et al. (2012a). When representative agents sufficiently care about environmental quality, a higher production with carbon emission externality contributes to a lower realized utility. From the empirical side, Bernard et al. (2022) documents that households have willingness to pay for offsetting carbon externalities. Bretscher et al. (2020a) also documents that combining the external habit formation with recursive preference is more feasible to match with asset pricing quantitatively.

Conditional on this basic setting, the SDF of the model from Euler Equation follows:

$$M_{t,t+1} = \beta\{\frac{\widetilde{C}_{t+1}}{\widetilde{C}_t}\}^{-\frac{1}{\psi}} \Big(\frac{U_{t+1}}{E_t[U_{t+1}^{1-\gamma}]^{\frac{1}{1-\gamma}}}\Big)^{\frac{1}{\psi}-\gamma} \tag{14}$$

The derivation of the formula is included in the Appendix. In contrast to the canonical stochastic discount factor, the model now includes a non-separable low-carbon consumption ratio which is directly exposed to transition risk.

### 1.4.2 Intermediate Good Sector

There are continuous intermediate firms with 1 unit of density in the economy. Each of the firms belongs to group B, a brown sector or group G, green sector. In contrast to brown firms, the green firms have lower transition risk exposure. The production functions for both types follow Cobb-Douglas forms:

$$Y_{i,t} = (\underbrace{Z_t \Omega(\widetilde{S}_t)}_{\widetilde{A}_t} L_t^i)^{\alpha} (\mathcal{E}_t^i)^{1-\alpha}, \text{ where } i \in \{B, G\} \tag{15}$$

In order to quantify the distinct impact of systematic risk on pricing, I decompose the economic growth into different driving forces. $Z_t$ denotes the non-stationary productivity risk over the business cycle ($\Delta z_t = \Delta log(Z_t)$). $\Omega(\widetilde{S}_t)$ denotes the emission damage function. The function takes the form of $e^{-\mu \widetilde{S}_t}$ in Golosov et al. (2014). In practice, many other uncertainties might affect the propagation of emission externalities on real economics as Nordhaus and Moffat (2017) and Cai and Lontzek (2019). The procedure might follow a dynamic temperature accumulation process or belief in rare disasters (climate tipping events). However, since these economic channels usually take hundreds of years to materialize over the economy, this paper highlights a relatively short-term period of pricing performance. Therefore, the model is simplified to keep track of the carbon footprint externalities directly as Hassler et al. (2018). They document a way to unify the risk impact within the form of the exponential damage function quantitatively as in 15:

$$\Delta log(Z_t) = \mu_Z + z_t^l + \sigma_a \epsilon_{a,t} \tag{16}$$

$$z_t^l = \rho_z z_{t-1}^l + \epsilon_{z,t} \tag{17}$$

where $\epsilon_{a,t}$ denotes the transitory productivity shocks.

Hassler et al. (2021) argues the importance of the elasticity of substitution between the green capital and brown capital within a neoclassical framework. Therefore, $\mathcal{E}_t^i$ denotes the aggregate capital exploited by intermediate firm $i$ for production. It's in the form of the Dixit-Stiglitz

capital inputs for both green and brown firms, where

$$\mathcal{E}_t^i = \{(K_t^i)^{\frac{\sigma-1}{\sigma}}(1-\nu_i) + (F_{i,t})^{\frac{\sigma-1}{\sigma}}\nu_i\}^{\frac{\sigma}{\sigma-1}}, \text{where } i \in \{B, G\} \tag{18}$$

Where $\sigma$ denotes the elasticity of substitution between the green capital and energy capital. $\nu_i$ refers to the intermediate input shares. $\nu_i$ is in (0,1) for brown sector (B, high CRE) and $\nu_i = 0$ for green sector (G, low CRE). $\sigma$ denotes the elasticity of the substitution between the green capital and brown capital. The brown capital is assumed to be in the form of fossil fuel, $F_{i,t}$. Another critical assumption for the production function is that only brown (fossil fuel, $F_{i,t}$) capital produces the emission externalities. The green capital has similar dynamic pattern as physical capital and has no emission externalities.

For simplicity, the paper focuses on an equilibrium model with symmetric assumption for intermediate firms. The density for firms in green and brown sectors are denoted as $\gamma_B \equiv \int_{i\in B} di$ and $\gamma_G \equiv \int_{i\in G} di$ respectively with $\gamma_B + \gamma_G = 1$. Each firm is to maximize the firm valuation with manipulation of its green and brown capital investment:

$$V_{i,t} = \max_{K_{t+j+1}^i, F_{i,t+j+1}, I_{t+j}^i, L_{t+j}^i} E_t\{\sum_{l=0}^{+\infty} M_{t,t+j} D_{i,t+j}\} \tag{19}$$

The dividends are defined as:

$$D_{i,t} = \begin{cases} P_t^i Y_{i,t} - I_t^i - P_t^F F_{i,t} - Tax_{i,t} - W_t L_t^i, & \text{when } i = B \\ P_t^i Y_{i,t} - I_t^i - W_t L_t^i, & \text{when } i = G \end{cases} \tag{20}$$

$P_t^F$ equals the aggregate price of the brown capital (fossil fuel capital) in the energy market. $\tau_{i,t}$ represents the perceived carbon taxation in brown sector. $\tau_{B,t} = \tau_t^* p_{B,t}$, where $p_{B,t}$ represents the perceived probability that economics switch into carbon regulatory regime. $\tau^*$ refers to a consensus of potential carbon taxation rate proposed by authorities or economists like Hassler et al. (2018). For non-negative concern, the odd of probability that a federal-level climate

regulatory risk is imposed, $w_t = log(\frac{p_{B,t}}{1-p_{B,t}})$, follows an AR(1) process as:

$$w_t = (1-\rho_w)w_{ss} + \rho_w w_{t-1} + \epsilon_{w,t} \tag{21}$$

$$p_{B,t} = \frac{1}{1+e^{-w_t}} \tag{22}$$

$$Tax_{B,t} = p_{B,t}\tau^* \times \text{Carbon Emission}_{B,t} \tag{23}$$

A critical assumption in this firm value maximization problem is the timeline of production performance. In decentralized economics, firms have no acknowledges of their emission dynamics and have no technique to identify their real productivity ($Z_t$) from the observed $\widetilde{A}_t$. Therefore, they cannot internalize their emission externalities accordingly. Firms only allocate labor and capital resources in the economy while taking the aggregate economic growth, $\widetilde{A}_t$ as given. They have no capacities to identify how much their decision on brown capital could ruin economic growth negatively.

For green capital accumulation process, both sectors have the same depreciated ratio rate, $\delta$, even though in reality they might face distinct utilitization rate:

$$K^i_{t+1} = (1-\delta)K^i_t + \Phi^i(I^i_t, K^i_t), i \in \{G, B\} \tag{24}$$

$K^i_t$ denotes the green capital for firm $i$ at time t. $\Phi^i$ is the convex capital cost function. The convex cost functions are parametrized as Jermann (1998):

$$\Phi^i(I^i_t, K^i_t) := \Big[\frac{\alpha_1}{1-\xi}(\frac{I^i_t}{K^i_t})^{1-\xi} + \alpha_2\Big]K^i_t \tag{25}$$

The cost functions are symmetric with the same adjustment coefficients, $\alpha_1$ and $\alpha_2$, for $i \in \{B, G\}$ implying that high (low) exposed sector has the same investment ratio at the steady states. Following the Massimiliano Croce (2014), the capital adjustment has properties for $G(\frac{I}{K}) \equiv \frac{I}{K} - \frac{\Phi^i(I,K)}{K}$ at the steady states where $G(\bar{i}) = 0$ and $G(\bar{i})' = 1$, where $\bar{i}$ equals to $(\frac{I}{K})_{SS}$ around the steady states.

With these assumptions, firms optimize their market valuations for $K^i_{t+j+1}, F_{i,t+j+1}, I^i_{t+j}, L^i_{t+j}$ under the perfect competition. All the firms take the input and output prices as given, assuming that there is no market power within each sector. Therefore, the cross-sectional difference among firms depends on their exposure to climate regulatory risk and feasibilities to access the brown capital.

### 1.4.3 Energy Sector

In contrast to Hsu et al. (2022), the model's target is to endogenize carbon emissions with climate policy exposure. I introduce the energy market for brown capital to realize this endogeneity in the form of carbon taxation. The energy sector in this economy internalizes its brown capital supply by maximizing its firm valuation, $V_t^F$. The optimization problem of the energy entrepreneur follows Bornstein et al. (2022) with the assumption that energy firms have no access to equity financing. The rationality of the structure is to match with Hotelling rule for energy price. There is only one firm producing the energy supply in this model, even though, in reality, the market power exists with heterogeneous sectors.

In each period t, the goal of the energy entrepreneur is to maximize its firm valuation by endogenizing green capital investment ($I_t^F$), energy extraction ($\theta_t$), exploration ($X_{t+1}$) and reserving ($R_{t+1}$), etc.

$$V_t^F = \max_{X_{t+j+1}, R_{t+j+1}, I_{t+j}^F, \theta_{t+j}} E_t\{\sum_{l=0}^{+\infty} M_{t,t+j}\Pi_{t+j}^o\} \tag{26}$$

$\Pi_{t+j}^o$ denotes the energy entrepreneur profit at time $t+j$. It's defined as the net of energy extraction profits:

$$\Pi_t^o = \underbrace{P_t\theta_t R_t}_{\text{Energy Revenue}} - \underbrace{I_t^F}_{\text{Energy Investment}} - \underbrace{\Phi^F(\theta_t)R_t}_{\text{Energy Cost}} \tag{27}$$

$\Phi^F(\theta_t)$ stands for a convex energy extraction cost to capture the real extraction behavior of the energy supply sector. $R_t$ is defined as the energy reserves. To match with the energy sector more, the model also includes two other types of capital as Bornstein et al. (2022) to capture the time variation of the energy supply sectors over the steady states in the long run. The first capital is the energy reserves with dynamics as follows:

$$R_{t+1} = (1-\theta_t)R_t + \lambda X_t \tag{28}$$

Intuitively, the energy reserve is defined as the sum of new exploration, extraction deduction and last energy reserves, where $X_t$ refers to the exploration capital. Intuitively, the entrepreneur holds the exploration capital to set up new oil wells. The dynamics of the exploration capital follows:

$$X_{t+1} = (1-\lambda)X_t + (I_t^F)^{\xi}(L_t^a)^{1-\xi} \tag{29}$$

$L_t^a$ refers to the accessible land resources for energy reserve exploration. To make the economy over the balanced growth path as Colacito et al. (2018), the model assumes that the normalized land resources follow a dynamic stationary process:

$$log(\frac{L_t^a}{Z_{t-1}}) = (1-\rho_L)L_{ss}^a + \rho_L log(\frac{L_{t-1}^a}{Z_{t-2}}) + \rho_a a_t + \epsilon_{l,t} \tag{30}$$

Intuitively, the term $\rho_a a$ in $L_t^a$ dynamics could refer to Ready (2018). It could capture the feedback effect of economic growth on the energy supply.

### 1.4.4 Final Good Sector

A unique final consumption good, $Y_t$, is made from continuous intermediate goods. The market for the final is assumed to be in perfect competition with the CES aggregate production function:

$$Y_t = \big[\int_i Y_{i,t}^{\frac{\sigma_y-1}{\sigma_y}} di\big]^{\frac{\sigma_y}{\sigma_y-1}} \tag{31}$$

Under the symmetric setting for two sectors (high or low), the production aggregator is simplified as:

$$Y_t = \big[\gamma_G Y_{G,t}^{\frac{\sigma_y-1}{\sigma_y}} + \gamma_B Y_{B,t}^{\frac{\sigma_y-1}{\sigma_y}}\big]^{\frac{\sigma_y}{\sigma_y-1}}, \text{where } \gamma_G + \gamma_B = 1 \tag{32}$$

The target of the final good sector is to optimize its profits given the input price of $P_t^B$ and $P_t^G$:

$$\max_{Y_t^G, Y_t^B} \Big[P_t Y_t - \int P_t^i Y_t^i di\Big] \tag{33}$$

Because of homogeneous and perfect competition assumptions, the optimal allocation is obtained with zero profit. The final good price is normalized to $P_t \equiv 1$. The optimal condition is given by:

$$\frac{P_t^G}{P_t^B} = \Big(\frac{Y_t^G}{Y_t^B}\Big)^{-\frac{1}{\sigma_y}} \tag{34}$$

The price of the final good follows $P_t = \big[\gamma_G (P_t^G)^{1-\sigma_y} + \gamma_B (P_t^B)^{1-\sigma_y}\big]^{\frac{1}{1-\sigma_y}}$.

### 1.4.5 Carbon Cycle

The innovative part of the paper is to incorporate the canonical (IAM) integrated assessment model (Nordhaus (1992)) into the long-run risk asset pricing framework. The carbon emission production is assumed to be linear in the brown capital as Hassler et al. (2018):

$$e_{B,t} = \varphi_1 \frac{F_{B,t}}{Z_{t-1}} \tag{35}$$

Only the high CRE exposed (brown) sector generates carbon emissions at $e_t \equiv e_{B,t}\gamma_B$. The aggregate carbon emission is defined as a non-linear aggregator:

$$CE_t = \Big(\int_{i\in B} e_{i,t}di\Big)^{\zeta}\Big(\frac{\Gamma_t}{Z_{t-1}}\Big)^{1-\zeta} = (\gamma_B e_{B,t})^{\zeta}\Big(\frac{\Gamma_t}{Z_{t-1}}\Big)^{1-\zeta}, \text{where } \zeta \in \big(0,1\big) \tag{36}$$

where $\Gamma_t$ denotes the other climate uncertainty that potentially affects the carbon dynamics. To guarantee the economy over the balanced growth path, the logarithmic $\frac{\Gamma_t}{Z_{t-1}}$ follows an AR(1) process. $\Gamma_t$ tracks the carbon emission multipliers at atmosphere. Intuitively, this variable, $\Gamma_t$, measures alternative channels affecting the carbon accumulation. Scientifically, it could denote the uncertainty of superficial sea carbon absorption. Cai and Lontzek (2019) also documents that distinct earth layers could have various effects on atmosphere carbon concentration. Given this setting, the accumulation process of aggregate carbon emissions in the atmosphere is defined as AR(1) process:

$$\widetilde{S}_{t+1} = \rho_S \widetilde{S}_t + CE_{t+1} \tag{37}$$

### 1.4.6 Government

The government in the decentralized economy mainly transfers surplus of climate policy taxation and government debt issuance to the representative household. It has no social optimal view to realize the first best. Therefore, its budget constraint are as follows:

$$TR_t := Tax_{B,t} + B_t - R_{f,t}B_{t-1} \tag{38}$$

where $TR_t$ is defined as a lump-sum transfer to the households. It might be controversial to assume this form of federal-level carbon taxation. However, there is quantitative evidence

indicating that implications of US carbon taxation exist as Tomar (2022) or macro fundamentals are affected historically (Fried (2018b)) by aggregate carbon tax rate. In consistence with the previous equation, 22 and 23, the intensity of expected carbon price is defined as:

$$\tau_t \equiv \tau^* \frac{e^{w_t}}{1 + e^{w_t}} \tag{39}$$

where $\tau^*$ is interpreted as the aggregate expected carbon price from an official environmental report (IPCC) or rigorous academic research such as 45.5 dollars per tone of carbon emissions in some regions (Hassler et al. (2018)).

### 1.4.7 Market Equilibrium

According to the budget constraint of the representative agent and the profit definition for different sectors, the good market clearing is easily derived as the following[5]:

$$Y_t = C_t + I_t + \Psi_t^R, \text{where } I_t^{GB} := \int_i I_{i,t} di = \gamma_G I_t^G + \gamma_B I_t^B \tag{40}$$

$\Psi_t^R$ denotes the extraction cost of the energy sector, $\psi(\theta_t)R_t$. $I_t$ refers to the aggregate investment of the economics. The aggregate investment includes green investments in brown, green, and energy sectors:

$$I_t \equiv I_t^{GB} + I_t^o \tag{41}$$

Since the economy assumes that there are two sectors, each under a firm symmetric setting. Therefore, the energy market clearing condition could be derived as:

$$\int_{i \in B} F_{i,t} di = \gamma_B F_t = \theta_t R_t \tag{42}$$

Plus the market clearing condition for labor, $L_t = \int_i L_t^i di = 1$. Concerning the asset market, there is a market clearing condition for holding equities $\omega_{i,t} = \gamma_i$ where $i \in \{B, G\}$ and government bond market is at zero net supply.

[5]Details of the model are included in the Appendix 1.10.5

## 1.5 Model Solution

### 1.5.1 Investment and Equity Return

Solve the dynamic valuation problem for individual firms of the two sectors. Based on the Bellman equation for the firm in the green sector, the first order condition for individual firm green capital ($K^G_{t+1}$) are the following:

$$Q^G_t = E_t\big[M_{t,t+1}\big(\frac{P^G_{t+1}\alpha Y^G_{t+1}}{K^G_{t+1}} + Q^G_{t+1}[(1-\delta) + \Phi^G_{t+1} - (\Phi^G_{t+1})'\frac{I^G_{t+1}}{K^G_{t+1}}]\big)\big] \tag{43}$$

Firms in the brown sector also make the optimal decision for green capital, $K^B_{t+1}$:

$$Q^B_t = E_t\big[M_{t,t+1}\big(\frac{P^B_{t+1}\alpha(1-\nu)Y^B_{t+1}}{\mathcal{E}_{t+1}^{1-\frac{1}{\sigma}}(K^B_{t+1})^{\frac{1}{\sigma}}} + Q^B_{t+1}[(1-\delta) + \Phi^B_{t+1} - (\Phi^B_{t+1})'\frac{I^B_{t+1}}{K^B_{t+1}}]\big)\big] \tag{44}$$

Based on these first-order conditions, the investment returns in $B$ and $G$ sectors are derived:

$$\begin{aligned} R^B_{I,t+1} &\equiv \frac{1}{Q^B_t}\big(\frac{P^B_{t+1}\alpha(1-\nu)Y^B_{t+1}}{\mathcal{E}_{t+1}^{1-\frac{1}{\sigma}}(K^B_{t+1})^{\frac{1}{\sigma}}} + Q^B_{t+1}[(1-\delta) + \Phi^B_{t+1} - (\Phi^B_{t+1})'\frac{I^B_{t+1}}{K^B_{t+1}}]\big) \\ R^G_{I,t+1} &\equiv \frac{1}{Q^G_t}\big(\frac{P^G_{t+1}\alpha Y^G_{t+1}}{K^G_{t+1}} + Q^G_{t+1}[(1-\delta) + \Phi^G_{t+1} - (\Phi^G_{t+1})'\frac{I^G_{t+1}}{K^G_{t+1}}]\big) \end{aligned} \tag{45}$$

Since the model does not explicitly identify the difference between physical capital and green capital, the $R^G_{t+1}$ is similar to the physical investment return in Massimiliano Croce (2014). The equation 45 also sheds light on the importance of capital complementarity for investment return between brown firms and green firms. When the complementarity between green and brown capital strengthens with lower $\sigma$ ($\sigma \to 0$), the aggregate capital, $\mathcal{E}_t^{1-\frac{1}{\sigma}}$, could become a negative force pushing down the brown firm investment return as transition risk intensifies. In contrast, as $\sigma \to 1^+$, the $\mathcal{E}_t$ could drive investment return positively as transition risk intensifies.

As for the equity return, the equation 19 is rewritten in the Bellman Equation and The price of market risk is derived, following production-based asset pricing in 47. form:

$$V_{i,t} = D_{i,t} + E_t[M_{t,t+j}V_{i,t+1}], \tag{46}$$

Then, I rearrange the terms and define the ex-dividend price: $P^S_{i,t} := V_{i,t} - D_{i,t}$. Then, the Bellman equations 46 can be transferred into:

$$E_t[M_{t,t+1}\frac{V_{i,t+1}}{V_{i,t} - D_{i,t}}] = 1 \rightarrow E_t[M_{t,t+1}R^i_{t,t+1}] = 1, i \in \big(G, B\big) \tag{47}$$

The equity premium refers to the: $R_{i,t,t+1} = \frac{P^S_{i,t+1}+D_{i,t+1}}{P^S_t}$ for green and brown firms. Then, the Euler equation can be transferred into:

$$E_t[R^i_{t,t+1}] - R_{f,t+1} = -Cov_t(R_{i,t+1}, M_{t,t+1})R_f \tag{48}$$

The expression shows that the cross-sectional pricing implication depends on the negativity or positivity of the marginal utility with the cash flow channel (realized value).

### 1.5.2 Climate Adjusted SDF

Since I include the emission externality into the preference. A generalized recursive preference including carbon emission externality is demanding as Bretscher et al. (2020b) and Swanson (2018), accounting for the externalities from the production emission. Follow the Bansal and Yaron (2004b), redefine the discount rate channel via redefining the wealth accumulation process: $\widetilde{W}_{t+1} := (\widetilde{W}_t - \widetilde{C}_t)R^W_{t,t+1}$, where $R^W_{t,t+1}$ denotes the overall returns of the asset holding for the household[6]. $\widetilde{W}_t$ denotes the carbon adjusted household total wealth as: $\widetilde{W}_{t+1} := a_t R_{t,t+1} + B_t R^f_{t,t+1} + TR_{t+1} + \Pi^o_{t+1} + W_{t+1}L_{t+1} - \Psi S_{t+1}$. Since the model assumes a non-separable additive utility function with emission externalities as Bretscher et al. (2020a), the approximate log SDF at steady state ($C_{ss}$, $S_{ss}$) is derived as following 1.10.5:

$$m_{t,t+1} \approx \frac{1-\gamma}{1-\frac{1}{\psi}}log(\beta) - \underbrace{\frac{1}{\psi}\kappa_{c,ss}\frac{1-\gamma}{1-\frac{1}{\psi}}\Delta c_{t+1}}_{\text{Consumption Sensitivity (+)}} + \underbrace{\frac{\Psi}{\psi}\kappa_{s,ss}\frac{1-\gamma}{1-\frac{1}{\psi}}\Delta s_{t+1}}_{\text{Emission Sensitivity (+)}} + \underbrace{\frac{\frac{1}{\psi}-\gamma}{1-\frac{1}{\psi}}r^W_{t,t+1}}_{\text{Wealth Level Effect (-)}} \tag{49}$$

where $s_t \equiv log(S_t)$ , $\kappa_{c,ss} = \frac{C_{ss}}{\widetilde{C}_{ss}}$ and $\kappa_{s,ss} = \frac{S_{ss}}{\widetilde{C}_{ss}}$. $C_{ss}$ refers to the aggregate consumption at a steady state. $\widetilde{C}_{ss}$ denotes the non-negative environmental adjusted consumption at a steady

[6]The aggregate carbon adjusted wealth return is therefore defined as $R^W_{t,t+1} = \frac{a_t R_{t,t+1} + B_t R^f_{t,t+1} + TR_{t+1} + \Pi^o_t + W_{t+1}L_{t+1} - \Psi S_{t+1}}{a_t + B_t}$, where $a_t$ is defined as total asset from budget constraint: $a_t \equiv \int_i w_i P_{i,t} di$ and $R_{t,t+1} = \frac{\int_i w_{i,t+1}(P_{i,t+1}+D_{i,t+1})}{\int_i w_{i,t+1}P_{i,t}}$

state. Consistent with the literature on early resolution of long-run business risk, $\psi > 1 > \frac{1}{\gamma}$, the agent dislikes the shocks to long-run growth over utility continuation.

After decomposing the SDF in 49, the approximated log SDF shows three different driving forces to push the discount rates while transition risks intensify. When a positive climate policy shock occurs, consumption growth is prone to decrease because of inflexibility in substituting the brown capital. Consumption growth could contribute to a negative effect over SDF given $\gamma > 1$. Meanwhile, as carbon emissions are mitigated in response to the climate policy intensification, with investor risk aversion ($\gamma > 1$), emission growth tends to be negative, positively affecting the SDF. Simultaneously, an intensified transition risk reduces the carbon emission level, $S_{t+1}$, contributing to a higher environmental adjusted return according to 1.10.5. Since $\psi > 1 > \frac{1}{\gamma}$, this wealth effect tends to drive the SDF downward.

## 1.6 Quantitative Implication

Given the previous section's analytic insight, I solve the model with careful quantitative design. The target sample for moment evaluation is from 1996 to 2019 since the model tries to capture the endogeneity of carbon emission growth concerning climate policy risk. It's not until the recent decades that US aggregate carbon growth shows an average negative trend. The fundamental variables involve aggregate consumption and output from the US with series from the BEA (Burea of Economic Analysis). The series of carbon emission growth, $\Delta ca_t$, and oil extractions are from EIA public website. The energy extraction rate, $\theta_t$, is approximated with oil production divided by oil reserves at an annual frequency.

### 1.6.1 Calibration

The economy is calibrated with two sectors (green and brown). Each sector has homogeneous firms respectively with density $\gamma_B$ and $\gamma_G$ ($\gamma_B + \gamma_G = 1$). The model is calibrated around the steady state through Dynare++ with perturbation order=2. The key determinants of the model solution are the brown capital ($F_{SS}$) and energy extraction rate ($\theta_{SS}$). The detailed procedure and some analytic properties for steady states calculation are introduced in the Appendix 1.18.

Concerning the tuning parameters, the calibrated parameters are carefully designed to match macro fundamentals as in Table 1.3. The long-run growth, $\mu_Z$, is set to match the US real growth of about 1.8%. The other parameters are first calibrated with reference to the paper of Fried (2018a), Bornstein et al. (2022) and Massimiliano Croce (2014). The labor share in the production function is set at $\alpha = 0.285$, which is a similar parameter as Fried (2018a). More importantly, as Fried (2018a) and Hassler et al. (2021) documents the importance of Leontief' property for complementary capital inputs, the model calibration for $\sigma$ and $\sigma_y$ are 0.03 and 0.08, respectively to guarantee that the brown input is not easily substituted with green input. Besides these referred parameters, some other parameters are calibrated, targeting at matching the aggregate moments on carbon emissions and energy extraction. $\rho_\Gamma$ and $\Gamma_{ss}$ are calibrated to match with the negative carbon emission growth ($-0.08\%$). Intuitively, this is because $\rho_\Gamma$ determines the long-term impact of the exogenous forces driving carbon emission growth. $\pi_1$ identifies the energy sector production ratio with an unconditional moment 17.6% according to the data, as a higher $\pi_1$ production sensitivity implies a higher brown capital cost for brown capital demanding firms.

Table 1.4 summarizes the aggregate moment matching and asset pricing performance with the given parameters. For the aggregate moments, the model predictions on macroeconomic fundamentals, including consumption growth (1.81%), output growth (1.84%), energy production rate (22.81%) and consumption volatility (2.48%), match with the aggregate real data on consumption growth (2.22%), output growth (2.06%), energy extraction (17.60%) and consumption growth volatility (2.15%) quite well. Notably, the model generates a negative carbon emission growth (-0.07%) compared with conditional real data (-0.08%) from 1996 to 2019. Compared to alternative frameworks like Hsu et al. (2022), this model provides a way to match the conditional negative moments rather than the unconditional through endogenous carbon emission.

For asset performance, in order to match the risk-free rate and stock return jointly, the model requires to consider both household and production economic characteristics jointly. The intertemporal elasticity of substitution is $\psi = 3$, and the risk aversion is set to $\gamma = 20$, since these parameters mostly determine the exposure to exogenous risk through discount rate channel for risk-free rate. The volatility ratio between long-run and short run-risk is calibrated to be comparable with Massimiliano Croce (2014) around ten to match with the real data, and the implication of their levels matches with the risk-free rate. $\rho_w$ and $\sigma_w$ are also attentionally calibrated to determine the cross-sectional equity premium and carbon emission growth, as the

persistence and economic magnitude of $\sigma_w$ determine the brown firm climate regulatory exposure relative to green firms through cash flows. $\sigma_w$ is calibrated regarding the non-negativity of brown equity premium and risk-free rate jointly. With these overall settings, the model predicts a 5.32% equity premium with 12.92% market return volatility that are comparable with aggregate market performance. More importantly, the model drives the unobserved moment predictions on cross-sectional premiums that's more close to reality as 1.1. The model in Table 1.4 predicts a positive equity premium (4.29%) between green and brown firms. The brown firms generate a 3.75% equity premium, while the green firms have an 8.04% equity premium. Importantly, due to the high volatility of transition risk, the simulated volatility of brown firms is much higher (25.07%) than the green firms (9.39%).

Finally, Table 1.4 conducts sensitivity analysis with alternative parameter setting. Compared with the benchmark model in column (2), the premium between brown and green firms is much smaller. This alternative setting is quantitatively contradictory to the cumulative premium in Figure 1.1. When the elasticity of substitution between brown and green capital is much higher, as in column (3), it is more flexible for brown firms to hedge against the policy risk via green capital. Therefore, the impact of transition risk over both firms is much smaller in column (3) through cash flow channel, contributing to a lower premium. In column (4), the premium between brown and green firms is smaller due to a less positive correlation between brown cash flow and marginal utility when exposed to transition risk in the lower $\Psi$ ($\Psi = 0.5$) case. In column (5), with a lower transition risk volatility, the premium between green and brown firms is of smaller economic magnitude and close to 0 since both firms are affected jointly by transition risk with a smaller economic magnitude.

### 1.6.2 Model Prediction

Given the predefined calibration parameters, the paper explores the properties of firm characteristics in response to an intensive carbon tax rate at the steady state. Figure A.1.19 in appendix shows the relationship between climate policy intensity and fundamentals at the steady states. As climate policy intensifies, the share of brown firms over the economy is getting smaller and the consumption adjusted by low-carbon preference is going up over balanced growth path.

Consistent with the two-period static case, the model with heterogeneous sectors indicates a

similar decomposition conceptually:

$$E(R^i - R^f) \approx \underbrace{\beta_i^1 \lambda^1}_{\text{Productivity Risk}} + \underbrace{\beta_i^2 \lambda^2}_{\text{Transition Risk}} \tag{50}$$

The propagation of the transition risk and productivity risk could work jointly through the marginal utility and cash flows as section 1.5. Then, it is important to identify the signs of the price of risks respectively and account for its consistency with empirical findings on carbon risk. The equation 50 refers to an approximated decomposition of the expected return over productivity risk premium and climate regulatory risk premiums.

**Climate Regulatory Risk**

Figure 1.6 shows the impulse response function to climate regulatory risk. The left figures show pricing performance for higher low-carbon preference (high $\Psi$). When a transition risk is realized, investors demand a risk compensation for higher risk exposure but the environmental low-carbon demand offsets this risk channel. Overall, the stochastic discount factor becomes smaller, as in Panel A. Meanwhile, since the realized climate regulatory risk increases the brown capital cost, the profitability of brown firms reduces conditional on a lower substitutability between brown and green capital. The lower elasticity of substitution between green and brown capital implies a lower capital shift from brown to green as Appendix Figure 1.16. According to the Euler Equation transformation:

$$E_t(R^B_{t,t+1} - R^G_{t,t+1}) = -R^f_{t,t+1} Cov_t(R^B_{t,t+1} - R^G_{t,t+1}, M_{t,t+1}) \tag{51}$$

The realized premium between brown and green firms becomes negative, and the SDF becomes smaller when exposed to transition risk. The price of transition risk turns out to be negative as 51.

On the contrary, when the preference for low-carbon economics is small ($\Psi = 0$), investors still require risk compensations for holding high CRE exposed (brown) assets as the right graphs in Figure 1.6. In this case, the price of transition risk is positive because of the negative correlation between marginal utility and realized premium.

Figure 1.7 shows the fundamental response to the realized climate regulatory shock. As climate policy tightens, the brown capital price (tax inclusive) increases in Panel A. The overall eco-

nomic output decreases in spite of a small economic magnitude in Panel B. Both the brown and green firms go through divestment in Panel C and D. This economic contraction from positive transition shock further reduces the demand for brown capital. It contributes to less investment in the energy sector, as in Panel G. In contrast, the carbon accumulation growth decreases with higher environmental penalization in Panel H. This transition shock also reduces the supply from the energy sector, as in Panel I. Overall, conditional on the low substitutability between brown and green capital, the transition shock moderates the overall economy contributing to a negative price of transition risk. This prediction is testable once the climate regulatory exposure is defined in the following section.

**Productivity Risk**

In Figure 1.8, the model shows the quantitative implication over long-run productivity risk (Panel A-C) and short-run productivity risk (Panel D-F). When exposed to the long-run productivity risk, the stochastic discount factor in Panel A becomes smaller through the risk compensation channel with higher aggregate consumption. The offsetting effect from low-carbon preference still exists, but because of the magnitude, it is dominated by the risk compensation mechanism. Meanwhile, conditional on lower substitutability between brown and green capital, brown (higher CRE) firms are prone to reduce overall expenditure using brown capital. This operation corresponds to a higher realized market valuation as in Panel B, and the equity premiums between brown and green firms are higher in Panel C. Overall, when exposed to the long-run productivity risk, the model predicts a positive price of long-run productivity risk and a positive price of carbon risk. This prediction is consistent with Bolton and Kacperczyk (2021a). In contrast, when facing a short-run positive shock, the stochastic discount factor is smaller because of improved consumption through risk compensation, even though the dominated low carbon preference effect still coexists. Concerning the cash flow, the short-run productivity risk has a trivial economic effect to differentiate the two sectors in Panel E and F.

Figure 1.9 explores the fundamental response to the long-run productivity risk. As the long-run productivity shock realizes positively at time 1, the tax-inclusive energy price increases in Panel A due to higher brown capital demand. The overall economy is positively affected by higher output (Panel B) for both sectors (Panel C and Panel D). Therefore, the demand for physical and brown capital has been amplified. The investments go up in Panel E, and Panel

F. Energy sector investment also increases with higher production demand, as in Panel G. The accumulative carbon growth also increases in Panel H, even though it has a trivial economic magnitude. These findings imply that when the firms are exposed to the long-run productivity risk, brown firms earn higher expected returns and display a positive price of carbon risk as Bolton and Kacperczyk (2021a).

**Elasticity of Substitution**

Hassler et al. (2021) exploits a CES production function and finds that the elasticity of substitution between brown and green capital is close to Leontief's property. In this paper, I extend it to asset pricing literature and explore the importance of brown and green capital complementarity over cross-sectional pricing.

Figure 1.10 Panel A explores the equity premiums with different elasticity of substitution between green capital and brown capital. When the economy is exposed to climate regulatory shock, the premiums between brown (high CRE) and green (low CRE) firms under low elasticity are more negative than the premiums with high elasticity. This difference is because when firms cope with climate regulatory risk, the ones with higher capital elasticity are more flexible in substituting the brown with green ones[7]. Firms, therefore, are less exposed to climate regulatory risk. This finding generates an implication on cross-sectional asset pricing that is consistent with Bolton and Kacperczyk (2021c). Carbon premiums are lower when there is more renewable energy substitute (higher elasticity) in the economy.

Figure 1.10 Panel B explore the equity premiums in response to the long-run productivity risk. At the steady state, the brown contribution to the production increase is higher as the complementarity between brown and green capital strengthens:

$$\frac{\partial\left(\frac{\partial lnY_t^B}{\partial lnF_t}\right)}{\partial\sigma}|_{ss} < 0, \text{ with } \sigma > 0^8 \tag{52}$$

Thus, similar to the first figure on climate regulatory risk, the economy with lower elasticity between the green and brown capital shows higher positive equity premiums between brown

[7]Figure A.1 1.16 displays the impulse response of green share, $K_t^G/F_t$, with different levels of brown and green capital complementarity . The green share is higher as the elasticity of substitution between green and brown is higher

[8] $\frac{\partial\left(\frac{\partial lnY_t^B}{\partial lnF_t}\right)}{\partial\sigma}|_{ss} = -\frac{1}{\sigma^2} log(\frac{K_{ss}^B}{F_{ss}})\left(\frac{K_{ss}^B}{F_{ss}}\right)^{\frac{\sigma-1}{\sigma}} \times \left[\frac{1}{\left(\frac{K_{ss}^B}{F_{ss}}\right)^{\frac{\sigma-1}{\sigma}}\nu+1-\nu}\right]^2 \nu \times (1-\nu) \times \alpha < 0, \text{ where } \frac{K_{ss}^B}{F_{ss}} > 1$

(high CRE) and green (low CRE) firms. Intuitively, the economic channel is that when the elasticity is lower, there is a strong complementarity between the brown capital and green capital for brown (high CRE) firms. Therefore, brown (higher CRE) firms have less proportional demand for expensive green capital than green (lower CRE) firms[9]. With higher brown capital demand, brown (high CRE exposed) firms also invest more than the green (low CRE) firms[10], contributing to higher market valuation.

**Book-to-Market**

The economy could also identify exogenous source of the book-to-market variation. Within the economy, the book-to-market ratio for brown (high CRE) firms is defined as:

$$BM_t^B := \frac{K_t^B + P^F F_{B,t}}{V_t^B - D_t^B} \text{ and } BM_t^G := \frac{K_t^G}{V_t^G - D_t^G} \tag{53}$$

where according to our definitions, $P_t^i \equiv V_t^i - D_t^i$, where $i \in \{B, G\}$ Figure 1.11 shows that when exposed to long-run productivity, the brown (high CRE) firms and green (low CRE) firms do not differentiate the book-to-market ratio in Panel B. However, when exposed to climate regulatory risk, the marginal investment Q for brown (high CRE) firms is lower than the investment Q for green (low CRE) firms. This contributes to a positive difference on the book-to-market ratio between brown (high CRE) and green (low CRE firms) as in Panel A. The prediction is consistent with Bolton and Kacperczyk (2021c), documenting that carbon emissions positively correlate with the book-to-market ratio with or without industry fixed effect.

### 1.6.3 Endogenous Climate Policy

The previous sections introduce an economy with independent climate regulatory and productivity risks. However, it is not necessarily the case that ad-hoc government operates. The government might design the climate regulatory policy from a social welfare view. The next question is what's the price of risk within this first-best economy ? The following experimen-

[9]The price of brown capital is only 36.6% to 38% of the green capital depending on how to specify the calibration

[10]Figure 1.9 Panel (E) and Panel (F)

tation designs an economy under the social welfare view and derives the analytic formula for endogenous carbon taxation in the form of Pigovian taxation.

### Carbon Price

**Proposition 3** *Under the previous assumptions, the optimal carbon tax is defined as 54 by optimizing the social planner problem in appendix 1.10.5. The optimal taxation is defined as:*

$$\tau_t^* = \frac{\lambda_t^S}{\lambda_t \gamma_B}(1-\zeta)\{\frac{\Gamma_t}{\phi_1 F_t \gamma_B}\}^{\zeta} \tag{54}$$

*where $\lambda_t^S$ denotes the value of carbon emission storage in the whole economics:*

$$\lambda_t^S = \underbrace{\lambda_t \Psi}_{Preference\ Cost} + \underbrace{\lambda_t(1-\alpha)\mu\frac{Y_t}{Z_{t-1}}}_{Production\ Cost} + \underbrace{\rho_S E_t\{M_{t,t+1}(\frac{U_{t+1}/\widetilde{C}_{t+1}}{U_t/\widetilde{C}_t})^{-\frac{1}{\psi}}\lambda_{t+1}^S\}}_{Utility\ Continuation} \tag{55}$$

*$\lambda_t$ denotes the household marginal utility of the consumption. $\rho_S$ refers to the AR(1) carbon accumulation coefficient.*

According to proposition 3, the price of carbon concentration is divided into three parts as in equation 55. The first component implies the direct effect through the preference channel. A higher environmental preference $\Psi$ indicates more emission externalities through a household preference channel. The production cost tracks the emission externalities through the production channel at a given $\mu$. A higher $\mu$ implies that damage on production is more serious given the carbon concentration. Finally, the utility continuation documents the discounted future carbon price. It tracks the dynamic effect of carbon price in economics.

Given the shadow price of carbon concentration, the optimal carbon taxation is defined as 54, where the optimal climate taxation is proportional to both the shadow price of carbon concentration ($\frac{\lambda_t^S}{\lambda_t}$) and marginal carbon production from the brown sector ($\{\frac{\Gamma_t}{\phi_1 F_t \gamma_B}\}^{\zeta}$). Figure 1.12 documents the optimal carbon taxation under this optimal design. Compared with the short-run risk, long-run risk matters more for cross-sectional equity premiums under optimal carbon taxation. When a positive long-run shock materializes, the brown (higher CRE exposed) firms generate higher stock returns. Simultaneously, the stochastic discount factor becomes smaller in response to the positive shock. This combined effect consistently implies that the price of

the long-run productivity risk is positive. As carbon emissions increase, the price of carbon risk is still positive when completely endogenous carbon taxation prevails.

## 1.7 Data

The paper exploits both the data at the firm level and the aggregate level. For the firm-level information, the financial accounting data is from the annual COMPUSTAT and matches the stock market performance from CRSP. The innovative text data on climate regulatory exposure (CRE) is from Baz et al. (2022a) and connects it with WRDS through CIK. The details of data construction are below.

### 1.7.1 SEC Filings

The sample period ranges from January 2009 to December 2015. This sample period is because the information on climate risk is not sufficiently disclosed before 2008. After 2008 as Baz et al. (2022a) documents, investors began requesting climate information disclosure through the SEC petition system. After Paris Agreement, since president Trump's election contributed to the deregulation of climate policy, the empirical test focuses on a stable CRE period.

There are 58,473 10-K filings in total, as reported in Table 2.1. On average, every year in our sample includes around 8,021 10-K filings for around 7930 individual firms. In addition, the average length of the document is increasing year by year, ranging from 500.64 to 7300.06 words. We use the industry classification according to the Fama-French definitions, excluding financial and banking services (Fama and French (1992)).

To construct a climate regulatory exposure measure, I implement the measure from Baz et al. (2022a) and modify their equations for sentiment (as inHassan et al. (2019a) and Malmendier et al. (2020)), as follows:

$$CRE_{i,t} \equiv \frac{1}{B_{i,t}} \left( \sum_{b=1}^{B_{i,t}} \mathbb{I}_{\{b \in CCV\}} \times S_b(r) \right), \tag{56}$$

where $B_{i,t}$ refers to the length of the document, $\mathbb{I}_{\{b\in CCV\}}$ denotes the indicator whether the term $b$ belongs to the Climate Change Vocabularies (CCV)[11]. The innovative contribution of the algorithm comes from the $S_b(\cdot)$ function; it is sum of the number of GPV terms that are within a distance of $r$ words from word $b$. Formally, $S_b(r)$ is defined as:

$$S_b(r) \equiv \sum_{c=b-r}^{b+r} \mathbb{I}_{\{c\in GPV\}}, \tag{57}$$

where $\mathbb{I}_{\{c\in GPV\}}$ is an indicator variable equal to 1 if the c$^{\text{th}}$ term in the document is in Government Policy Vocabularies (GPV)[12], and zero otherwise. The algorithm looks for government policy terms that fall within a distance $r$ from climate risk terms.[13]

For example, given a specific word distance (e.g., $r$ equals 10), the algorithm will search around the key phrase representing climate risk and record the count of government policy phrases within this word distance. The idea is that if climate risk topic is being discussed, and GPV terms are also mentioned near this climate risk discussion, it becomes more likely that the firm is worried about climate regulatory exposures. Figure 1.13 displays the cross-industry variation of the CRE based on the first digit of the SIC industry code. It is intuitively consistent that mining (including oil and coal) has the most serious regulatory concerns, and transportation, including gas and electricity, comes second. The first row of Table 1.2 shows the summary of statistics for the climate regulatory exposure conditional on different environmental preference level ($CRE_{i,t} \times MCCC_t$) with 7.9 basis points on average ($10^{-4}$) and 1.94 basis points ($10^{-4}$) standard deviation. The CRE measure in second row displays 7.5 basis points ($10^{-4}$) on average with a standard deviation of 1.74 bps (basis points, $10^{-4}$). In addition, in order to control other text-implied variables, I also include the text implied $Constraint_{i,t}$ (financial constraint), $Legislation_{i,t}$ (legislation uncertainty) and $Complexity_{i,t}$ (document complexity) following Loughran and McDonald (2016)[14]. These text-implied variables are constructed via normalizing the dictionary count with document length, $\frac{1}{B_{i,t}} \sum_{b=1}^{B_{i,t}} \mathbb{I}_{\{b\in Dict_{topic}\}}$, where '$Dict_{topic}$' refers to the topics on financial constraint, legislation and complex dictionaries respectively.

[11]CCV refers to a climate change dictionary from Chou and Kimbrough (2019a)

[12]GPV refers to the dictionary for government policy concerns. It's a dictionary used in Koijen et al. (2016)

[13]In what follows, we use a value of $r$ equal to 10 following the Malmendier et al. (2020).

[14]Their public data code is shared through link:`https://sraf.nd.edu/textual-analysis/`

### 1.7.2 Stock Returns and Firm Fundamentals

As for financial accounting data and equity information, the paper merges the CRSP with the COMPUSTAT from 2009 to 2015. Table 1.2 shows the average monthly stock returns, $R^i_{t+1}$, with 2%. $Leverage_{i,t}$ refers to the book leverage of the firm. $Size_{i,t}$ represents the firm's market valuation ($10^9$ dollars) in year t. $BM_{i,t}$ denotes the book-to-market ratio. $Investment_{i,t}$ refers to capital expenditure normalized by the total asset. $Vol_{i,t}$ means the stock's historical volatility. It is constructed with a one-year rolling window for monthly returns. $MOM_{i,t}$ denotes cumulative stock return with a one-year rolling window. $Beta_{i,t}$ stands for the market $\beta_{i,t}$ from CAPM model with estimators from one-year rolling window. In addition, the paper also exploits the firm-level carbon emission data from Trucost. $Scope^1_{i,t}$ means total direct carbon emission. $Scope^2_{i,t}$ denotes the firm-level indirect carbon emission.

## 1.8 Empirical Results

The paper investigates the testable implications of the model from section 1.6.2. First, the negative (positive) sign of the transition risk price depends on realized high (low) low-carbon preference towards climate issue. Second, firms consider climate regulatory ex-ante and reduce carbon emissions as the climate policy tightens. Third, consistent with Bolton and Kacperczyk (2021c), the paper investigates the relationship between the Book-to-Market ratio and CRE. This helps identify a positive statistical significance as the model predicts.

### 1.8.1 Portfolio Evidence

Table 1.5 shows the results of the portfolio sorting based on the $CRE_{i,t}$ exposure. Since the expected return is relatively hard to measure in reality, I first use the future realized return as an approximate expected return.

The low portfolio return denotes the returns with the bottom 30% of CRE (green) equity premiums based on the 10-K information disclosed in the last calendar year. The high portfolio refers to the top 30% of CRE exposed (brown) portfolio returns. An "H-L" denotes a long short portfolio over the high (brown) against the low (green). Panel A shows that with Newway-West

standard errors of 15 lags, a long-short portfolio is statistically significant and of great economic magnitude with -1% monthly return (annualized -12 %). However, there is a thought in asset pricing that equally-weighted portfolio construction is only affected by small firms. Then, I adjust the portfolio balance with value-weighted schemes.

Panel B in Table 1.5 shows the CRE sorted portfolio based on value-weighted stock returns. The low quantile (green,bottom 30%) displays a 1.7% monthly expected stock return, and the high quantile (brown, top 30%) shows a 1% stock return. Overall, a long-short trading strategy could produce a statistically significant -8.4% annualized return.

Table 1.5 Panel C compares the other fundamentals and explore whether there are monotonic relationship between firm fundamentals and climate regulatory exposure. The variables, including investment, leverage, complexity, size, and sale, show monotonic relationship with CRE. Therefore, it is plausible that the one-way sorting portfolio return is driven by these confounding variables. To cope with this issue, I do two-way portfolio sorting in Table 1.6, Table 1.7 and Table A.1.13. Similar to the previous definitions, the low and high denote a bottom 30% (green) and top 30% (brown) quantiles of the variable. Table 1.6 shows that controlling the sale in low (high) quantile, a long-short trading strategy with an equally weighted scheme earns 15.6% (6%) significant portfolio return. For a value-weighted scheme, a long-short trading strategy realizes a 12% (7.2%) return conditional on a low (high) sale quantile in Panel B. For the double sorting on size and CRE in Panel C, the paper shows that a long-short equally weighted trade strategy realizes 13.2% (8.4%) annualized return conditional on small (large) size. Similarly, a double sorting based on value-weighted scheme earns a 7.2% (8.4%) conditional on low (high) market valuation quantiles.

Table 1.7 shows the double sorting on CRE and leverage (Investment) results. An equally weighted CRE-investment sorted portfolio earns 10.8% and 12% annualized return conditional on lower and higher investment ratio, respectively. For the value-weighted case, even though in the low investment, the CRE sorted long-short trading strategy could merely earn 3.6% annualized return, under the high investment, this long-short trading strategy could still realize with 12% annualized return. In Panel C and Panel D, the main result still holds with both statistical significance and great economic magnitude. Panel C documents 14.4% and 8.4% for the long-short equally weighted portfolios conditional on low and high leverage, respectively. Panel D shows that a similar trading strategy but with a value-weighted scheme realizes 13.2% and 4.8% given lower and higher leverage, respectively. In sum, this double sorting method shows

that the higher CRE (brown) could potentially predict a lower stock return monotonically. This relationship cannot be due to the confounding variables, including size, sale, investment, and leverage. The Appendix also shows that complexity-CRE sorting does not change our main result as well in Table A.1.13.

### 1.8.2 Abnormal Return

Given the portfolio analysis, the next question is whether existing time-varying factors could explain the previously documented portfolio returns. Alternatively, it's interpreted as an undiversified abnormal return. The paper exploits different time series regressions based on the factor model to eliminate the contributions from existing factors.

Table 1.8 shows the time series asset pricing test and adjusts the standard errors with Newey-West 15 lags. Panel A displays the result with CAPM model. After controlling for the market factor, the portfolio's abnormal returns decrease monotonically from 6% (low) to -6% (high) annualized return. A long-short trading strategy shows statistically significant 13.2% (equally weighting) and 13.2% (value weighting) annualized premiums. Panel B explores the return performance over a four-factor model, including market, size, value, and momentum. The undiversified risk return decreases from lower to higher CRE exposure for equally-weighted (value-weighted) strategy, and a long-short trading strategy earns 12% (8.4)% annualized returns. Panel C investigates its robustness with even more factors, including the operating profitability and investment factors. The monotonic statistical significant relationship is robust against the strategy types regardless of equally-weighted or value-weighted schemes. A long-short portfolio between low (green) and high (brown) CRE firms realizes 9.6% and 12%, respectively. This portfolio analysis indicates that our documented CRE-driven returns cannot be explained by the benchmark factor model and are probably due to undiversified regulatory risk.

### 1.8.3 Predictive Panel Regression

However, given the previous regression, it is still probable that our main results might be jointly driven by unobserved confounding variables. In order to show the robustness of the result, the paper implements a predictive regression as Bansal and Shaliastovich (2013). The

main regression design follows:

$$\frac{1}{T}\sum_{j=1}^{T} R^i_{t+j-1,t+j} = \beta_0 + \beta_1 CRE_{i,t-1} \times MCCC_t + \beta_2 CRE_{i,t-1} + Control_{i,t-1} + \beta_3 MCCC_t + \epsilon_{i,t} \tag{58}$$

$\frac{1}{T}\sum_{j=1}^{T} R^i_{t+j-1,t+j}$ denotes the cumulative returns starts from time t to t+T. I choose $T = 3, 6, 9, 12$ months to calculate the 3-month, 6-month, 9-month, and 12-month future average raw returns. The paper follows the strategy as Fama and French (1992). In order to materialize before the realization of stocks, the paper collects the text information from the last calendar year, guaranteeing that there is at least a six-month gap between the 10K filing disclosure and stock return realization. Therefore, the disclosed climate regulatory risk should be realized before the return. $MCCC_t$ denotes the aggregate climate preference measure as Pastor et al. (2021a). For the control variables, I control similar confounding variables as portfolio sorting method, including leverage, book-to-market ratio, investment rate, market valuation, historical volatility, market beta, momentum, text-implied complexity, text-implied constraint, and text-implied legislation risk.

Table 1.9 shows the regressions with different time horizons. From the empirical regression with $CRE_{i,t}$, there exists a risk compensation finding on CRE. For instance, when $T = 3$, the exposure to CRE in column (1) and column (2) increases the future stock return. With one standard deviation of 1.74 ($10^{-3}$) increase in $CRE_{i,t}$, the future predictive monthly return in column (1) (column (2)) will go up by 1.04% (1.31%) with both statistical significance and great economic magnitude.

More importantly in Table 1.9, for the interaction effect between $MCCC_t$ and $CRE_{i,t-1}$, there exists strong negative coefficients across different time horizons. For a 3-month average future return, with one standard deviation change of the variable, $MCCC_{t-1} \times CRE_{i,t-1}$, the future average return decreases by 1.34% or 1.53% depending on how to specify the controls. The negative effect is stronger while calculating the 6-month average future returns with 1.64% and 1.61%. The 9-month average future returns are 1.22% (column 5) and 1.26% (column 6), conditional on the specification of the controls. This negative effect is even weaker after 12 months but it is still statistically significant and of great economic magnitude, with 1.16% and 1.20% in column (7) and column (8). Overall, the paper shows that when exposed to climate regulatory exposure, investors have a joint effect of risk for compensation and demand for a low-carbon environment. When the general preference is more intensive as $MCCC_t$, a higher

exposure to climate regulatory risk reduce the expected returns.

In the Table 1.10, the paper also shows an alternative way to test this negative relationship through Fama-French cross-sectional regression, controlling more confounding variables, including financial stress, leverage etc. A one standard deviation change over CRE implies at least 23 basis point reduction over monthly stock returns depending on how to specify the controlling variables. The conditional price of climate regulatory risk from Fama-Macbeth regression is negatively correlated with $MCCC_t$ in Table 1.16. In addition, as Bolton and Kacperczyk (2021c) argues that carbon emissions are directly linked to the book-to-market ratio, to rule out the potential effect from carbon accumulation, Table A.1.14 and A.1.15 control the potential variation with direct carbon emissions and indirect carbon emissions, respectively. Even though future 9-month average return in column (5) and (6) are insignificantly negative when climate attention spikes, our main results still hold with great economic magnitude when $T = 3, 6, 12$.

### 1.8.4 Event Study

Given the discussion above, the paper confirms the implication of climate policy on cross-sectional asset pricing. The next unresolved question concerns the story's sensibility around critical historical events. In Figure 1.14, the paper studies realized stock performance around two key events. The first is the Paris Agreement in Dec, 2015, when the international greenhouse gas reduction commitment is signed. Countries, including the U.S., enforce their climate regulation and restrict carbon emissions. With this regulatory shock to the U.S., the figure shows that the firms with higher climate regulatory exposure tend to be lower than the firms with higher CRE.

Similarly, the overall market is shocked unexpectedly by the climate deregulation event in the U.S., Donald Trump's election in Dec 2016, since President Trump claims that his status on climate change is a denier. In response, firms with higher CRE (brown) realize high stock returns than those with lower CRE (green) as Figure 1.15. The details of the event analysis with Diff-in-Diff are intensively discussed in the other paper, Baz et al. (2022a).

### 1.8.5 Carbon Emission

In contrast to Hsu et al. (2022), this paper is based on the assumption that firms take climate policy ex-ante and tend to reduce their carbon emissions when policy tightens. The paper explores the relationship between text-implied climate policy concern and carbon emissions (scope 1 and scope 2):

$$Carbon_{i,t} = \beta_0 + \beta_1 CRE_{i,t} \times MCCC_t + Control_{i,t} + \epsilon_{i,t} \tag{59}$$

Table 1.11 shows that carbon emissions are reduced with higher $CRE$. Dependent on how to specify the control variables, one standard deviation change of $CRE_{i,t} \times MCCC_t$ contributes to 13.32 (13.49) tons of $CO_e^2$ reduction in direct greenhouse gas emission without (with) confounding variables as control (column (1) and (2)). This coefficient indicates that when media attention to climate issues is relatively high, a higher firm's perceived climate regulatory risk tends to reduce direct carbon emissions in a statistically significant method. Concerning the indirect emissions, even though indirect greenhouse gas (column (3) and column (4)) displays with trivial significance, the negative relationship still holds. In Table 1.11 column (5) and column (6), the regressions show the statistically significant climate regulatory impact on greenhouse gas emission, including indirect emission and direct emission. The estimators are both statistically significant and of great economic magnitude. One standard deviation change in $CRE_{i,t} \times MCCC_t$ contributes to 14.36 (14.58) tons of greenhouse gas reduction in direct and indirect emission in column (5) (column (6)).

These negative findings are significant when the regressions include the industry-level fixed effect indicating that all these results are prominent at the within-industry level variation. In the appendix Table 1.17, the paper also shows the regressions without industry-level fixed effect, and the relationship is positively significant. This result implies that a higher CRE at the cross industry level implies a higher carbon emission. Overall, this evidence shows that conditional on a higher social preference for a low-carbon environment, firms endogenize their carbon emission decision with climate policy risk.

### 1.8.6 Book-to-Market Ratio

The model shows that due to the climate policy risk, firms with higher climate policy exposure (brown) have a higher book-to-market ratio with lower marginal Q of clean capital investment. This prediction is consistent with Bolton and Kacperczyk (2021c) that higher carbon emissions contribute to a higher book-to-market ratio. In this section, I provide further empirical evidence to convince this channel. The following regression explores the relationship between the book-to-market ratio and CRE exposure:

$$log(BM_{i,t}) = \beta_0 + \beta_1 CRE_{i,t} \times MCCC_t + \beta_2 Carbon_{i,t} \times MCCC_t + Control_{i,t} + \epsilon_{i,t} \quad (60)$$

Table 1.12 documents that conditional on a sensitive environmental preference (high $MCCC_t$), a higher firm $CRE_{i,t}$ contributes to a higher book-to-market ratio.The result is also robust as Bolton and Kacperczyk (2021c) regardless of whether the industry fixed effect is included or not. Moreover, in order to control the instantaneous variation from the carbon emission, I do include the interaction terms of $Carbon_{i,t} \times MCCC_t$, $Carbon_{i,t}$. Our empirical result indicates that the high book-to-market ratio at carbon-intensive sectors is due to their climate policy risk exposure.

## 1.9 Conclusion

In this paper, I propose a framework to show how climate policy risk is reflected in the cross-sectional asset price and how investors perceive these climate policy shocks by using additive low-carbon preference. This innovative addition indicates that Media Climate Change Concerns in Ardia et al. (2020) is closer to an interpretation of virtue signaling than Pastor et al. (2021a) documents since investors sacrifice risk compensation for higher environmental quality. The model's strength helps identify the exogenous source for the positive price of carbon risk in Bolton and Kacperczyk (2021a). Moreover, the model sheds light on the negative price of transition risk for U.S firms when climate concerns spike. The price of transition risk is identified by the firm carbon emissions involving cash flow and marginal utility channels. These findings highlight the propagation of transition risk over the economy through both cash flow and investor awareness channels.

In addition, the paper proposes a measure of firm-level climate regulatory risk based on SEC 10K filings and documents a significant negative abnormal premium within cross-sectional equity pricing. Firms adjust their carbon emission decision according to the perceived climate regulatory risk. Firms with low climate regulatory exposure (green) carry an 8.4% or 12% premium relative to the higher (brown) conditional on how to specify the control. Moreover, the innovative measure helps explain the finding in Bolton and Kacperczyk (2021c) on book-to-market ratio and transition risk relationship. Based on this framework, other friction or firm environmental characteristics can be included in the future to explain climate policy or carbon premiums.

# Tables

Table 1.1: Summary of Statistics for 10K Raw Information

| Year | Firm Num | Doc Num | Doc.Ave.Length | Ind Num |
|---|---|---|---|---|
| 2009 | 9346 | 9463 | 5500.64 | 48 |
| 2010 | 8802 | 8938 | 5702.34 | 48 |
| 2011 | 8481 | 8604 | 6038.08 | 48 |
| 2012 | 8090 | 8173 | 6263.85 | 48 |
| 2013 | 7748 | 7845 | 6564.35 | 48 |
| 2014 | 7692 | 7788 | 6928.94 | 48 |
| 2015 | 7590 | 7662 | 7300.06 | 48 |

Note: The table shows the summary of statistics for the document filings in our sample. **Firm Num** refers to the number of firms in our sample counted through the unique cik number. **Doc Num** is the number of the reported SEC 10-K filings. **i**s for the average length of the document. **Ind Num** represents the number of Fama-Kennetch defined 48 industry count

Table 1.2: Summary of Statistics for Firm Characteristics

| Variables | Average | Q25 | Median | Q75 | Max | Std.Dev |
|---|---|---|---|---|---|---|
| $CRE_{i,t} \times MCCC_t(10^{-3})$ | 0.79 | 0.00 | 0.06 | 0.53 | 78.19 | 1.94 |
| $CRE_{i,t}(10^{-3})$ | 0.75 | 0.00 | 0.06 | 0.53 | 35.71 | 1.74 |
| $Complexity_{i,t}(10^{-3})$ | 16.11 | 11.85 | 15.59 | 19.51 | 225.81 | 6.93 |
| $Constraint_{i,t}(10^{-3})$ | 9.38 | 7.47 | 9.24 | 11.24 | 48.54 | 3.19 |
| $Legislation_{i,t}(10^{-3})$ | 19.51 | 14.66 | 17.93 | 21.41 | 157.89 | 11.32 |
| $Leverage_{i,t}$ | 0.19 | 0.01 | 0.15 | 0.31 | 0.93 | 0.18 |
| $BM_{i,t}$ | 0.78 | 0.32 | 0.56 | 0.93 | 52.28 | 1.15 |
| $Investment_{i,t}$ | 0.05 | 0.01 | 0.03 | 0.06 | 0.82 | 0.06 |
| $Size_{i,t}(10^9)$ | 0.01 | 0.00 | 0.00 | 0.00 | 0.75 | 0.03 |
| $Vol_{i,t}$ | 0.13 | 0.07 | 0.11 | 0.16 | 1.81 | 0.09 |
| $Beta_{i,t}$ | 0.01 | 0.00 | 0.01 | 0.02 | 1.49 | 0.10 |
| $MOM_{i,t}$ | 0.16 | −0.10 | 0.15 | 0.39 | 6.74 | 0.53 |
| $R^i_{t,t+1}$ | 0.02 | −0.05 | 0.01 | 0.07 | 7.88 | 0.15 |
| $Scope^1_{i,t}$ | 286.56 | 2.21 | 7.74 | 44.14 | 15700.05 | 1224.36 |
| $Scope^2_{i,t}$ | 42.18 | 2.50 | 7.84 | 28.73 | 1664.48 | 120.43 |
| $(Scope^1 + Scope^2)_{i,t}$ | 328.74 | 6.13 | 20.96 | 91.01 | 15702.76 | 1263.82 |

Note: The table 1.2 show the summary of statistics for the main regression in the paper. $MCCC_t$ refers to the aggregate media attention to climate change from news in Ardia et al. (2020). $CRE_{i,t}$ denotes the Climate Regulatory Exposure (CRE) in Baz et al. (2022a). $Complexity_{i,t}$ refers to the dictionary-based text complexity measure from Loughran and McDonald (2016). $Constraint_{i,t}$ refers to the text-implied financial constraint from Loughran and McDonald (2016). $Legislation_{i,t}$ is the text-implied legal uncertainty with reference to dictionary in Loughran and McDonald (2016). $Leverage_{i,t}$ denotes the firm book leverage. $BM_{i,t}$ is the book-to-market ratio. $Investment_{i,t}$ is the firm investment rate. $Size_{i,t}$ denotes the firm market valuation. $Vol_{i,t}$ is the firm last year volatility. $Beta_{i,t}$ denotes the market exposure from CAPM model with observations from last year. $MOM_{i,t}$ is the cumulative stock raw return from the last year. $R^i_{t,t+1}$ stands for the stock raw return from t to t+1. $Scope^1_{i,t}$ is the firm-level greenhouse gas direct emission, $10^4$ tons of $CO^2_e$. $Scope^2_{i,t}$ refers to the firm-level greenhouse gas indirect emission, $10^4$ tons of $CO^2_e$. $(Scope^1 + Scope^2)_{i,t}$ is the summation of scope 1 and scope 2 greenhouse gas emissions, $10^4$ tons of $CO^2_e$.

Table 1.3: Calibration Parameter in the Model

| Parameter | Symbol | Value | Source |
|---|---|---|---|
| **Household** | | | |
| Subjective Discount Rate | $\beta$ | 0.991 | Croce et al. (2012) |
| Elasticity of intertemporal substitution | $\psi$ | 3 | |
| Risk aversion | $\gamma$ | 20 | |
| Carbon sensitivity | $\Psi$ | 0-1 | |
| **Final Good Sector** | | | |
| Elasticity of Intermediate Good Demand | $\sigma_y$ | 0.08 | Fried (2018b)Hassler et al. (2021) |
| Density of High CRE Firms | $\gamma_B$ | 0.34 | |
| Density of Low CRE Firms | $\gamma_G$ | 0.66 | |
| **Intermediate Firm** | | | |
| Capital Input Share | $\alpha$ | 0.285 | Fried (2018b) |
| Depreciation Rate | $\delta$ | 12% | |
| Damage Function Multiplier | $\mu$ | $5 \times 10^{-6}$ | |
| capital cost function coefficient | $\eta$ | 1/3 | Massimiliano Croce (2014) |
| Total Labor | $L$ | 1 | |
| Long Run Risk Coefficient | $\rho_z$ | 0.9 | Massimiliano Croce (2014) |
| Long run growth | $\mu_Z$ | 1.8% | Output Growth Moment in the U.S. |
| Stationary Growth Risk Std Dev | $\sigma_a$ | 0.0116% | |
| Nonstatoinary Growth Risk Std Dev | $\sigma_z$ | 0.131% | $\sigma_a/\sigma_z$, Massimiliano Croce (2014) |
| **Brown Firm** | | | |
| Brown capital Share | $\nu$ | 0.2-0.9 | |
| Elasticity of Capital Demand | $\sigma$ | 0.03 | Fried (2018b)Hassler et al. (2021) |
| **Energy Supplier** | | | |
| Growth Sensitivity to Land | $\rho_a$ | 0.1 | Ready (2018) |
| Land Resource (Steady State) | $\rho_a$ | 0.05 | Bornstein et al. (2022) |
| Exploration Share to Energy Reserve | $\lambda$ | 0.9 | Bornstein et al. (2022) |
| Energy investment share | $\xi$ | 0.15 | |
| Energy cost convex coefficient | $\pi_1$ | 0.09 | |
| **Carbon Cycle** | | | |
| immediate carbon generation | $\xi$ | 0.675 | |
| Average Alternative Green Gas | $\Gamma_{ss}$ | 0 | |
| Alternative carbon coefficient | $\rho_\Gamma$ | 0.489 | |
| Carbon generation share ($S_t$) | $\rho_S$ | 0.612 | |
| Anternative Carbon Generation Risk Std Dev | $\sigma_\Gamma$ | 0.337 | |
| **Government** | | | |
| Expected Carbon Price | $\tau^*$ | 36 | Hassler et al. (2018) |
| Carbon price probability coefficient | $\rho_w$ | 95.90% | |
| Carbon Price Probability Std Dev | $\sigma_w$ | 24.24% | |

Note: Table 1.3 shows calibration parameters used in the model. All the parameters are implemented to calibrate for the benchmark model. For the other recalibration comparison including the 'No Regulation', this benchmark model parameter setting is used as well.

Table 1.4: Macro Fundamentals and Key Moment Simulation

| | Data | Benchmark | High $\sigma$ | Low $\Psi$ | Low $\sigma_w$ |
|---|---|---|---|---|---|
| **Asset Pricing Implication** | | | | | |
| $E(r_B - r_f)$ | | 3.75% | 5.19% | 1.76% | 6.37% |
| $E(r_G - r_f)$ | | 8.04% | 5.35% | 1.49% | 6.58% |
| $E(r_m - r_f)$ | 5.66% | 5.32% | 5.29% | 1.58% | 6.51% |
| $E(r_B - r_G)$ | | -4.29% | -0.16% | 0.27% | -0.22% |
| $\sigma_{r_B}$ | | 25.07% | 8.98% | 23.1% | 12.96% |
| $\sigma_{r_G}$ | | 9.39% | 6.94% | 2.57% | 9.16% |
| $\sigma_{r_m}$ | 17.82% | 12.92% | 7.44% | 9.01% | 9.68% |
| **Macro Fundamentals** | | | | | |
| $E(\Delta c_t)$ | 2.22% | 1.81% | 1.80% | 1.80% | 1.80% |
| $E(\Delta y_t)$ | 2.06% | 1.84% | 1.83% | 1.81% | 1.83% |
| $E(\Delta ca_t)$ | -0.08% | -0.07% | -0.08% | -0.10% | -0.08% |
| $E(\theta_t)$ | 17.60% | 22.81% | 12.64% | 22.83% | 21.99% |
| $\sigma_{\Delta c}$ | 2.15% | 2.48% | 1.94% | 0.79% | 2.38% |
| $\sigma_{\Delta y}$ | 2.18% | 0.72% | 1.25% | 0.72% | 0.34% |
| $\rho_{\Delta c, \Delta ca}$ | 72.25% | 82.58% | 77.45% | 74.63% | 82.55% |

Note: Table 1.4 compares the observed quantitative moments with the model predictions. The column (1) summarizes the aggregate fundamental moments. $r_B$ refers to the brown firm stock return. $r_G$ is the green firm stock return. $r_m$ means the market stock return (SP500 return). $\sigma_{r_B}, \sigma_{r_G}$ and $\sigma_{r_m}$ denotes the annualized volatility for brown firm return, green firm return and market return respectively. $\Delta y$ is the output growth of the economy. $\Delta c$ refers to the consumption growth. $\Delta ca$ represents the carbon growth of the whole economy. $\theta_t$ refers to the oil production rate as approximate to energy extraction.The benchmark column (2) refers to the model simulated given Table 1.3. The column (3) denotes the model simulation when the elasticity between brown and green capital for brown firms is relatively higher (0.26). The column (3) refers to the model simulation result when the preference for low carbon is 0.5. The column (4) refers to the model simulated when the climate regulatory risk is less volatile with $\sigma_w = 10.39\%$ $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 1.5: One-Way Portfolio Sorted Result

Panel A: Equal-Weighted Sorted Portfolios

| | *Dependent variable: EW Portfolio* | | | |
|---|---|---|---|---|
| | Low | Middle | High | H -L |
| $Portfolio\ Return_{i,t}$ | $0.020^{***}$ | $0.016^{**}$ | 0.010 | $-0.010^{***}$ |
| | (0.007) | (0.006) | (0.006) | (0.003) |

Panel B: Value-Weighted Sorted Portfolios

| | *Dependent variable: VW Portfolio* | | | |
|---|---|---|---|---|
| | Low | Middle | High | H - L |
| $Portfolio\ Return_{i,t}$ | $0.016^{***}$ | $0.014^{***}$ | $0.010^{*}$ | $-0.007^{**}$ |
| | (0.005) | (0.005) | (0.006) | (0.003) |

Panel C: Portfolio Sorting Characteristics

| | Low (Bottom 30%) | Middle | High (Top 70%) |
|---|---|---|---|
| $Size_{i,t}$ | 0.005 | 0.007 | 0.009 |
| $Sale_{i,t}$ | 3023.096 | 6085.853 | 11854.342 |
| $Leverage_{i,t}$ | 0.196 | 0.198 | 0.254 |
| $Complexity_{i,t}$ | 268.783 | 200.402 | 179.322 |
| $Investment_{i,t}$ | 0.038 | 0.041 | 0.098 |
| $BM_{i,t}$ | 0.822 | 0.832 | 0.822 |
| $Constraint_{i,t}$ | 162.561 | 119.220 | 135.961 |
| $Legistlation_{i,t}$ | 305.962 | 237.103 | 243.652 |
| $MOM_{i,t}$ (Momentum) | 0.147 | 0.170 | 0.148 |
| $Beta_{i,t}$ (Market Beta) | 0.013 | 0.013 | 0.012 |
| $Vol_{i,t}$ (Volatility) | 0.133 | 0.117 | 0.117 |

Note: Table 1.5 shows the one-way portfolio sorting result based on the CRE proposed measure from the last calendar year. Panel A denotes the portfolio rebalance through equally weighted strategy. Panel A refers to the portfolio construction with value-weighted strategy. The 'High' implies that the portfolio is constructed based on top 70% of CRE ranked stocks. The 'Low' means that the portfolio is constructed with bottom 30% of CRE ranked stocks. Consistently, Panel C denotes the firm summary statistics sorted based on climate regulatory exposure (CRE). $Size_{i,t}$ denotes the firm market valuation ($10^9$ dollars). $Sale_{i,t}$ refers to the firm sales. $BM_{i,t}$ is the book-to-market ratio. $Investment_{i,t}$ is the firm investment rate. $Complexity_{i,t}$ refers to the dictionary-based text complexity measure from Loughran and McDonald (2016). $Constraint_{i,t}$ refers to the text-implied financial constraint from Loughran and McDonald (2016). $Legislation_{i,t}$ is the text-implied legal uncertainty with reference to dictionary in Loughran and McDonald (2016). $Leverage_{i,t}$ denotes the firm book leverage. $Vol_{i,t}$ is the firm last year volatility. $Beta_{i,t}$ denotes the market exposure from CAPM model with observations from last year. $MOM_{i,t}$ is the cumulative stock raw return from the last year. All standard errors are adjusted with Newey–West method with 15 lags. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

Table 1.6: Two-Way Portfolio Sorted Result

Panel A: Double Sorted Portfolios: Sale & CRE (EW)

| | *Dependent variable: Equally-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Sale* | | | *High Sale* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.023*** | 0.010 | -0.013*** | 0.019*** | 0.013** | -0.005* |
| | (0.007) | (0.008) | (0.004) | (0.006) | (0.006) | (0.003) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel B: Double Sorted Portfolios: Sale & CRE (VW)

| | *Dependent variable: Value-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Sale* | | | *High Sale* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.016** | 0.006 | -0.010* | 0.016*** | 0.010* | -0.006** |
| | (0.007) | (0.007) | (0.006) | (0.005) | (0.005) | (0.003) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel C: Double Sorted Portfolios: Size & CRE (EW)

| | *Dependent variable: Equally-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Size* | | | *High Size* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.023*** | 0.012 | -0.011** | 0.019*** | 0.011* | -0.007** |
| | (0.007) | (0.008) | (0.004) | (0.006) | (0.006) | (0.003) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel D: Double Sorted Portfolios: Size & CRE (VW)

| | *Dependent variable: Value-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Size* | | | *High Size* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.015** | 0.010 | -0.006* | 0.016*** | 0.010* | -0.007** |
| | (0.007) | (0.007) | (0.003) | (0.005) | (0.005) | (0.003) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Table 1.6 shows the double sorting portfolio result with CRE. Panel A shows the result of double sorting with sale and climate regulatory exposure (CRE). The portfolio is balanced annually with equally-weighting scheme. Panel B displays the double sorting with both sale and CRE. The portfolio is balanced every year with value weights. Panel C refers to the result sorted based on stock size and CRE disclosure with equally weighted portfolios. Panel D shows the double sorting with size and CRE measure. The portfolio is constructed with value weighted method. In the Table, 'high' refers to the portfolio with top 70th percentiles of CRE (sale, size) sorted stocks and 'low' refers to the portfolio with bottom 30th percentiles of CRE (sale, size) sorted stocks. 'L-H' implies a long-short trading strategy by longing the low (bottom 30%) CRE exposed stocks and shorting high (top 70%) CRE exposed stocks. All standard errors are adjusted with Newey–West method with 15 lags. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.7: Two-Way Portfolio Sorted Result

Panel A: Double Sorted Portfolios: Investment & CRE (EW)

| | *Dependent variable: Equally-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Investment* | | | *High Investment* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.023*** | 0.014** | -0.009*** | 0.019*** | 0.009 | -0.010** |
| | (0.006) | (0.007) | (0.003) | (0.007) | (0.008) | (0.005) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel B: Double Sorted Portfolios: Investment & CRE (VW)

| | *Dependent variable: Value-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Investment* | | | *High Investment* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.014*** | 0.012** | -0.003 | 0.018*** | 0.008 | -0.010** |
| | (0.005) | (0.005) | (0.003) | (0.005) | (0.006) | (0.004) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel C: Double Sorted Portfolios: Leverage & CRE (EW)

| | *Dependent variable: Equally-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Leverage* | | | *High Leverage* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.023*** | 0.011 | -0.012*** | 0.019*** | 0.013* | -0.007* |
| | (0.006) | (0.007) | (0.004) | (0.007) | (0.007) | (0.004) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel D: Double Sorted Portfolios: Leverage & CRE (VW)

| | *Dependent variable: Value-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Leverage* | | | *High Leverage* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.020*** | 0.009 | -0.011** | 0.014*** | 0.011** | -0.004* |
| | (0.005) | (0.007) | (0.004) | (0.005) | (0.005) | (0.002) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Note: Table 1.7 shows the double sorting portfolio result with CRE. Panel A shows the result of double sorting with sale and climate regulatory exposure (CRE). The portfolio is balanced annually with equally-weighting scheme. Panel B displays the double sorting with both sale and CRE. The portfolio is balanced every year with value weights. Panel C refers to the result sorted based on stock size and CRE disclosure with equally weighted portfolios. Panel D shows the double sorting with size and CRE measure. The portfolio is constructed with either equally weighted or value weighted method. In the Table, 'high' refers to the portfolio with top 70th percentiles of CRE (sale, size) sorted stocks and 'low' refers to the portfolio with bottom 30th percentiles of CRE (sale, size) sorted stocks. 'L-H' implies a long-short trading strategy by longing the low (bottom 30%) CRE exposed stocks and shorting high (top 70%) CRE exposed stocks. All standard errors are adjusted with Newey–West method with 15 lags. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.8: Portfolio Based Asset Pricing Test

Panel A: Asset Pricing Test in CAPM Model

| | *Dependent variable: Equity Premium* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Equal-Weighted | | | | Value-Weighted | | | |
| | L | M | H | H-L | L | M | H | H-L |
| $\alpha^{CRE}$ | $0.005^{*}$ | $-0.001$ | $-0.005$ | $-0.011^{***}$ | $0.004^{***}$ | 0.001 | $-0.004$ | $-0.008^{**}$ |
| | (0.003) | (0.002) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) | (0.003) |
| $\beta^{MKT}$ | $1.236^{***}$ | $1.332^{***}$ | $1.341^{***}$ | 0.105 | $0.963^{***}$ | $1.033^{***}$ | $1.071^{***}$ | 0.108 |
| | (0.064) | (0.054) | (0.076) | (0.072) | (0.035) | (0.034) | (0.053) | (0.069) |

Panel B: Asset Pricing Test in Carhart Four-Factor Model

| | *Dependent variable: Equity Premium* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Equal-weighted | | | | Value-Weighted | | | |
| | L | M | H | H-L | L | M | H | H-L |
| $\alpha^{CRE}$ | $0.007^{***}$ | 0.001 | $-0.003$ | $-0.010^{***}$ | $0.004^{**}$ | 0.001 | $-0.004$ | $-0.007^{**}$ |
| | (0.002) | (0.001) | (0.003) | (0.003) | (0.002) | (0.001) | (0.002) | (0.003) |
| $\beta^{MKT}$ | $1.011^{***}$ | $1.077^{***}$ | $1.093^{***}$ | 0.083 | $0.992^{***}$ | $1.013^{***}$ | $1.024^{***}$ | 0.032 |
| | (0.051) | (0.033) | (0.072) | (0.087) | (0.042) | (0.038) | (0.063) | (0.084) |
| $\beta^{SMB}$ | $0.722^{***}$ | $0.716^{***}$ | $0.505^{***}$ | $-0.217$ | $-0.039$ | $-0.175^{***}$ | 0.108 | 0.147 |
| | (0.082) | (0.054) | (0.116) | (0.139) | (0.067) | (0.062) | (0.102) | (0.135) |
| $\beta^{HML}$ | $-0.101$ | 0.087 | 0.049 | 0.150 | $-0.152^{**}$ | $0.160^{***}$ | $-0.058$ | 0.094 |
| | (0.080) | (0.053) | (0.115) | (0.137) | (0.066) | (0.060) | (0.100) | (0.133) |
| $\beta^{MOM}$ | $-0.212^{***}$ | $-0.153^{***}$ | $-0.285^{***}$ | $-0.074$ | $-0.062^{**}$ | $-0.034$ | $-0.095^{**}$ | $-0.033$ |
| | (0.038) | (0.025) | (0.054) | (0.065) | (0.031) | (0.029) | (0.047) | (0.063) |

Panel C: Asset Pricing Test in Six Factor Model (Carhart+Fama Five)

| | *Dependent variable: Equity Premium* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Equal-Weighted | | | | Value-Weighted | | | |
| | L | M | H | H-L | L | M | H | H-L |
| $\alpha^{CRE}$ | $0.007^{***}$ | 0.001 | $-0.004$ | $-0.010^{***}$ | $0.004^{**}$ | 0.001 | $-0.004^{*}$ | $-0.008^{**}$ |
| | (0.002) | (0.001) | (0.003) | (0.003) | (0.002) | (0.001) | (0.002) | (0.003) |
| $\beta^{MKT}$ | $0.987^{***}$ | $1.091^{***}$ | $1.108^{***}$ | 0.121 | $0.982^{***}$ | $1.036^{***}$ | $1.049^{***}$ | 0.067 |
| | (0.051) | (0.034) | (0.075) | (0.088) | (0.043) | (0.038) | (0.065) | (0.086) |
| $\beta^{SMB}$ | $0.669^{***}$ | $0.746^{***}$ | $0.537^{***}$ | $-0.132$ | $-0.060$ | $-0.124^{**}$ | 0.151 | 0.211 |
| | (0.084) | (0.055) | (0.122) | (0.143) | (0.070) | (0.061) | (0.106) | (0.140) |
| $\beta^{HML}$ | $-0.168$ | $0.115^{*}$ | 0.086 | 0.254 | $-0.182^{**}$ | $0.224^{***}$ | $-0.074$ | 0.107 |
| | (0.102) | (0.067) | (0.149) | (0.174) | (0.085) | (0.075) | (0.129) | (0.171) |
| $\beta^{MOM}$ | $-0.220^{***}$ | $-0.150^{***}$ | $-0.281^{***}$ | $-0.061$ | $-0.066^{**}$ | $-0.026$ | $-0.102^{**}$ | $-0.036$ |
| | (0.039) | (0.026) | (0.057) | (0.067) | (0.033) | (0.029) | (0.049) | (0.065) |
| $\beta^{RMW}$ | $-0.291^{**}$ | $0.168^{*}$ | 0.178 | $0.469^{**}$ | $-0.116$ | $0.279^{***}$ | 0.252 | $0.368^{*}$ |
| | (0.131) | (0.087) | (0.192) | (0.224) | (0.110) | (0.096) | (0.166) | (0.220) |
| $\beta^{CMA}$ | 0.134 | $-0.046$ | $-0.069$ | $-0.204$ | 0.061 | $-0.128$ | 0.092 | 0.031 |
| | (0.174) | (0.115) | (0.255) | (0.298) | (0.146) | (0.128) | (0.221) | (0.292) |

Note: Table 1.8 provides time series regression of CRE sorted portfolios on existing factor regression. Panel A shows time series regression based on CAPM model with both equally weight and value weight. Panel B display the time series regression of the CRE sorted portfolio return on market, value, size and momentum factors with both equally weight and value weight. Panel C shows the time series regressions with CRE sorted portfolios on market, value, size, momentum, profitability and investment from Fama-Kenneth website. In the Table, 'high' refers to the portfolio with top 70th percentiles of CRE (sale, size) sorted stocks and 'low' refers to the portfolio with bottom 30th percentiles of CRE (sale, size) sorted stocks. 'L-H' implies a long-short trading strategy by longing the low (bottom 30%) CRE exposed stocks and shorting high (top 70%) CRE exposed stocks. All standard errors are adjusted with Newey–West method with 15 lags. $^{*}p$ <0.1; $^{**}p$ <0.05; $^{***}p$ <0.01.

## Table 1.9: Predictive Equity Return Regression

| | *Dependent variable: Cumulative Stock Return* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\frac{CAR(0,3)}{3}$ | | $\frac{CRR(0,6)}{6}$ | | $\frac{CRR(0,9)}{9}$ | | $\frac{CRR(0,12)}{12}$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $CRE_{i,t-1} \times MCCC_t$ | −6.882** | −7.865** | −8.324** | −8.465** | −6.301* | −6.587* | −5.979** | −6.199** |
| | (2.849) | (3.201) | (3.253) | (3.285) | (3.719) | (3.757) | (2.937) | (2.996) |
| $CRE_{i,t-1}$ | 6.036* | 7.572** | 7.760** | 8.246** | 5.687 | 6.191 | 5.987* | 6.369* |
| | (3.082) | (3.450) | (3.345) | (3.486) | (4.036) | (4.133) | (3.215) | (3.305) |
| $MCCC_t$ | 0.117* | 0.136* | 0.009 | 0.005 | 0.014 | 0.018 | 0.050*** | 0.055*** |
| | (0.065) | (0.068) | (0.020) | (0.020) | (0.018) | (0.019) | (0.013) | (0.014) |
| $Leverage_{i,t-1}$ | | 0.011* | | −0.001 | | 0.003 | | 0.009** |
| | | (0.006) | | (0.004) | | (0.005) | | (0.004) |
| $BM_{i,t-1}$ | | 0.012*** | | 0.007*** | | 0.006*** | | 0.006*** |
| | | (0.003) | | (0.002) | | (0.001) | | (0.001) |
| $Investment_{i,t-1}$ | | −0.013 | | −0.024 | | −0.022 | | −0.003 |
| | | (0.032) | | (0.017) | | (0.015) | | (0.014) |
| $Size_{i,t-1}$ | | 0.052* | | 0.047** | | −0.0003 | | −0.005 |
| | | (0.028) | | (0.020) | | (0.017) | | (0.016) |
| $Beta_{i,t-1}$ | | 0.101 | | 0.164** | | 0.110* | | 0.142** |
| | | (0.147) | | (0.077) | | (0.065) | | (0.066) |
| $MOM_{i,t-1}$ | | 0.004 | | 0.008*** | | 0.006** | | 0.005* |
| | | (0.002) | | (0.003) | | (0.002) | | (0.003) |
| $Vol_{i,t-1}$ | | −0.046** | | −0.064*** | | −0.034*** | | −0.023* |
| | | (0.017) | | (0.016) | | (0.013) | | (0.013) |
| $Complexity_{i,t}$ | | 0.048 | | −0.012 | | −0.179 | | −0.104 |
| | | (0.166) | | (0.118) | | (0.116) | | (0.130) |
| $Constraint_{i,t}$ | | −1.171*** | | −0.249 | | −0.170 | | −0.434* |
| | | (0.281) | | (0.201) | | (0.265) | | (0.219) |
| $Legislation_{i,t}$ | | −0.001 | | −0.015 | | −0.012 | | −0.127** |
| | | (0.105) | | (0.055) | | (0.051) | | (0.057) |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Adjusted $R^2$ | 0.164 | 0.183 | 0.119 | 0.139 | 0.096 | 0.106 | 0.079 | 0.090 |
| Observations | 9,051 | 8,944 | 9,546 | 9,361 | 11,429 | 11,166 | 11,854 | 11,524 |
| Adjusted $R^2$ | 0.164 | 0.183 | 0.119 | 0.139 | 0.096 | 0.106 | 0.079 | 0.090 |

Note: Table 1.9 shows the predictive regression with $CRE_{i,t-1} \times MCCC_t$ as in equation 58. Column (1) and (2) show the regressions with T=3, $\frac{1}{T}\sum_{j=1}^{T} R^i_{t+j-1,t+j} = \beta_0 + \beta_1 CRE_{i,t-1} \times MCCC_t + \beta_2 CRE_{i,t-1} + Control_{i,t-1} + \beta_3 MCCC_t + \epsilon_{i,t}$. The column (3) and (4) are regressions with $T = 6$. The column (5) and (6) are regressions with $T = 9$. The column (7) and (8) are regressions with $T = 12$. The control variables in these regressions are the same as Table 1.2. 'Industry FE' implies the fixed effect with Fama French 48 industries. 'Year FE' means the year level fixed effect. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.10: Fama-MacBeth Cross Sectional Regression for Asset Pricing Test

| | *Dependent variable: Stock Return* | | | | | |
|---|---|---|---|---|---|---|
| | $Return_{t,t+1}$ | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $CRE_{i,t}$ | $-1.448^{*}$ | $-1.451^{*}$ | $-1.374^{*}$ | $-1.423^{**}$ | $-1.397^{**}$ | $-1.344^{**}$ |
| | (0.829) | (0.842) | (0.821) | (0.657) | (0.654) | (0.669) |
| $Complexity_{i,t}$ | | 0.035 | | | | 0.021 |
| | | (0.076) | | | | (0.078) |
| $Constraint_{i,t}$ | | 0.087 | | | | $-0.096$ |
| | | (0.172) | | | | (0.167) |
| $Legislation_{i,t}$ | | $-0.031$ | | | | $-0.029$ |
| | | (0.050) | | | | (0.042) |
| $Vol_{i,t}$ | | | $-0.001$ | | $-0.002$ | $-0.006$ |
| | | | (0.015) | | (0.013) | (0.014) |
| $Beta_{i,t}$ | | | $0.136^{**}$ | | $0.135^{**}$ | $0.134^{**}$ |
| | | | (0.062) | | (0.061) | (0.061) |
| $MOM_{i,t}$ | | | $0.002^{*}$ | | | $0.002^{*}$ |
| | | | (0.001) | | | (0.001) |
| $Leverage_{i,t}$ | | | | 0.006 | 0.006 | 0.007 |
| | | | | (0.006) | (0.006) | (0.006) |
| $BM_{i,t}$ | | | | $0.003^{**}$ | $0.003^{*}$ | $0.003^{*}$ |
| | | | | (0.001) | (0.001) | (0.001) |
| $Investment_{i,t}$ | | | | $-0.017$ | $-0.017$ | $-0.016$ |
| | | | | (0.015) | (0.015) | (0.014) |
| $Size_{i,t}$ | | | | 0.002 | 0.007 | 0.004 |
| | | | | (0.038) | (0.032) | (0.032) |
| $\alpha$ | $0.019^{***}$ | $0.018^{***}$ | $0.017^{***}$ | $0.015^{***}$ | $0.014^{***}$ | $0.015^{***}$ |
| | (0.006) | (0.007) | (0.005) | (0.006) | (0.005) | (0.005) |
| Observations | 103,950 | 103,950 | 103,713 | 103,156 | 102,921 | 102,921 |
| $R^2$ | 0.145 | 0.147 | 0.155 | 0.160 | 0.167 | 0.170 |

Table 1.10 shows the Fama-MacBeth regression on $CRE_{i,t}$. The regression exploits the full sample from Jan, 2009 to Dec, 2015. The regression is running without the finance service and banking industries. The dependent variable is the firm-level stock return. The control variables in these cross-sectional regressions have the same definition as Table 1.2. All standard errors are adjusted with Newey–West method with 15 lags. $^{*}p$ <0.1; $^{**}p$ <0.05; $^{***}p$ <0.01.

Table 1.11: Carbon Emission Regression and Climate Regulatory Exposure

| | *Dependent variable: Carbon Emission* | | | | | |
|---|---|---|---|---|---|---|
| | $Scope^1_{i,t}$ | | $Scope^2_{i,t}$ | | $Scope^1_{i,t} + Scope^2_{i,t}$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $CRE_{i,t} \times MCCC_t$ | −6,816.244* | −6,955.845** | −584.228 | −559.455 | −7,400.473** | −7,515.301** |
| | (3,501.353) | (3,375.259) | (469.670) | (431.895) | (3,424.783) | (3,348.012) |
| $Leverage_{i,t}$ | | 207.148 | | −0.556 | | 206.593 |
| | | (185.044) | | (7.844) | | (184.178) |
| $BM_{i,t}$ | | 6.842 | | 0.933 | | 7.775 |
| | | (6.255) | | (0.677) | | (6.761) |
| $Investment_{i,t}$ | | −711.428** | | −86.853 | | −798.281** |
| | | (349.053) | | (81.914) | | (356.451) |
| $Size_{i,t}$ | | 64.695 | | 12.371 | | 77.065 |
| | | (318.497) | | (142.104) | | (376.477) |
| $Vol_{i,t}$ | | 32.572 | | −6.696 | | 25.876 |
| | | (39.405) | | (5.164) | | (40.628) |
| $Beta_{i,t}$ | | −268.259 | | 0.474 | | −267.786 |
| | | (186.129) | | (38.223) | | (199.202) |
| $MOM_{i,t}$ | | 1.309 | | −0.248 | | 1.061 |
| | | (5.886) | | (1.342) | | (6.772) |
| $Complexity_{i,t}$ | | −269.410 | | 135.476 | | −133.934 |
| | | (679.562) | | (187.306) | | (675.670) |
| $Legislation_{i,t}$ | | 876.284 | | 37.799 | | 914.083 |
| | | (1,040.607) | | (66.090) | | (1,068.536) |
| $Constraint_{i,t}$ | | 4,680.616 | | −453.631 | | 4,226.985 |
| | | (5,572.212) | | (649.461) | | (5,848.164) |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 39,705 | 39,473 | 39,705 | 39,473 | 39,705 | 39,473 |
| Adjusted $R^2$ | 0.960 | 0.960 | 0.888 | 0.888 | 0.958 | 0.958 |

Note: Table 1.11 shows the carbon emission regression on $CRE_{i,t} \times MCCC_t$ in equation 59. The control in these regressions are the same as previous definition in Table 1.2 'Industry FE' implies the fixed effect with Fama French 48 industries. 'Year FE' means the year level fixed effect. In column (1) and (2), the regressions are direct greenhouse gas emission ($Scope^1_{i,t}$) on $CRE_{i,t} \times MCCC_t$. Column (3) and (4) refer to the regressions with indirect greenhouse gas emission on $CRE_{i,t} \times MCCC_t$. The regressions in column (5) and (6) refer to the sum of greenhouse gas direct and indirect emission on $CRE_{i,t} \times MCCC_t$. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.12: Book-to-Market Ratio and CRE regression

| | *Dependent variable: Book-to-Market Ratio* | | | |
|---|---|---|---|---|
| | $log(BM_{i,t})$ | | | |
| | $Scope^1_{i,t}$ | | $Scope^2_{i,t}$ | |
| | (1) | (2) | (3) | (4) |
| $CRE_{i,t} \times MCCC_t$ | 50.807*** | 53.964*** | 48.142** | 51.801*** |
| | (17.426) | (17.358) | (20.625) | (18.678) |
| $Scope^j \times MCCC_t$ | −0.00001 | −0.00001 | | |
| | (0.00001) | (0.00001) | | |
| $Scope^j \times MCCC_t$ | | | 0.0003** | 0.0002 |
| | | | (0.0002) | (0.0001) |
| $CRE_{i,t}$ | −43.581 | −45.367 | 22.308 | −43.172 |
| | (29.876) | (30.179) | (33.345) | (31.097) |
| $MCCC_t$ | −0.062* | −0.066** | −0.074** | −0.074** |
| | (0.032) | (0.031) | (0.036) | (0.033) |
| $Scope^1$ | 0.00003*** | 0.00003*** | | |
| | (0.00001) | (0.00001) | | |
| $Scope^2$ | | | 0.0001 | −0.00003 |
| | | | (0.0003) | (0.0002) |
| Control | Yes | Yes | Yes | Yes |
| Industry FE | Yes | Yes | No | No |
| Year FE | Yes | Yes | Yes | Yes |
| Observations | 39,473 | 39,705 | 39,473 | 39,705 |
| Adjusted $R^2$ | 0.296 | 0.268 | 0.139 | 0.086 |

Note: Table 1.12 shows the Book-to-Market regression on $CRE_{i,t} \times MCCC_t$ in equation 60. Similar to Bolton and Kacperczyk (2021c), the dependent variable is the logarithmic of book-to-market ratio. The control in these regressions are the same as previous definition in Table 1.2. In column (1) and (2), 'Scope' refers to control with scope 1, direct carbon emission. These columns imply that whether we control the direct carbon emission effect or not. In column (3) and (4), 'Scope' denotes the indirect carbon emissions of scope 2. These columns indicates that whether we control the indirect carbon emission or not. 'Industry FE' implies the fixed effect with Fama French 48 industries. 'Year FE' means the year level fixed effect. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.




Figure 1.1: Cross-Industry Cumulative Raw Return from 2009 to 2015

Note: The figure shows the industry-level accumulative raw return from Jan, 2009 to Dec, 2015. The high industries refer to Bolton et al. (2020) top 4 highest carbon emission industries following 1st digit of SIC code including utility, gas, electricity and transportation. Others contain all the other industries excluding the high ones.

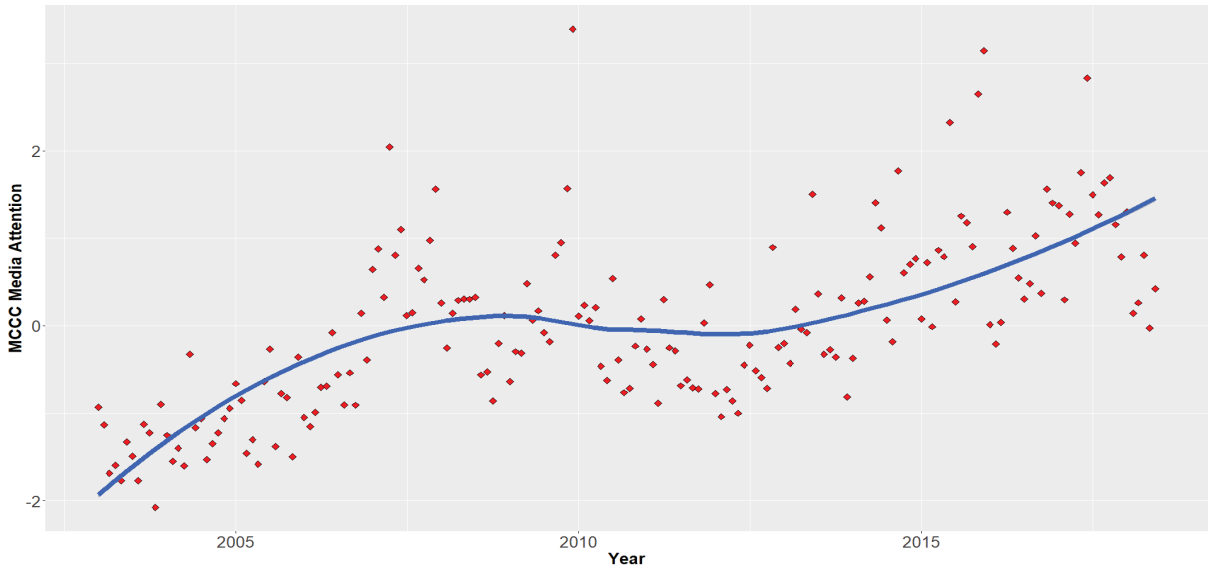



Figure 1.2: Media Climate Change Concern from 2000 to 2019

Note: The figure shows the industry-level accumulative raw return from Jan, 2009 to Dec, 2015. The high industries refer to Bolton et al. (2020) top 4 highest carbon emission industries following 1st digit of SIC code including utility, gas, electricity and transportation. Others contain all the other industries excluding the high ones. The points are fitted with smooth curve.

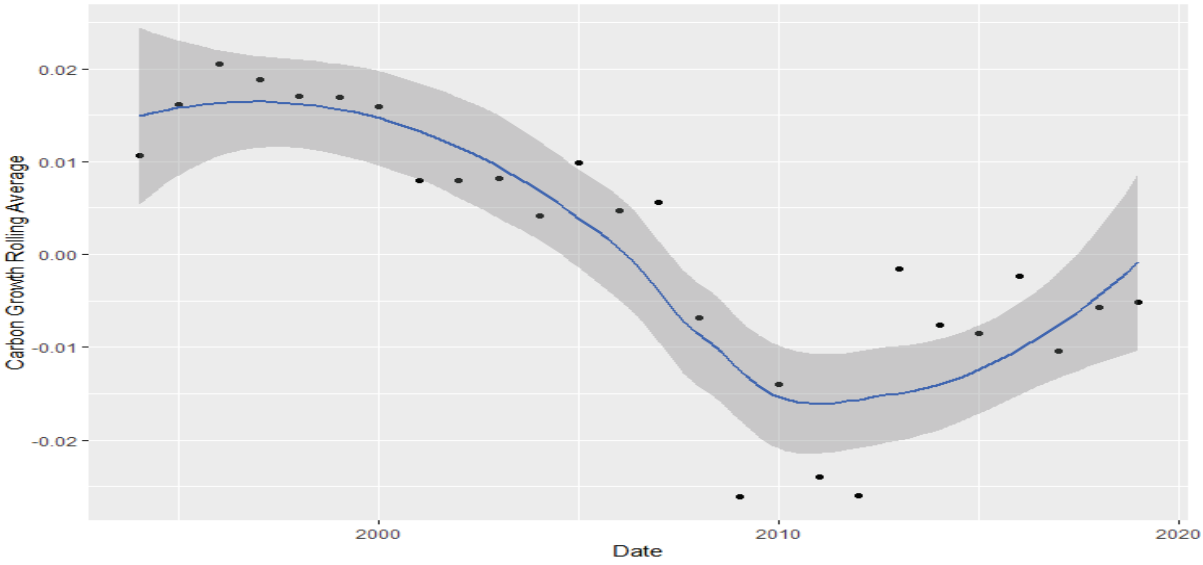



Figure 1.3: Rolling Average of Carbon Growth from 1990 to 2019

Note: Figure 1.3 denotes the rolling average of the U.S carbon growth from 1990 to 2019. The curve is drawn with 4-year rolling average carbon emission. The blue line refers to the fitted function and shaded area denotes the confidence interval of the fitted estimator.

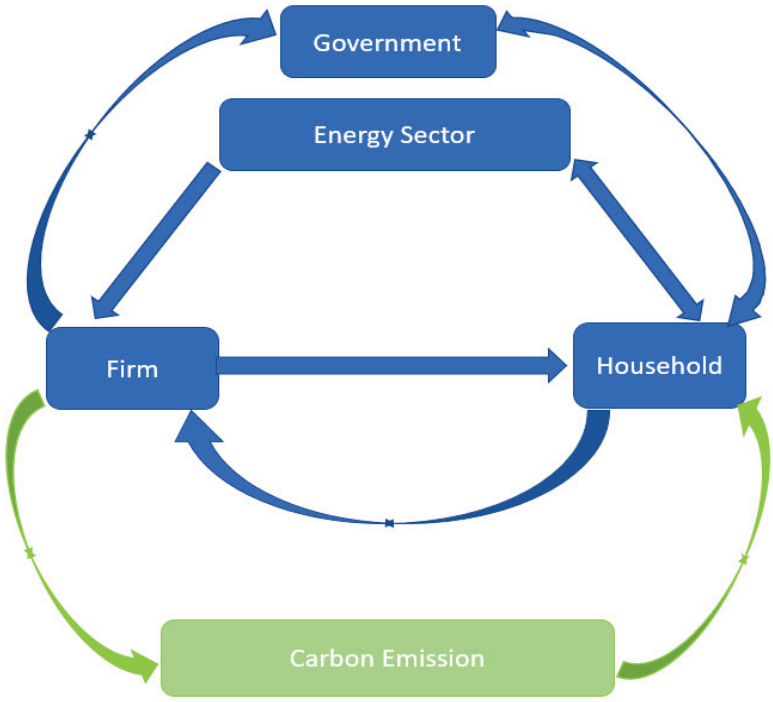



Figure 1.4: Multi-Sector Dynamic Model

Note: Figure 1.4 denotes the resource flow of agent in this economy. The arrows from energy refer to the facts that brown firms have dirty capital demand and energy profits go back to household. The arrows from firm imply that firms produce final goods for consumption, generate carbon emissions and pay for the government carbon taxation. The arrows from household indicate that households provide labor for production, purchase the equities from financial market and buy the government bond. The arrows from carbon emission mean that carbon emissions has negative externalities over household utility directly through investor awareness channel.

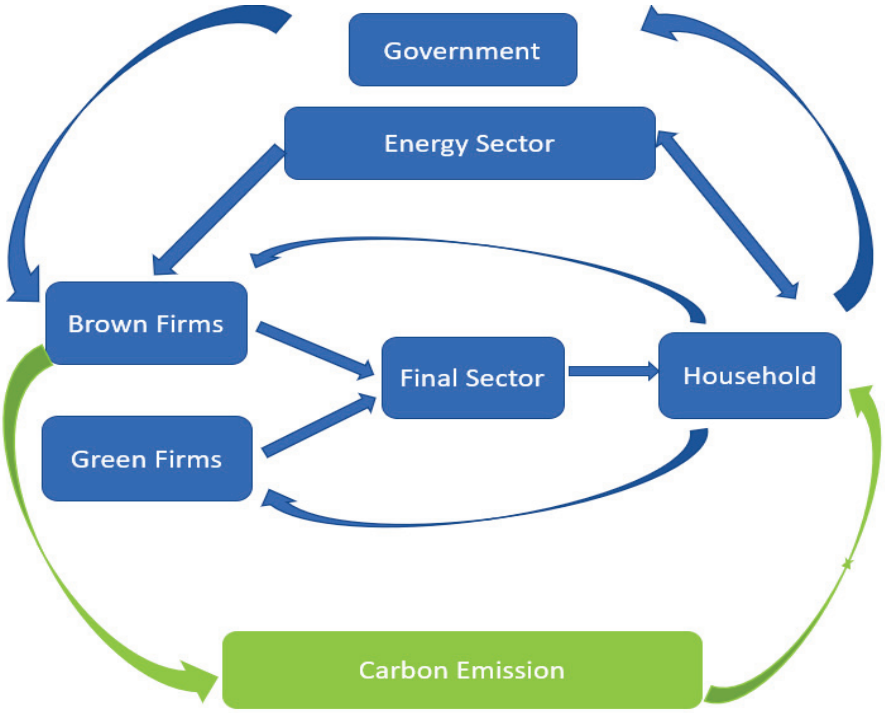



Figure 1.5: Multi-Sector Dynamic Model

Note: Figure 1.5 denotes the resource flow of agent in this economy. The arrows from energy refer to the facts that brown firms have dirty capital demand and energy profits go back to household. The arrows from brown firms denote that brown firms produce intermediate goods delivered to final good sector, generate carbon emissions and pay the carbon taxation to the government. The arrows from green firms denote that green firms also generate intermediate goods to final sector. The arrows from final sector implies that final good sector produces final goods for consumption. The arrows from household indicate that households provide labor for production, purchase the equities from financial market and buy the government bond. The arrows from carbon emission mean that carbon emissions has negative externalities over household utility directly through investor awareness channel.

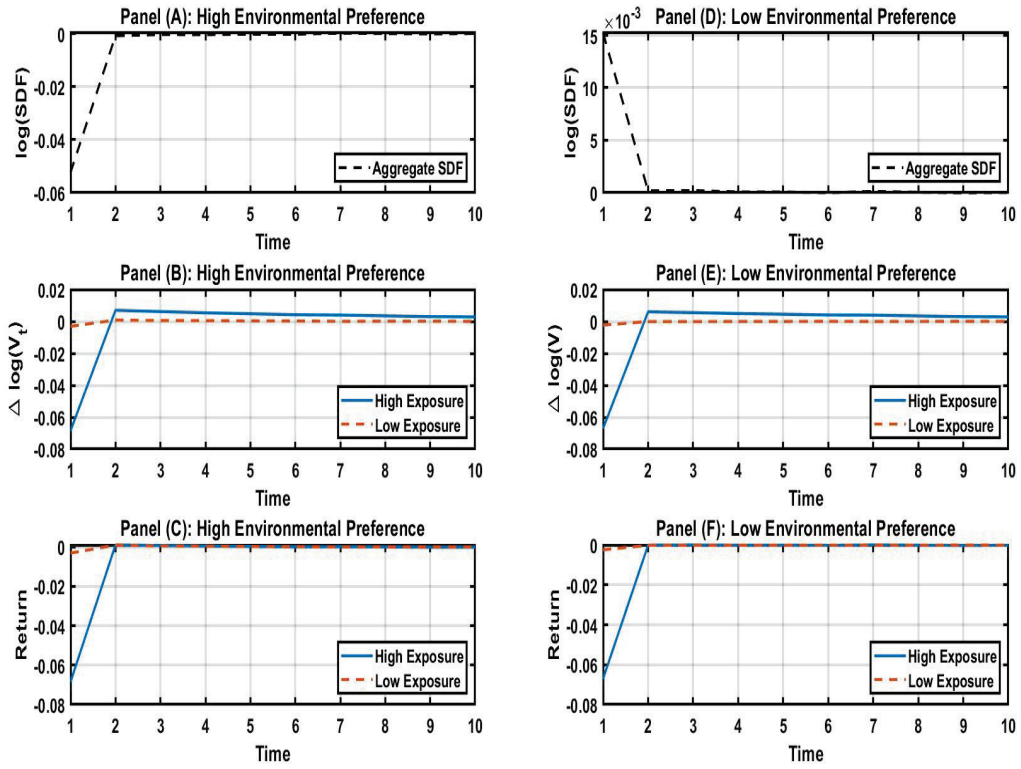



Figure 1.6: Impulse Response Function of Pricing Performance to Climate Transition Risk

Note: Figure 1.6 compares the impulse response function given different environmental preference. For the figures on the left columns with high environmental preference, the impulse response figure is drawn with $\Psi = 1$. Figures on the right is calibrated under the low environmental preference, $\Psi = 0$. $\Delta log(V_t)$ denotes the firm market valuation growth.

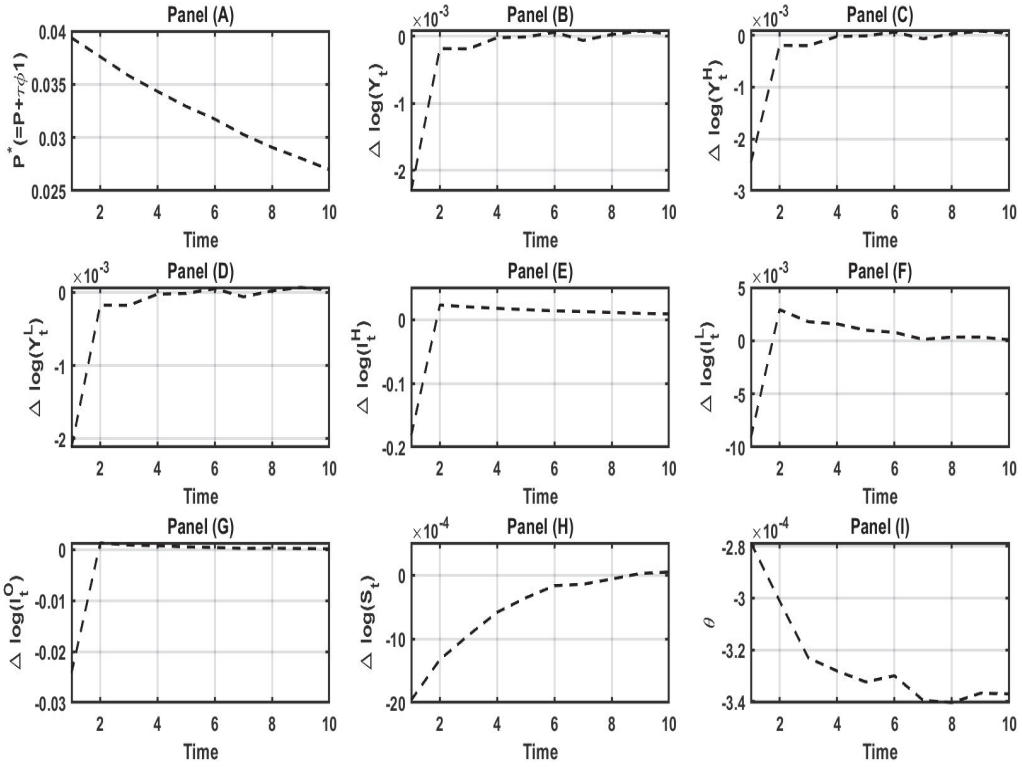
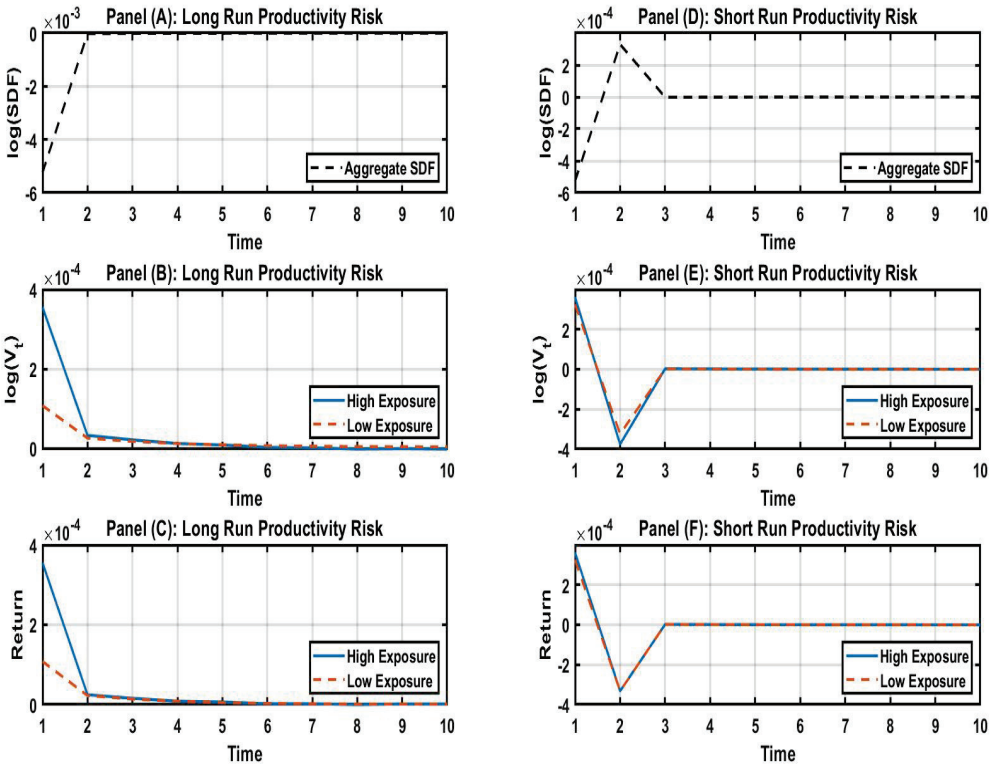



Figure 1.7: Fundamental Response to Climate Policy Shock

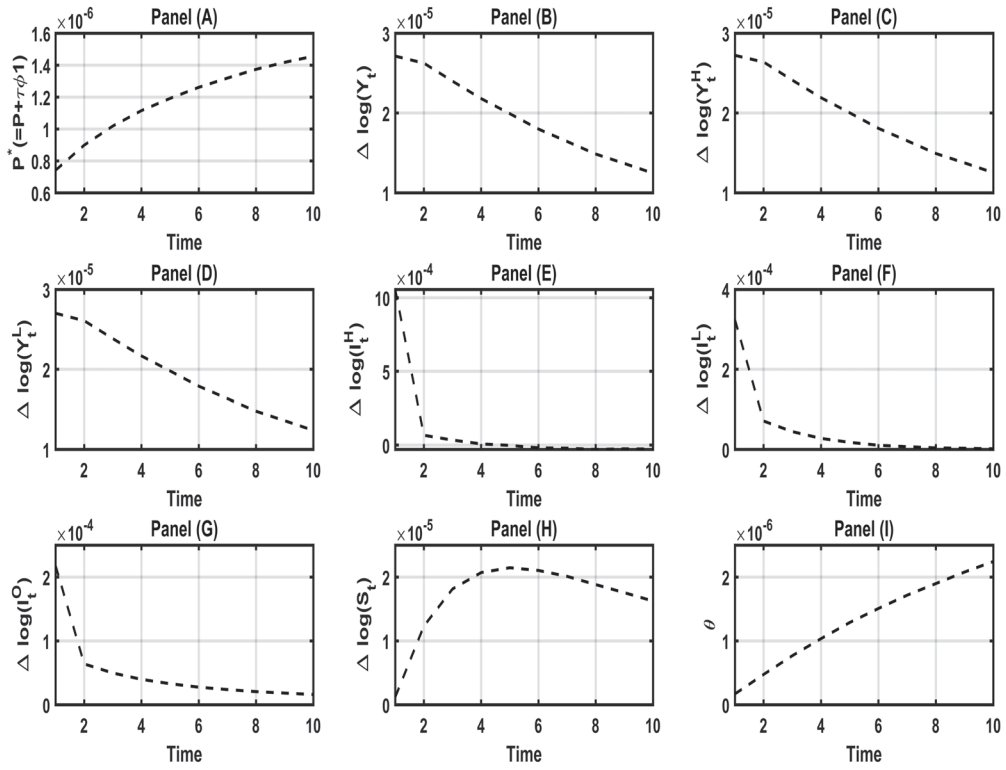
Note: Figure 1.7 displays the fundamental impulse response functions to the climate policy risk, $\epsilon_{w,t}$. $P^*$ refers to the climate tax included price. $\Delta log(Y^t)$ denotes the total output growth. $\Delta log(Y_t^B)$ refers to the high CRE output growth. $\Delta log(Y_t^G)$ is the output growth of the lower CRE firms. $\Delta log(I_t^B)$ denotes the high CRE exposed firm investment growth. $\Delta log(I_t^G)$ denotes the high CRE exposed firm investment growth. $\Delta log(I_t^O)$ denotes energy sector entrepreneur investment growth. $\Delta log(M_t)$ refers to the carbon accumulation growth. $\theta_t$ is the extraction rate of the energy supply.

Figure 1.8: Impulse Response to Productivity Risk (LRR vs SRR)



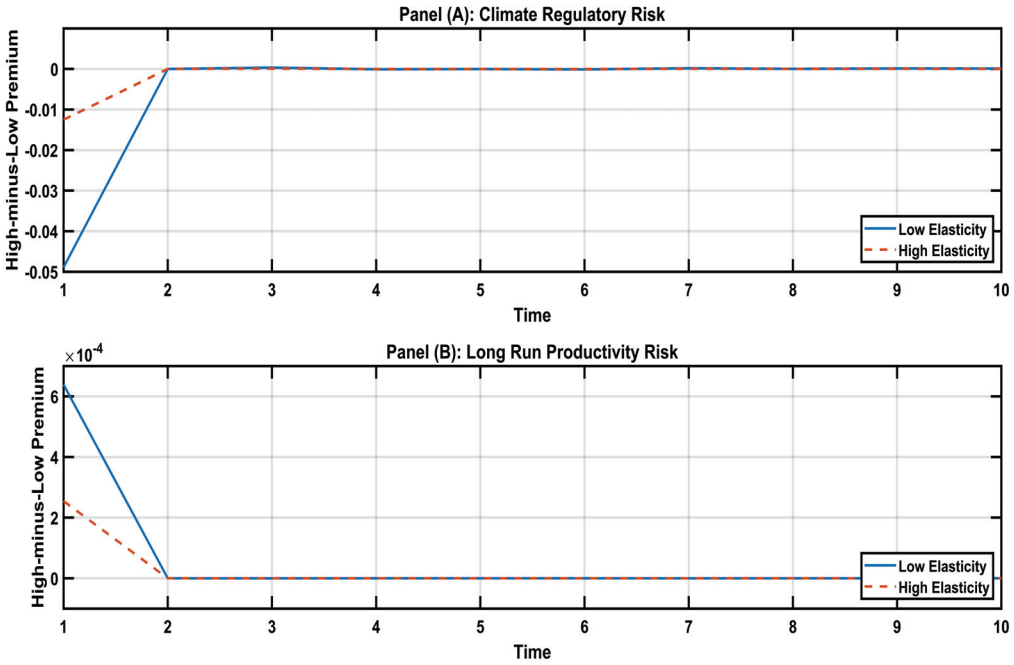

Note: Figure 1.8 denotes the impulse response function to productivity risk. The figures on the left column show the impulse response to positive long run productivity risk in equation 17 ($\epsilon_{z,t}$). The figures on the right show the positive response to short run risk 16 ($\epsilon_{a,t}$). $\Delta log(V_t)$ denotes the firm market valuation growth.

Figure 1.9: Impulse Response to Long Run Risk



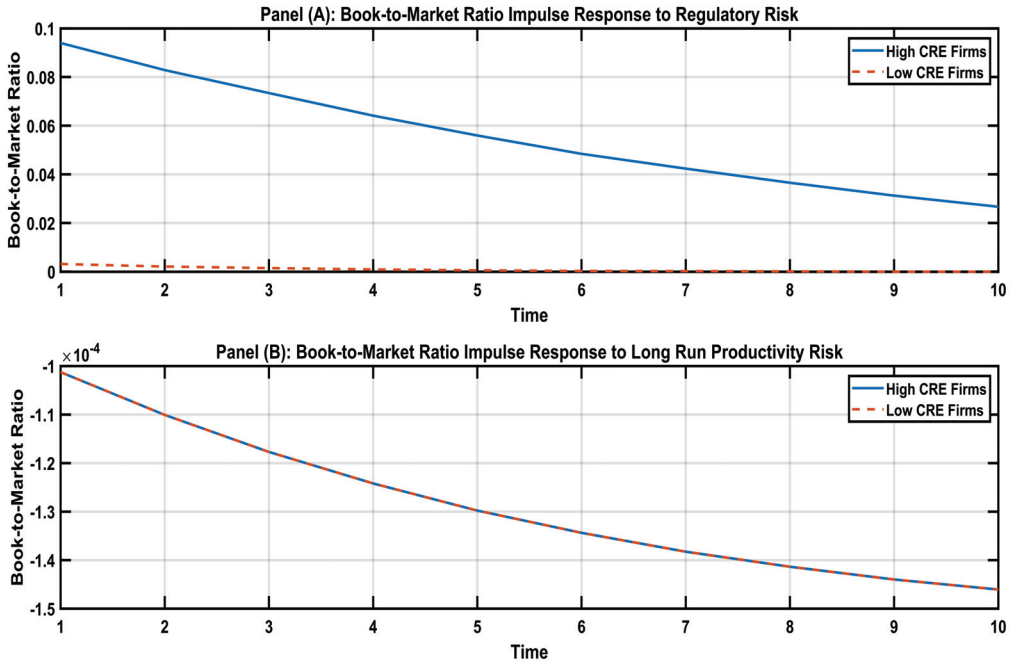

Note: Figure 1.9 displays the fundamental impulse response functions to the long run productivity risk, $\epsilon_{z,t} > 0$. $P^*$ refers to the climate tax included price. $\Delta log(Y^t)$ denotes the total output growth. $\Delta log(Y_t^B)$ refers to the high CRE output growth. $\Delta log(Y_t^G)$ is the output growth of the lower CRE firms. $\Delta log(I_t^B)$ denotes the high CRE exposed firm investment growth. $\Delta log(I_t^G)$ denotes the high CRE exposed firm investment growth. $\Delta log(I_t^O)$ denotes energy sector entrepreneur investment growth. $\Delta log(M_t)$ refers to the carbon accumulation growth. $\theta_t$ is the extraction rate of the energy supply.

Figure 1.10: Impulse Response Function and Elasticity of Substitution



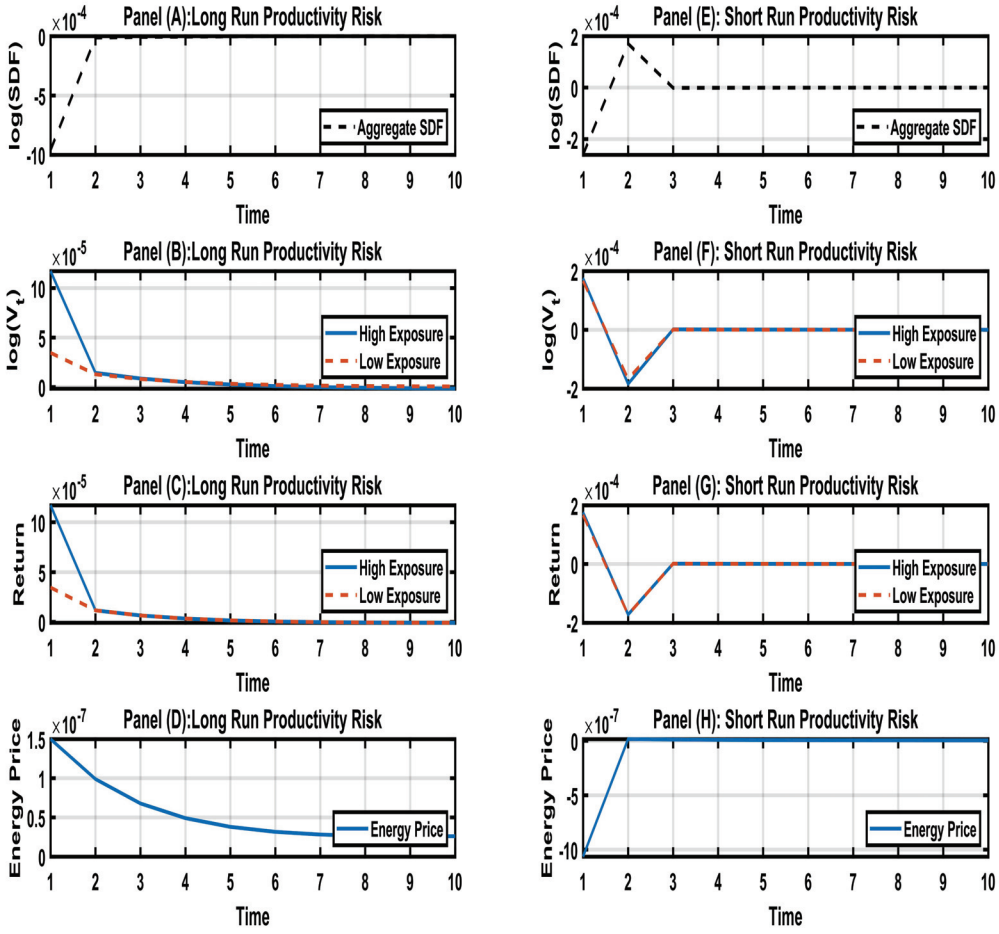

Note: Figure 1.10 displays the equity risk premium impulse response function to positive long-run productivity risk $\epsilon_{z,t} > 0$ and climate regulatory risk, $\epsilon_{w,t} > 0$. The y axis represents the equity premium with high exposed firm equity return minus low exposed firm equity return. The solid line refers to the benchmark economy with lower elasticity ($\sigma = 0.06$) between brown and green capital. The dot line shows the response with higher elasticity, $\sigma = 0.26$

Figure 1.11: Book-to-Market Impulse Response to Risks




Note: Figure 1.10 displays the Book-to-Market impulse response function to positive climate regulatory risk, $\epsilon_{w,t} > 0$ and long-run productivity risk $\epsilon_{z,t} > 0$. The y axis represents the Book-to-Market ratio for high CRE exposed firms and low CRE exposed firms respectively. The solid line refers to the firms with higher climate policy exposure. The dot line shows the response of firms with lower climate policy exposure.

Figure 1.12: Impulse Response to Productivity Risk with Endogenous Climate Policy




Note: Figure 1.12 shows impulse response to productivity risk with endogenous energy price in equation 54. The left figures show the response to the long-run productivity risk with $\epsilon_{z,t} > 0$ and the right-side figures display the result with short-run productivity risk, $\epsilon_{a,t} > 0$. $log(SDF)$ denotes the logarithmic of the stochastic discount factor. $\Delta log(V_t)$ refers to the growth of firm market valuations. Energy price is the price of energy capital in the market including the optimal tax from 54.

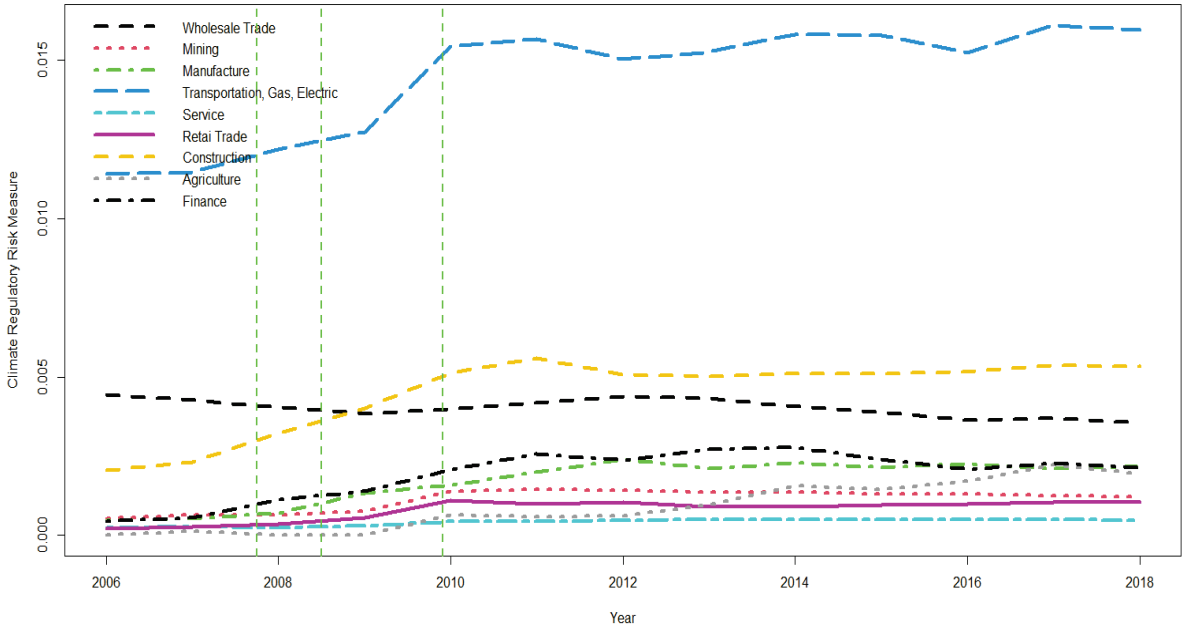
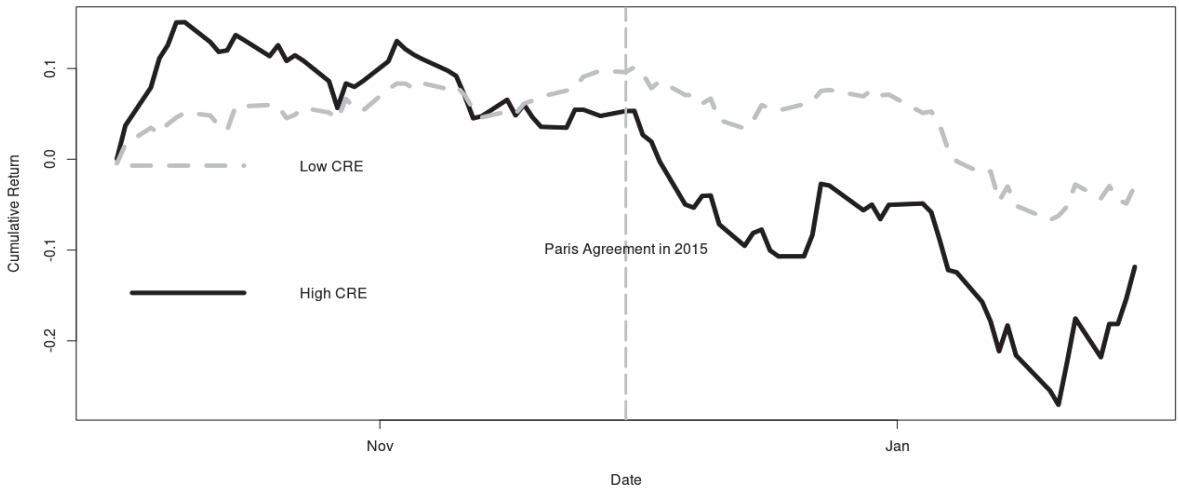



Figure 1.13: Cross-Industry Variation of the CRE

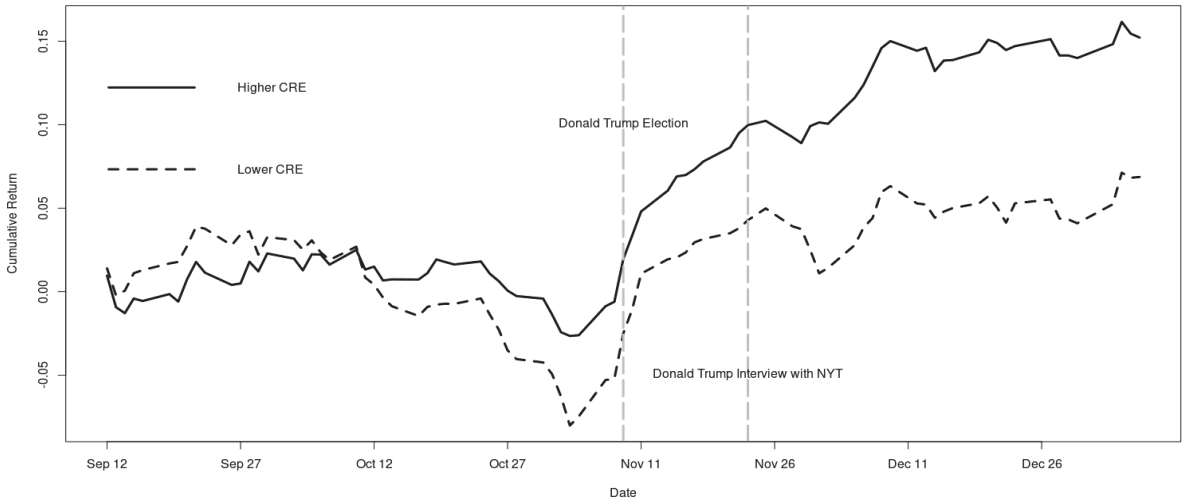
Note: The figure plot the CRE aggregate across industry level. The definitions of the industry in this figure is based on the first two digits of the 6 digit Standard Industrial Classification (SIC). The vertical lines stands for the timing that investors start to request for climate disclosure through SEC petitions.

Figure 1.14: Stock Return around Climate Regulatory Events: Paris Agreement




Note: Figure 1.14 is the figure for tracking cumulative daily stock return around a climate regulatory event, Paris Agreement in Dec, 2015. The solid line is the equally-weighted portfolio of stock returns with higher CRE exposure. The dotted line denotes a equally-weighted portfolio of stock returns with lower CRE exposure.

Figure 1.15: Stock Return around Climate Deregulatory Events: Trump Election

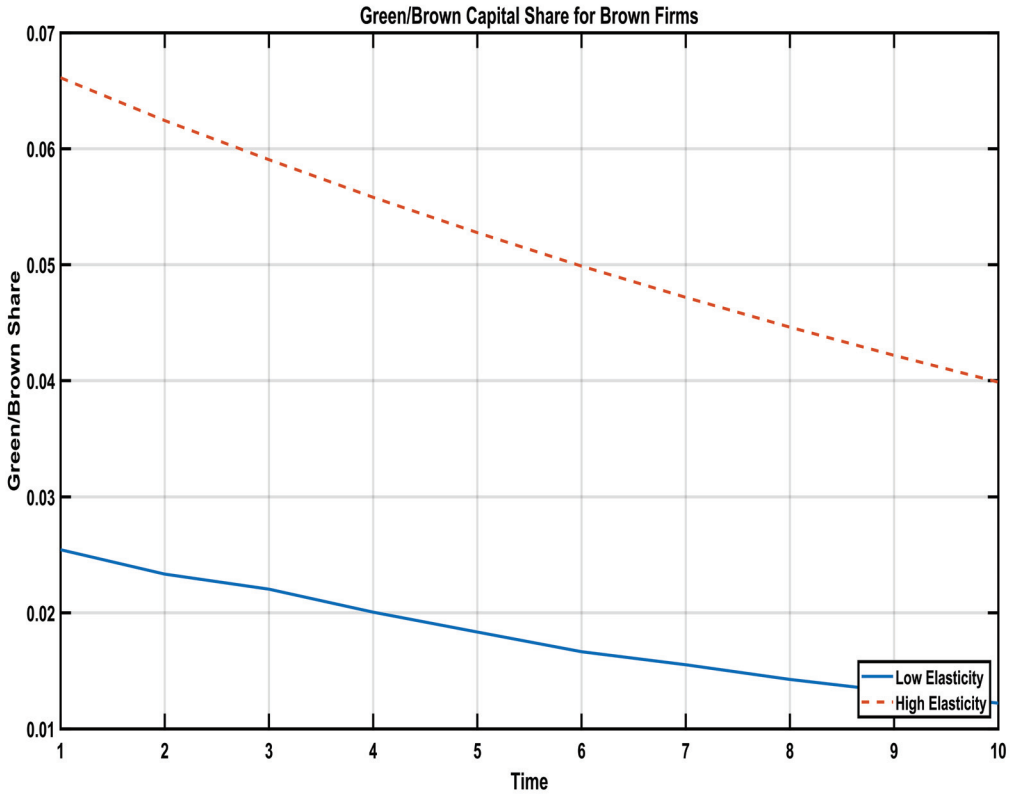



Note: Figure 1.15 is the figure for tracking cumulative daily stock return around a climate deregulatory event, Donald Trump election in Nov, 2016. The solid line is the equally-weighted portfolio of stock returns with higher CRE exposure. The dotted line denotes the equally-weighted portfolio of firm stock returns with lower CRE exposure.

## 1.10 Appendix A.

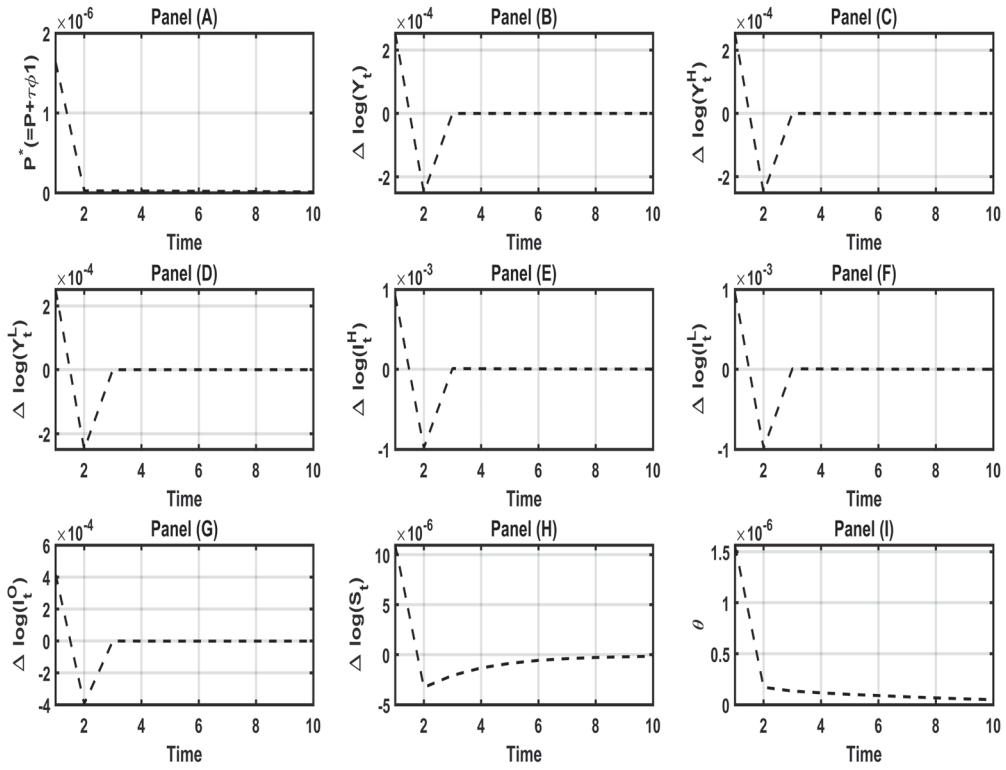
### 1.10.1 Figure

Figure 1.16: Fundamental Impulse Response to Climate Regulatory Risk



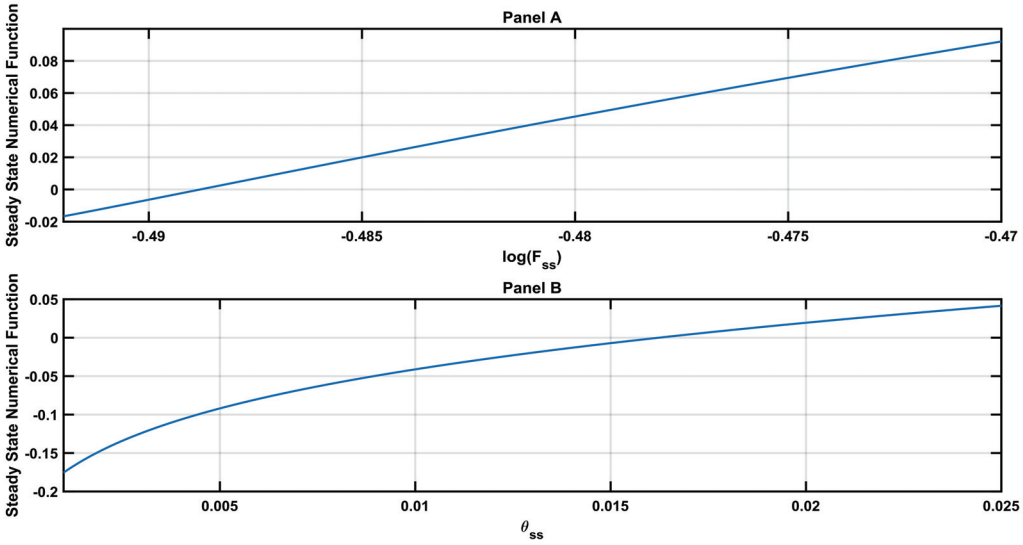

Note: Figure 1.16 displays the changes of Green-to-Brown share for brown firms in response to a shock from the climate regulatory risk. The blue line indicates the impulse response with $\sigma = 0.06$. The red dot line refers to the impulse response with $\sigma = 0.26$.

Figure 1.17: Fundamental Impulse Response to Short Run Productivity Risk



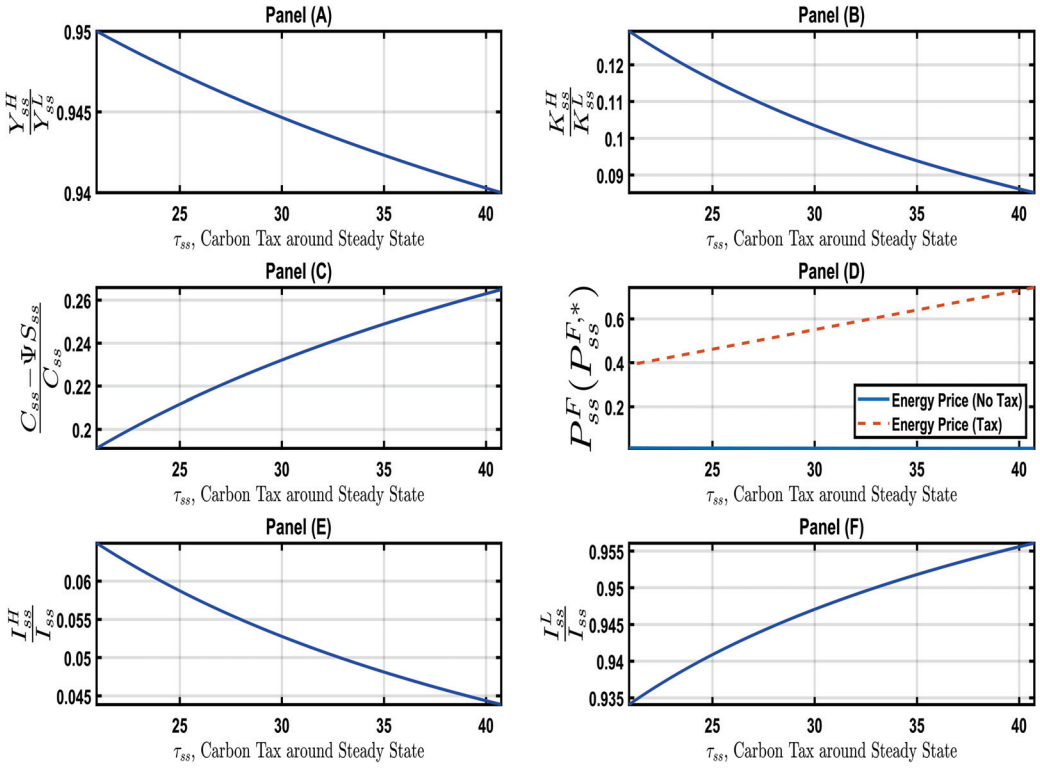

Note: Figure 1.17 displays the fundamental impulse response functions to the short run productivity risk, $\epsilon_{a,t} > 0$. $P^*$ refers to the climate tax included price. $\Delta log(Y^t)$ denotes the total output growth. $\Delta log(Y_t^B)$ refers to the high CRE output growth. $\Delta log(Y_t^G)$ is the output growth of the lower CRE firms. $\Delta log(I_t^B)$ denotes the high CRE exposed firm investment growth. $\Delta log(I_t^G)$ denotes the low CRE exposed firm investment growth. $\Delta log(I_t^O)$ denotes energy sector entrepreneur investment growth. $\Delta log(M_t)$ refers to the carbon accumulation growth. $\theta_t$ is the extraction rate of the energy supply.

Figure 1.18: Numerical Solutions for Determinant State Variables: $log(F_{ss})$ and $\theta_{ss}$




Note: Figure 1.18 displays the core steady state determinant variables for energy extraction, $\theta_{ss}$ and brown capital at steady state, $log(F_{ss})$. The figure above shows the determinant function for steady state , $log(F_{ss})$, where horizontal line with y=0 is the steady state point. The figure below displays the determinant function for $\theta_{ss}$, where horizontal line y=0 denotes the steady state point.

Figure 1.19: Fundamental and Realized Carbon Price at Steady State




Note: Figure 1.18 denotes the values at the steady states. The x-axis refers to the carbon intensity around the steady state, $\tau_{ss} := \tau^* * \frac{exp(w_{ss})}{exp(w_{ss}+1)}$. $\frac{Y_{ss}^B}{Y_{ss}^G}$ refers to the ratio of high transition risk exposed firms to low transition risk exposed firms. $\frac{C_{ss}-\Psi S_{ss}}{C_{ss}}$ denotes as the consumption surplus given $\Psi$. $\frac{K_{ss}^B}{K_{ss}^G}$ stands for the green capital ratio between high against low CRE exposed firms. $\frac{I_{ss}^B}{I_{ss}}$ refers to the ratio of high exposed. investment and total physical investment (including energy sector investment). $\frac{I_{ss}^G}{I_{ss}}$ is the fraction of low CRE exposed firm investment over total physical investment. $P^{F*}$ refers to the energy price of $F_t$ at steady state including the climate taxation. $P_{ss}^F$ refers to the energy capital price without climate taxation.

Table 1.13: Two-Way Portfolio Sorted Result

Panel A: Double Sorted Portfolios: Investment & CRE (VW)

| | *Dependent variable: Value-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Complexity* | | | *High Complexity* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.017*** | 0.010* | -0.006* | 0.017*** | 0.010* | -0.007** |
| | (0.005) | (0.006) | (0.004) | (0.005) | (0.006) | (0.003) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Panel B: Double Sorted Portfolios: Complexity & CRE (EW)

| | *Dependent variable: Equally-Weighted Portfolio Return* | | | | | |
|---|---|---|---|---|---|---|
| | *Low Complexity* | | | *High Complexity* | | |
| | Low CRE | High CRE | H-L | Low CRE | High CRE | H-L |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Portfolio Return | 0.020*** | 0.014** | -0.006** | 0.022*** | 0.010 | -0.012*** |
| | (0.006) | (0.006) | (0.003) | (0.007) | (0.008) | (0.005) |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |

Table A.1.13 shows the double sorting portfolio result with CRE. Panel B shows the result of double sorting with text complexity and climate regulatory exposure (CRE). The portfolio is balanced annually with equally-weighting scheme. Panel A displays the double sorting with both text complexity and CRE. The portfolio is balanced every year with value weights. In the Table, 'high' refers to the portfolio with top 70th percentiles of CRE (sale, size) sorted stocks and 'low' refers to the portfolio with bottom 30th percentiles of CRE (sale, size) sorted stocks. 'L-H' implies a long-short trading strategy by longing the low (bottom 30%) CRE exposed stocks and shorting high (top 70%) CRE exposed stocks. All standard errors are adjusted with Newey–West method with 15 lags. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.14: Predictive Equity Return Regression Controlling Direct Carbon Emission

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\frac{CAR(0,3)}{3}$ | | $\frac{CRR(0,6)}{6}$ | | $\frac{CRR(0,9)}{9}$ | | $\frac{CRR(0,12)}{12}$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $CRE_{i,t-1} \times MCCC_t$ | −12.780** | −13.129** | −6.859*** | −7.040*** | −8.981 | −8.968 | −10.827** | −10.772* |
| | (5.154) | (5.388) | (2.470) | (2.522) | (5.420) | (5.609) | (5.226) | (5.354) |
| $MCCC_t \times Scope^2_{i,t}$ | 0.00001 | 0.00001 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 | 0.00000 |
| | (0.00001) | (0.00001) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) |
| $CRE_{i,t-1}$ | 11.546** | 12.307** | 6.240** | 6.657** | 9.057 | 9.304 | 11.602** | 11.879* |
| | (5.040) | (5.554) | (2.356) | (2.564) | (5.885) | (6.318) | (5.762) | (6.113) |
| $MCCC_t$ | 0.137 | 0.132 | 0.014 | 0.016 | 0.010 | 0.013 | 0.056*** | 0.056*** |
| | (0.102) | (0.104) | (0.031) | (0.032) | (0.022) | (0.024) | (0.014) | (0.014) |
| $Scope^1_{i,t}$ | −0.00000 | −0.00001 | −0.00000 | −0.00000 | −0.00001* | −0.00001 | −0.00001 | −0.00000 |
| | (0.00001) | (0.00001) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) | (0.00000) |
| Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2,751 | 2,731 | 2,905 | 2,873 | 3,454 | 3,409 | 3,587 | 3,533 |
| Adjusted $R^2$ | 0.259 | 0.267 | 0.212 | 0.221 | 0.183 | 0.183 | 0.113 | 0.118 |

Note: Table 1.9 shows the predictive regression with $CRE_{i,t-1} \times MCCC_t$ as in equation 58. Column (1) and (2) show the regressions with T=3, $\frac{1}{T}\sum_{j=1}^{T} R^i_{t+j-1,t+j} = \beta_0 + \beta_1 CRE_{i,t-1} \times MCCC_t + \beta_2 CRE_{i,t-1} + Control_{i,t-1} + \beta_3 MCCC_t + \epsilon_{i,t}$. The column (3) and (4) are regressions with $T = 6$. The column (5) and (6) are regressions with $T = 9$. The column (7) and (8) are regressions with $T = 12$. The control variables in these regressions are the same as Table 1.2. 'Scope 1' refers to the direct carbon emissions from Trucost Database. 'Industry FE' implies the fixed effect with Fama French 48 industries. 'Year FE' means the year level fixed effect. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.15: Predictive Equity Return Regression Controlling Indirect Carbon Emission

| | *Dependent variable: Average Future Return* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $\frac{CAR(0,3)}{3}$ | | $\frac{CRR(0,6)}{6}$ | | $\frac{CRR(0,9)}{9}$ | | $\frac{CRR(0,12)}{12}$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $CRE_{i,t-1} \times MCCC_t$ | −12.345** | −12.669** | −6.739*** | −6.942*** | −8.495 | −8.525 | −10.384* | −10.358* |
| | (5.202) | (5.463) | (2.356) | (2.416) | (5.399) | (5.561) | (5.231) | (5.349) |
| $MCCC_t \times Scope^2_{i,t}$ | 0.00002 | 0.00001 | −0.00001 | −0.00002 | 0.00002 | 0.00002 | 0.00001 | 0.00001 |
| | (0.00004) | (0.00005) | (0.00002) | (0.00002) | (0.00002) | (0.00002) | (0.00001) | (0.00001) |
| $CRE_{i,t-1}$ | 11.104** | 11.805** | 6.107** | 6.538** | 8.551 | 8.853 | 11.100* | 11.445* |
| | (5.106) | (5.646) | (2.274) | (2.487) | (5.875) | (6.290) | (5.736) | (6.089) |
| $MCCC_t$ | 0.137 | 0.133 | 0.015 | 0.017 | 0.009 | 0.012 | 0.056*** | 0.055*** |
| | (0.103) | (0.105) | (0.031) | (0.032) | (0.023) | (0.024) | (0.014) | (0.014) |
| $Scope^2_{i,t}$ | −0.00002 | 0.00000 | 0.00001 | 0.00002 | −0.00002 | −0.00002 | −0.00002 | −0.00002 |
| | (0.00004) | (0.0001) | (0.00002) | (0.00002) | (0.00002) | (0.00002) | (0.00001) | (0.00001) |
| Control | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 2,751 | 2,731 | 2,905 | 2,873 | 3,454 | 3,409 | 3,587 | 3,533 |
| Adjusted $R^2$ | 0.259 | 0.267 | 0.212 | 0.221 | 0.182 | 0.182 | 0.112 | 0.118 |

Note: Table 1.9 shows the predictive regression with $CRE_{i,t-1} \times MCCC_t$ as in equation 58. Column (1) and (2) show the regressions with T=3, $\frac{1}{T}\sum_{j=1}^{T} R^i_{t+j-1,t+j} = \beta_0 + \beta_1 CRE_{i,t-1} \times MCCC_t + \beta_2 CRE_{i,t-1} + Control_{i,t-1} + \beta_3 MCCC_t + \epsilon_{i,t}$. The column (3) and (4) are regressions with $T = 6$. The column (5) and (6) are regressions with $T = 9$. The column (7) and (8) are regressions with $T = 12$. The control variables in these regressions are the same as Table 1.2. 'Scope 2' refers to the indirect carbon emissions from Trucost Database. 'Industry FE' implies the fixed effect with Fama-French 48 industries. 'Year FE' means the year level fixed effect. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.16: Climate Attention and Market Price of Climate Regulation

| | *Dependent variable: Time-Varying Market Price* | | | | | |
|---|---|---|---|---|---|---|
| | $\lambda_t^{CRE}$ | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $MCCC_t$ | −13.136* | −13.956* | −13.354* | −9.542* | −9.709* | −11.024** |
| | (6.938) | (7.024) | (6.860) | (5.419) | (5.390) | (5.480) |
| $Water.Drought_t$ | 6.189 | 6.500 | 6.185 | 3.113 | 3.089 | 4.057 |
| | (3.968) | (4.017) | (3.923) | (3.099) | (3.083) | (3.134) |
| $Forests_t$ | 6.171 | 6.383 | 6.208 | 5.295 | 5.321 | 5.381 |
| | (4.139) | (4.190) | (4.092) | (3.233) | (3.216) | (3.269) |
| $Carbon.Credits.Market_t$ | 6.014 | 6.252 | 6.114 | 6.643* | 6.627* | 6.914* |
| | (4.553) | (4.609) | (4.502) | (3.556) | (3.537) | (3.596) |
| $Loan_t$ | 0.006 | 0.006 | 0.006 | 0.005 | 0.005 | 0.005 |
| | (0.018) | (0.018) | (0.018) | (0.014) | (0.014) | (0.014) |
| $Consumption_t$ | 0.007 | 0.006 | 0.007 | 0.004 | 0.005 | 0.004 |
| | (0.012) | (0.012) | (0.011) | (0.009) | (0.009) | (0.009) |
| $Inflation_t$ | 1.669 | 2.169 | 1.509 | 2.203 | 2.156 | 2.667 |
| | (3.175) | (3.214) | (3.139) | (2.480) | (2.467) | (2.508) |
| $VIX_t$ | 0.596 | 0.647 | 0.570 | 0.509 | 0.491 | 0.496 |
| | (0.914) | (0.926) | (0.904) | (0.714) | (0.710) | (0.722) |
| $Car.Industry_t$ | 2.124 | 2.665 | 2.267 | −0.811 | −0.682 | 0.112 |
| | (4.178) | (4.230) | (4.131) | (3.263) | (3.246) | (3.300) |
| $Production_t$ | −1.548 | −1.479 | −1.633 | −0.714 | −0.798 | −0.813 |
| | (1.526) | (1.545) | (1.509) | (1.192) | (1.186) | (1.205) |
| *Constant* | 121.885 | 115.688 | 131.142 | 44.973 | 53.007 | 54.977 |
| | (143.884) | (145.667) | (142.262) | (112.381) | (111.787) | (113.654) |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 84 | 84 | 84 | 84 | 84 | 84 |
| Adjusted $R^2$ | −0.064 | −0.057 | −0.059 | −0.031 | −0.032 | −0.019 |

Note: Table 1.16 shows the time varying market price of $CRE_{i,t}$ regression on $MCCC_t$. The $\lambda_t$ is constructed with conditional cross-sectional regression following, $E_t(R_{i,t+1}) - R_{f,t+1} = \beta_{i,0} + \beta_{i,1}\lambda_t^{CRE} + Control_{i,t} + \epsilon_{i,t}$. In Ardia et al. (2020), $Water.Drought$ denotes the text-implied water drought variable. $Forest_t$ refers to the text-implied forest-topic variable. $Carbon.Credit.Market_t$ represents the topic variable of carbon credit market . $Loan_t$ means the media-implied loan topic variable. $Car.Industry$ denotes a unstructured text variable on car industry. For quantitative aggregate variables from Fred website, $Consumption_t$ denotes the US aggregate consumption. $Inflation_t$ refers to the market inflation. $VIX_t$ denotes the market implied volatility of the SP products. $Production_t$ is the aggregate output in the US. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 1.17: Carbon Emission and Climate Regulatory Exposure without Industry Fixed Effect

| | *Dependent variable: Carbon Emission* | | | | | |
|---|---|---|---|---|---|---|
| | $Scope^1$ | | $Scope^2$ | | $Scope^1 + Scope^2$ | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $CRE_{i,t} \times MCCC_t$ | 112,734.600* | 97,227.270** | 4,069.012 | 3,840.494 | 116,803.600* | 101,067.800** |
| | (62,126.690) | (42,564.530) | (2,675.764) | (2,635.873) | (62,899.240) | (42,813.420) |
| $Leverage_{i,t}$ | | 526.646 | | 30.368 | | 557.015 |
| | | (439.129) | | (24.830) | | (439.321) |
| $BM_{i,t}$ | | 59.648 | | 1.815 | | 61.463 |
| | | (55.398) | | (2.527) | | (56.905) |
| $Investment_{i,t}$ | | 120.784 | | 162.350** | | 283.135 |
| | | (773.229) | | (79.884) | | (813.014) |
| $Size_{i,t}$ | | 4,344.588 | | 834.535 | | 5,179.123 |
| | | (4,096.892) | | (553.843) | | (4,600.551) |
| $Vol_{i,t}$ | | −1,184.658 | | −83.439* | | −1,268.097 |
| | | (824.838) | | (47.914) | | (824.087) |
| $Beta_{i,t}$ | | −6,085.047** | | −106.389 | | −6,191.436** |
| | | (3,069.138) | | (185.679) | | (3,096.473) |
| $MOM_{i,t}$ | | 27.742 | | −3.606 | | 24.136 |
| | | (32.127) | | (2.408) | | (33.092) |
| $Complexity_{i,t}$ | | −9,083.412 | | 138.825 | | −8,944.587 |
| | | (6,073.223) | | (382.027) | | (6,032.396) |
| $Lesgislation$ | | −226.204 | | −194.381 | | −420.585 |
| | | (3,339.481) | | (402.268) | | (3,593.763) |
| $Constraint$ | | 25,144.920 | | −3,080.186** | | 22,064.730 |
| | | (26,977.020) | | (1,378.356) | | (27,465.870) |
| Industry FE | No | No | No | No | No | No |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 39,705 | 39,473 | 39,705 | 39,473 | 39,705 | 39,473 |
| Adjusted $R^2$ | 0.047 | 0.100 | 0.008 | 0.142 | 0.047 | 0.110 |

Note: Table 1.17 shows the carbon emission regression on $CRE_{i,t} \times MCCC_t$ in equation 59 without industry level fixed effect. The control in these regressions are the same as previous definition in Table 1.2 'Industry FE' implies the fixed effect with Fama French 48 industries. 'Year FE' means the year level fixed effect. In column (1) and (2), the regressions are direct greenhouse gas emission ($Scope^1_{i,t}$) on $CRE_{i,t} \times MCCC_t$. Column (3) and (4) refer to the regressions with indirect greenhouse gas emission on $CRE_{i,t} \times MCCC_t$. The regressions in column (5) and (6) refer to the sum of greenhouse gas direct and indirect emission on $CRE_{i,t} \times MCCC_t$. All standard errors are clustered at the industry level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

### 1.10.2 Proof 1: Two-Period Model Solution

Follow the basic settings of the two-period model, the key is to solve the brown capital procedure following the equilibrium conditions. The supply of brown capital follows the FOC:

$$P_1 = \pi_1 \frac{J_1}{\bar{R}}$$

Since the aggregate shock at period 0 is realized at time 1, $R_1(A_1, \tau_1)$ is realized as expected return conditional on the realization of $A_1$. The demand of brown capital FOC follows:

$$A_1 F_1^{\alpha-1} \alpha = P_1 + \tau\phi_1$$

With brown capital market clearing condition, we get $F_1 = J_1$ and the partial derivative w.r.t $\tau$ with $\phi_1 > 0$ and $\alpha \in (0, 1)$:

$$\frac{\partial F_1}{\partial \tau} = \frac{\phi_1}{A_1\alpha(\alpha-1)F^{\alpha-2} - \frac{\pi_1}{R_1}} < 0$$

In addition, there is:

$$\frac{\partial F_1}{\partial a} = \frac{e^a (F_1^{ss})^{\alpha-1}\alpha}{\pi_1/\bar{R} + \alpha(1-\alpha)e^{\mu_a}(F_1^{ss})^{\alpha-2}} > 0$$

. With market clearing condition for goods market,

$$C_1 = Y_1 - G(F_1; \bar{R})$$

The environment adjusted consumption is:

$$\begin{aligned} \tilde{C_1} &= Y_1 - G(F_1; \bar{R}) - \Phi S_1 \\ &= A_1 F_1^{\alpha} - G(F_1; \bar{R}) - \Phi\phi_1 F_1 \end{aligned} \tag{1.1}$$

From the Combined with the $\partial F_1/\partial\tau$ and the market clearing condition above:

$$\frac{\partial \tilde{C_1}}{\partial \tau} = \phi_1(\tau - \Phi)\frac{\partial F_1}{\partial \tau} = \begin{cases} < 0 & \text{when } \Phi < \tau_1 \\ > 0 & \text{when } \Phi > \tau_1 \end{cases}$$

where the sign of $\frac{\partial \tilde{C_1}}{\partial \tau}$ is dependent on the relative relationship between realized climate tax burden $\tau$ and preference sensitivity parameter, $\Phi$. Similarly, the marginal consumption at the market clearing decreases with higher realized climate tax risk at time 1, $\frac{\partial C_1}{\partial \tau_1} = \tau\phi_1 \frac{\partial F_1}{\partial \tau_1} < 0$. Define the brown capital price as $P_1^* \equiv P_1 + \tau\phi_1$. The tax-included brown capital price is increasing with a higher realized carbon price, $\frac{\partial P_1^*}{\partial \tau} = \frac{\partial^2 Y_1}{\partial^2 F_1}\frac{\partial F_1}{\partial \tau} > 0$, since $\alpha \in (0, 1)$.

### 1.10.3 Proof 2:

**Proof.** The proposition can be proved with first-order Taylor Expansion:

$$\begin{aligned} M_1 = e^{m_1} &\approx (\frac{\tilde{C}_1^{ss}}{\tilde{C}_0})^{-1} e^{-\frac{1}{\tilde{C}_1^{ss}}\frac{\partial \tilde{C}_1^{ss}}{\partial a}(a-\mu_a) - \frac{1}{\tilde{C}_1^{ss}}\frac{\partial \tilde{C}_1^{ss}}{\partial \tau}(\tau-\mu_\tau)} \\ &\approx (\frac{\tilde{C}_1^{ss}}{\tilde{C}_0})^{-1}(1 - \frac{1}{\tilde{C}_1^{ss}}\frac{\partial \tilde{C}_1^{ss}}{\partial a}(a-\mu_a) - \frac{1}{\tilde{C}_1^{ss}}\frac{\partial \tilde{C}_1^{ss}}{\partial \tau}(\tau-\mu_\tau)) \end{aligned}$$

where $(\frac{\partial \tilde{C}_1}{\partial a})_{ss} = e^{\mu_a}F_{ss}^{\alpha} + \frac{\partial F_1}{\partial a}\phi_1(e^{\mu_\tau} - \Phi)$ and $\frac{\partial \tilde{C}_1^{ss}}{\partial \tau} = (e^{\mu_\tau} - \Phi)\phi_1\frac{\partial F_1}{\partial \tau}$ ■

Similar to the proxy for SDF, the first-order approximation for realized return follows:

$$R_1 = e^{r_1} \approx \frac{D_1^{ss}}{P_0}(1 + \frac{1}{D_{ss}^1}\frac{\partial D_1}{\partial a}(a-\mu_a) + \frac{1}{D_{ss}^1}\frac{\partial D_{ss}^1}{\partial \tau}(\tau-\mu_\tau))$$

since, the expression for $D_1$ at the equilibrium for complete market:

$$D_1 = Y_1 - P_1F_1 - \tau_1\phi_1F_1 = (1-\alpha)Y_1$$

Therefore, with the linearization expressions, where $a \equiv log(A_1)$ and $\tau \equiv log(\tau_1)$, we have the results:

$$\begin{aligned} \kappa_{\tau,d} &= \frac{1}{D_{ss}^1}(\frac{\partial D_1}{\partial \tau})_{ss} = \frac{1}{D_{ss}^1}[\underbrace{e^{\mu_a}(F_1^{ss})^{\alpha-1}\alpha(\frac{\partial F_1}{\partial \tau})_{ss}]}_{(-)} \\ \kappa_{a,d} &= \frac{1}{D_{ss}^1}(\frac{\partial D_1}{\partial a})_{ss} = \frac{1}{D_{ss}^1}[\underbrace{e^{\mu_a}(F_1^{ss})^{\alpha} + e^{\mu_a}\alpha(F_1^{ss})^{\alpha-1}(\frac{\partial F_1}{\partial a})_{ss}]}_{(+)} \\ \kappa_{a,m} &= -\frac{1}{C_{ss}^1}(\frac{\partial \tilde{C}_1}{\partial a})_{ss} = -\frac{e^{\mu_a}(F_1^{ss})^{\alpha} + (e^{\mu_\tau} - \Phi)\phi_1(\frac{\partial F_1}{\partial a})_{ss}}{\tilde{C}_{ss}^1} \\ \kappa_{\tau,m} &= -\frac{1}{C_{ss}^1}(\frac{\partial \tilde{C}_1}{\partial \tau})_{ss} = -\frac{(e^{\mu_\tau} - \Phi)\phi_1(\frac{\partial F_1}{\partial \tau})_{ss}}{\tilde{C}_{ss}^1} \end{aligned}$$

where the $(\frac{\partial F_1}{\partial \tau})_{ss} = \frac{-\phi_1}{(1-\alpha)\alpha e^{\mu_a}F_1^{\alpha-2} + \pi_1/\bar{R}} < 0$ and $(\frac{\partial F_1}{\partial a})_{ss} = \frac{e^{\mu_a}(F_1^{ss})^{\alpha-1}\alpha}{\pi_1/\bar{R} + \alpha(1-\alpha)e^{\mu_a}(F_1^{ss})^{\alpha-2}} > 0$.

### 1.10.4 Decentralized Economics

For simplicity, the notation of the paper with subscript 'ss' refers to the economic fundamentals at the steady states.
In this section, I show some of the derivation for decentralized economics in this paper. The equilibrium is derived under an asymmetric setting for $\tau_{G,t} = 0$ and $\tau_{B,t} = \tau_t$ but homogeneity for firms within each sectors (B,G). The decentralized economics is defined as the following:

**Definition 1** *The market recursive equilibrium is defined as a joint optimal policy for households, corporations and government. Given the exogenous shocks, $\{A_t, w_t\}$, in economic growth and perception on government carbon taxation, the following conditions should hold for market equilibrium. 1). The optimal policies for households are determined among $\{C_t, w_{B,t+1}, w_{G,t+1}, B_{t+1}\}$ via maximizing the representative utility functions taking the asset prices as given. 2). Firms maximize their the market valuation via optimal policies on $\{I_t^B, I_t^G, K_{t+1}^B, K_{t+1}^G, L_{t+1}^B, L_{t+1}^G, Q_t^B, Q_t^G, Q_t^R, Q_t^X$ The prices related variables, including $\{P_t, P_t^B, P_t^G, V_{B,t}, V_{G,t}, D_t^B, D_t^G, P_t^F, R_t^f\}$ are obtained via market clearing conditions for equities, government bond, fossil fuel capitals and goods market. 4). Government finances through debt issuance and possible carbon taxation. It transfers the fiscal surplus to the households in the form of lump sum taxation, $TR_t$.*

### 1.10.5 Model Equation

Since the model includes a non-stationary growth for long-run economic growth, the state and controlled variables at time t are normalized with the aggregate growth level, $Z_{t-1}$ to revise the state variables over the balanced growth path. There are 17 state variables in the economy, which are denoted as

$$\mathcal{S}_t = \{\omega_t^B, \omega_t^G, B_t, \widehat{K}_t^B, \widehat{K}_t^G, Q_t^B, Q_t^G, \widehat{R_t}, \widehat{X_t}, Q_t^R, Q_t^X, a_t, \widehat{G_t^a}, \Gamma_t, \widehat{M_t}, w_t, Z_t, \widetilde{S}_t\}$$

where a subscript hat refers to the scaling. The list below shows the optimal conditions after normalization for the recursive equilibrium ($\widehat{K}_t^B = \frac{K_t^B}{Z_{t-1}}$).

#### Decentralized Economic Optimality

Household Optimization:
For $i \in \Big\{B, G\Big\}$:

$$E_t\Big(M_{t,t+1} R_{t,t+1}^i\Big) = 1$$

$$R_{t,t+1}^i = \frac{\widehat{V_{t+1}^i}}{\widehat{V_t^i} - \widehat{D}_t^i} e^{\Delta z_{t-1}}$$

$$M_{t,t+1} = \beta\big(\frac{\widehat{\widetilde{C}}_{t+1}e^{\Delta z_t}}{\widehat{\widetilde{C}}_t}\big)^{-\frac{1}{\psi}}\left(\frac{\widehat{U}_{t+1}}{E_t[\widehat{U}_{t+1}^{1-\gamma}]^{\frac{1}{1-\gamma}}}\right)^{\frac{1}{\psi}-\gamma}$$

$$\widehat{\widetilde{C}}_t \equiv \frac{\widetilde{C}_t}{Z_{t-1}}$$

$$R_{t+1}^f = \frac{1}{E_t\big[M_{t,t+1}\big]}$$

Firm Dynamics:

The firms in the H sector have the equilibrium conditions:

w.r.t $\widehat{I}_t^B$:

$$Q_t^B(\Phi_t^B)' = 1$$

w.r.t $L_t^B$:

$$\frac{P_t^B(1-\alpha)\widehat{Y}_t^B}{G_t^B} = \widehat{W}_t$$

w.r.t $Q_t^B$:

$$\widehat{K}_t^B e^{\Delta z_{t-1}} = \widehat{K}_{t-1}^B(1-\delta) + \Phi^B\widehat{K}_{t-1}^B$$

w.r.t $\widehat{K}_{t+1}^B$:

$$Q_t^B = E_t\big[M_{t,t+1}\big(\frac{P_{t+1}^B\alpha(1-\nu)\widehat{Y}_{t+1}^B}{\widehat{\mathcal{E}}_{t+1}^{1-\frac{1}{\sigma}}(\widehat{K}_{t+1}^B)^{\frac{1}{\sigma}}} + Q_{t+1}^B[(1-\delta)+\Phi_{t+1}^B-(\Phi_{t+1}^B)'\frac{\widehat{I}_{t+1}^B}{\widehat{K}_{t+1}^B}]\big)\big]$$

w.r.t $\widehat{F}_t$

$$\frac{P_t^B\alpha\nu\widehat{Y}_t^B}{\widehat{\mathcal{E}}_t^{1-\frac{1}{\sigma}}(\widehat{F}_t)^{\frac{1}{\sigma}}} = P_t^F + \tau_t\phi_1$$

$$P_t^{d,*} \equiv P_t^d + \tau_t\varphi_1$$

$$\widehat{\mathcal{E}}_t = \big[(1-\nu)(\widehat{K}_t^B)^{\frac{\sigma-1}{\sigma}} + \nu(\widehat{F}_t)^{\frac{\sigma-1}{\sigma}}\big]^{\frac{\sigma}{\sigma-1}}$$

$$\widehat{Y}_t^B = \big(\widehat{A}_t L_t^B\big)^{1-\alpha}(\widehat{K}_t^B)^{\alpha}$$

$$\widehat{A}_t = e^{\Delta z_t}\Omega(\widetilde{S}_t)$$

$\Phi_t$ is defined in the main content.

For the firms in the Green sector: w.r.t $\widehat{L}_t^G$

$$\frac{P_t^G(1-\alpha)\widehat{Y}_t^G}{L_t^G} = \widehat{W}_t$$

w.r.t $\widehat{K}^G_{t+1}$

$$Q^G_t = E_t\big[M_{t,t+1}\big(\frac{P^G_{t+1}\alpha\widehat{Y}^G_{t+1}}{\widehat{K}^G_{t+1}} + Q^G_{t+1}[(1-\delta) + \Phi^G_{t+1} - (\Phi^G_{t+1})'\frac{\widehat{I}^G_{t+1}}{\widehat{K}^G_{t+1}}]\big)\big]$$

w.r.t $\widehat{Q}^G_t$:

$$\widehat{K}^G_t e^{\Delta z_{t-1}} = \widehat{K}^G_{t-1}(1-\delta) + \Phi^G\widehat{K}^G_t$$

$$\widehat{Y}^G_t = \big(\widehat{A}_t L^G_t\big)^{1-\alpha}(\widehat{K}^G_t)^{\alpha}$$

Energy Sector:

w.r.t $\theta_t$

$$P^d_t = \phi(\theta_t)' + Q^R_t$$

w.r.t $\widehat{R}_{t+1}$

$$Q^R_t = E_t\big(M_{t,t+1}\big[P^F_{t+1}\theta_{t+1} - \phi(\theta_t) + Q^R_{t+1}(1-\theta_{t+1})\big]\big)$$

w.r.t $\widehat{X}_{t+1}$

$$Q^x_t = E_t\big[M_{t,t+1}\big(\lambda Q^R_{t+1} + (1-\lambda)Q^x_{t+1}\big)\big]$$

w.r.t $\widehat{I}^o_t$

$$Q^x_t = \frac{1}{\xi}\big(\frac{\widehat{I}^o_t}{\widehat{L}^a_t}\big)^{1-\xi}$$

where $\widehat{L}^a_t \equiv \frac{L^a_t}{Z_{t-1}}$.

w.r.t $Q^R_t$:

$$\widehat{R}_t e^{\Delta z_{t-1}} = (1-\theta_t)\widehat{R}_{t-1} + \lambda\widehat{X}_{t-1}$$

w.r.t $Q^x_t$:

$$\widehat{X}_t e^{\Delta z_{t-1}} = (1-\lambda)\widehat{X}_{t-1} + \lambda\big(\widehat{I}^o_t\big)^{\xi}\big(\widehat{L}^a_t\big)^{1-\xi}$$

The exogenous land resource after normalization follows:

$$log(\widehat{L}^a_t) = (1-\rho_L)L^a_{ss} + \rho_L log(\widehat{L}^a_{t-1}) + \rho_a a_t + \epsilon_{l,t}$$

Final Good Sector:

$$\frac{P^B_t}{P^G_t} = \Big(\frac{\widehat{Y}^B_t}{\widehat{Y}^G_t}\Big)^{-\frac{1}{\sigma}}$$

Under normalization assumption:

$$1 = \big[\gamma_B(P^B_t)^{1-\sigma} + \gamma_G(P^G_t)^{1-\sigma}\big]^{\frac{1}{1-\sigma}}$$

Carbon Emission Cycle:

$$\widetilde{S}_{t+1} = \rho_S\widetilde{S}_t + (\gamma_B e_{B,t})^{\zeta}\big(\widehat{\Gamma}_t\big)^{1-\zeta}$$

Market Clearing:

$$\widehat{C}_t = \widehat{Y}_t - \widehat{I}_t - \widehat{I}_t^o - \phi_t \widehat{R}_t$$

$$\gamma_B \widehat{F}_t = \widehat{R}_t \theta_t$$

$$L_t = 1$$

$$\widehat{I}_t \equiv \gamma_B \widehat{I}_t^B + \gamma_G \widehat{I}_t^G + \widetilde{I}_t^O$$

**Climate Adjusted SDF**

The decomposition of the EZ-climate follows the paper of Bansal and Yaron (2004b) around the steady state. The aggregate environmental quality adjusted wealth return is defined as

$$R^W_{t,t+1} = \frac{a_t R_{t,t+1} + B_t R^f_{t,t+1} + TR_{t+1} + \Pi^o_{t+1} + W_{t+1}L_{t+1} - \Psi S_{t+1}}{a_t + B_t}$$

, where $a_t$ is defined as total asset from budget constraint: $a_t \equiv \int_i w_i P_{i,t} di$ and $R_{t,t+1} = \frac{\int_i w_{i,t+1}(P_{i,t+1}+D_{i,t+1})}{\int_i w_{i,t+1}P_{i,t}}$. Then, from household perspective, the budget constraint becomes $\widetilde{W}_{t+1} = (\widetilde{W}_t - \widetilde{C}_t)R^W_{t,t+1}$, where environment adjusted wealth is defined as $\widetilde{W}_{t+1} := a_t R_{t,t+1} + B_t R^f_{t,t+1} + TR_{t+1} + \Pi^o_{t+1} + W_{t+1}L_{t+1} - \Psi S_{t+1}$. Follow the Bansal and Yaron (2004b), guess a linear solution for linear value function is expressed as $U_t = Z_t \widetilde{W}_t$. From the optimality regarding the variable, $\widetilde{C}_t$ and denote $X_t = \frac{\widetilde{C}_t}{\widetilde{W}_t}$. Then, the household optimization becomes:

$$Z_t^{1-\frac{1}{\psi}} = (1-\beta)\big[\frac{\widetilde{C}_t}{\widetilde{W}_t}\big]^{-\frac{1}{\psi}}$$

$$\mu^*_t = \big[\frac{1-\beta}{\beta}\bigg(\frac{\frac{\widetilde{C}_t}{\widetilde{W}_t}}{1-\frac{\widetilde{C}_t}{\widetilde{W}_t}}\bigg)^{-\frac{1}{\psi}}\big]^{\frac{1}{1-\frac{1}{\psi}}}, \text{ where } \mu^*_t = E_t[U^{1-\gamma}_{t+1}]^{\frac{1}{1-\gamma}}$$

Then the stochastic discount factor is therefore transformed into:

$$M_{t,t+1} = \beta^{\frac{1-\gamma}{1-\frac{1}{\psi}}}\left(\frac{\widetilde{C}_{t+1}}{\widetilde{C}_t}\right)^{-\frac{1}{\psi}\frac{1-\gamma}{1-\frac{1}{\psi}}}(R^W_{t,t+1})^{\frac{\frac{1}{\psi}-\gamma}{1-\gamma}}$$

**Market Clearing Condition**

According to equity market clearing, the model has:

$$\omega_{B,t} = \gamma_B \text{ and } \omega_{G,t} = \gamma_G \text{ with } \gamma_B + \gamma_G = 1$$

From the representative agent budget constraint and equity market clearing, there is:

$$C_t + B_t = B_{t-1}R_{f,t-1} + W_t L_t + \gamma_B D_{B,t} + \gamma_G D_{G,t} + TR_t + \Pi^o_t$$

Follow the $TR_t$ and $\Pi^o_t$:

$$C_t = W_t L_t + \gamma_B D_{B,t} + \gamma_G D_{G,t} - \gamma_B * Tax_{B,t} + P_t J_t - I^F_t - \Psi^R_t$$

Input the definitions of $D_{B,t}$ and $D_{G,t}$. Exploit the market clearing for $J_t = \gamma_B F_t$ and the labor market $L_t = \gamma_B L_{B,t} + \gamma_G L_{G,t}$. Finally, given the zero-profit condition of the final producer,

$Y_t = \gamma_B P_t^B Y_t^B + \gamma_G P_t^G Y_t^G$, the model generates the good market clearing condition as 40.

**Deterministic Steady State**

Given the first order conditions in the previous sections, the model is solved with Dynare++ at steady states. The core to solve the model at steady states is to calculate the fixed points over the dynamic state variables where $\mathcal{S}_t = \bar{\mathcal{S}}$.

According to the 1.10.5, there are equations for energy sector and production sector respectively. The economy for all equations could ultimately be simplified into two equations for $F_t$ and $\theta_t$ respectively. Then, the steady states are easier to be obtained with numerical approximations. Figure 1.18 shows that the numerical approximates for the key variables, $F_{ss}$ and $\theta_{ss}$, are globally stable within sensible numerical support.

Given these settings, Figure A.1.19 shows that with higher carbon tax intensity at the steady states, the output share of the brown sector is reduced against the green sector, as in Panel A. When the steady-state carbon taxation increases from 20 dollars/ton of carbon to 40 dollars/ton of carbon, the corresponding ratio of $Y^B/Y^G$ drops from 0.95 to 0.94. Intuitively, it is due to the lower substitutability between green and brown capital. Firms are inflexible in switching their energy resources and have to accept the higher penalized brown cost. Simultaneously, the higher penalization of the brown capital contracts the share of the brown firms against the green firms ($\frac{K_{SS}^B}{K_{SS}^G}$) as in Panel B.

Apart from firm characteristics, the characteristic on household consumption displays a different pattern on consumption-environmental surplus, $\frac{C_{ss}-\Psi S_{ss}}{C_{ss}}$ in Panel C. Due to a higher sensitivity to a low-carbon economy, the surplus is higher with tightened climate policy at steady state ranging from 20% to 25%. In Panel D, the model shows the energy price at the steady state. Even though the tax-exclusive price remains stable regardless of steady-state carbon tax values, the tax-inclusive shows a strong positive correlation with the higher carbon taxation rate. Finally, in Panel E, Figure 1.19 shows a strong negative correlation between brown firm investment shares, $\frac{I_{ss}^B}{I_{ss}}$, and steady-state carbon taxation. The investment share of high CRE (brown) firms decreases from 7% to 4.5% as carbon taxation rises from 20 dollars to 40 dollars per ton of $CO_e^2$[15]. In contrast, the green (low CRE) firm investment share increases as the climate policy risk intensifies in Panel F.

[15] $CO_e^2$ refers to carbon dioxide equivalent greenhouse gas according to Trucost

**Centralized Economics**

In this section, the model is applied to an economics with endogenous carbon tax policy to productivity risk. The state variables in this economics include:

$$\mathcal{S}_t^* = \{B_t, \widehat{K}_t^B, \widehat{K}_t^G, \lambda_t^B, \lambda_t^G, \widehat{R_t}, \widehat{X_t}, \lambda_t^R, \lambda_t^X, a_t, \widehat{L_t^a}, \Gamma_t, \widehat{S}_t, \lambda_t^S, Z_t\}$$

The social planner is to maximize the representative household in the economics:

$$\begin{aligned}
\mathcal{U}_t = &\{(1-\beta)\tilde{C}_t^{\,1-\frac{1}{\psi}} + \beta(E_t[U_{t+1}^{1-\gamma}])^{\frac{1-\frac{1}{\psi}}{1-\gamma}}\}^{\frac{1}{1-\frac{1}{\psi}}} + \sum_{i\in\{B,G\}} \lambda_{it}^i(K_t^i(1-\delta_c) + \Phi^i(I_t^i/K_t^i)K_t^i - K_{t+1}^i) \\
&+ \lambda_t([\gamma_B(Y_t^B)^{\frac{\sigma-1}{\sigma}} + \gamma_G(Y_t^G)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} - C_t - \int I_t^i di - \Phi^F R_t) \\
&+ \lambda_t^R[(1-\theta)R_t + \lambda X_t - R_{t+1}] + \lambda_t^X[X_t(1-\lambda) + (I_t^O)^\xi (L_t^a)^{1-\xi} - X_{t+1}] \\
&+ \lambda_t^N(1 - \gamma_B L_t^B - \gamma_G L_t^G) + \lambda_t^F(\theta_t R_t - \gamma_B F_t) + \lambda_t^{Y^B}((\widetilde{A}_t L_t^B)^{1-\alpha}(E_t)^\alpha - Y_t^B) \\
&+ \lambda_t^{Y^G}[(\widetilde{A}_t L_t^G)^{1-\alpha}(K_t^G)^\alpha - Y_t^G] + \lambda_t^S\big(S_t - \rho_S S_{t-1} - (\int \phi_1 F_t di)^{1-\zeta}(\Gamma_t)^\zeta\big)
\end{aligned}$$

where $E_t \equiv \big[\nu(F_t)^{\frac{\sigma-1}{\sigma}} + (1-\nu)(K_t^B)^{\frac{\sigma-1}{\sigma}}\big]^{\frac{\sigma}{\sigma-1}}$ as the previous definition. The social planner is able to maximize the welfare via internalizing the carbon emission externalities. Based on this target function, the first order condition for $I_t^i$ shows the result:

$$\frac{\lambda_t^i}{\gamma_i \lambda_t}(\Phi^i)' = 1$$

where marginal Q for $i$ is defined as $Q_t^i \equiv \frac{\lambda_t^i}{\gamma_i \lambda_t}$. The first order conditions for $Y_t^i$ define the price of intermediate products as in $P^i \equiv \frac{\lambda_t^{Y^i}}{\gamma_i \lambda_t}$. Then, maximizing the energy capital extraction with $\theta_t$, the FOC implies that:

$$\lambda_t(\Phi_t^R)' = \lambda_t^F + \lambda_t^R$$

where the price of the energy capital in decentralized could be defined as $P_t^F \equiv \frac{\lambda_t^F}{\lambda_t}$, $Q_t^R \equiv \frac{\lambda_t^R}{\lambda_t}$. With these price definitions, the FOCs for green capital H, L give the similar analytic results as the decentralized economy for intermediate output prices:

$$\lambda_t^{Y^i} = \lambda_t \gamma_i \big(\frac{Y_t}{Y_t^i}\big)^{1/\sigma}, \text{ where } P_t^i \equiv \frac{\lambda^{Y^i}}{\gamma_i \lambda_t} \text{ and } P_t \equiv 1$$

The pre-defined pricing variables are internally consistent within other FOCs including the green capital at $t+1$:

$$1 = \beta E_t\Big[\big(\frac{U_{t+1}}{E_t\big(U_{t+1}^{1-\gamma}\big)^{\frac{1}{1-\gamma}}}\big)^{\frac{1}{\psi}-\gamma}\big(\frac{\widetilde{C}_{t+1}}{\widetilde{C}_t}\big)^{-\frac{1}{\psi}}\left\{\frac{\lambda^{Y^i}}{\gamma_i\lambda_t}\frac{Y_{t+1}^i\alpha(1-\nu)}{E_{t+1}^{1-\frac{1}{\sigma}}(K_{t+1}^i)^{\frac{1}{\sigma}}} + \frac{\lambda_t^i}{\gamma_i\lambda_t}[(1-\delta) + \Phi_{t+1}^i - (\Phi_{t+1}^i)'\frac{\widehat{I}_{t+1}^i}{\widehat{K}_{t+1}^i}]\right\}\Big]$$

To derive the Pigovian tax for the carbon emission, the FOC for carbon accumulation price shows the shadow price of accumulated carbon emissions at the atmosphere as in :

$$\lambda_t^S = \underbrace{\lambda_t \Psi}_{\text{Preference Cost}} + \underbrace{\lambda_t (1-\alpha)\mu \frac{Y_t}{Z_{t-1}}}_{\text{Production Cost}} + \underbrace{\rho_S E_t \{ M_{t,t+1} \big( \frac{U_{t+1}/\widetilde{C}_{t+1}}{U_t/\widetilde{C}_t} \big)^{-\frac{1}{\psi}} \lambda_{t+1}^S \}}_{\text{Utility Continuation}}$$

The FOC for $F_t$ shows the price of carbon emission in the economy as:

$$\underbrace{\frac{\lambda_t^{Y^B}}{\gamma_B \lambda_t}}_{P_t^B} \frac{\alpha \nu Y_t^B}{E_t^{1-\sigma} F^{\frac{1}{\sigma}}} = \underbrace{\frac{\lambda_t^F}{\lambda_t}}_{P_t^F} + \phi_1 \underbrace{(1-\zeta) \frac{\lambda_t^S}{\lambda_t} \big( \frac{\Gamma_t}{\phi_1 F_t \gamma_B} \big)^{\zeta}}_{\text{Optimal Carbon Tax Price}}$$

## 1.11 Appendix B.

This appendix provides the details about the text mining dictionary for the bigram search of climate regulatory exposure (CRE). Some examples of the related text information are shown as below from 10K filings. The dictionary list below shows the top 30 words lists for climate change risk and government regulation correspondingly.

Table 1.18: Matched Top 30 GPV and CCV

| | CCV Top 30 | CCV Count | GPV Top 30 | CPV Count |
|---|---|---|---|---|
| 1 | sustainab | 66,935 | politic | 73,249 |
| 2 | ch4 | 55,537 | regulatory | 69,041 |
| 3 | ozone | 48,168 | act | 65,547 |
| 4 | renewable | 41,598 | Legislative | 64,822 |
| 5 | gwp | 38,085 | pederal | 61,560 |
| 6 | ghg | 36,319 | legislation | 57,066 |
| 7 | ods | 34,127 | accord | 56,450 |
| 8 | hfcs | 33,889 | reimbursement | 51,236 |
| 9 | co2 | 31,857 | policy | 50,351 |
| 10 | nox | 31,491 | subsidy | 48,770 |
| 11 | pfcs | 31,379 | congressional | 48,288 |
| 12 | ets | 30,750 | agreement | 47,109 |
| 13 | lcfs | 29,830 | government | 46,692 |
| 14 | methane | 29,677 | government approved | 46,668 |
| 15 | oils | 29,327 | legislatory | 46,192 |
| 16 | n2o | 28,051 | law | 45,052 |
| 17 | euets | 26,431 | regulation | 45,035 |
| 18 | green | 26,280 | legal | 43,097 |
| 19 | sf6 | 22,971 | government sponsored | 42,878 |
| 20 | temperatures | 21,939 | congress | 41,679 |
| 21 | nf3 | 21,558 | tax | 41,542 |
| 22 | climate | 21,255 | political | 35,461 |
| 23 | extinctions | 19,332 | governmental | 31,266 |
| 24 | rps | 18,348 | kyoto protocol | 2,083 |
| 25 | reserves | 16,424 | debt ceiling | 1,979 |
| 26 | emissions | 15,406 | price constraint | 1,800 |
| 27 | emit | 14,077 | government approval | 1,796 |
| 28 | low carbon fuel standards | 13,816 | governmental regulation | 1,751 |
| 29 | corporate average fuel economy | 13,720 | government deficit | 1,719 |
| 30 | nitrogen trifluoride | 12,845 | regulatory delay | 1,597 |

Note: the paper uses two dictionaries: CCV (climate change vocabulary) from Chou and Kimbrough (2019a) to capture climate-related terms, and GPV (government policy vocabulary) from Koijen et al. (2016) to capture firm-level risks related to government regulatory policy. Further details can be found in Baz et al. (2022a).

# 2. Firm-Level Climate Regulatory Exposure

Salim Baz, Lara Cathcart, Alexander Michaelides and Yi Zhang

## 2.1 Introduction

Climate risk is increasingly receiving substantial attention globally; recent surveys (for instance, Krueger et al. (2020)) show that institutional investors consider climate concerns as one of the most influential risks affecting investment decisions.[1] Consistent with this message, recent papers emphasize the role of climate risk in cross-sectional asset pricing (Berkman et al. (2019), Bolton and Kacperczyk (2020a), Engle et al. (2020a), and Ilhan et al. (2020)), in international asset pricing (Bolton and Kacperczyk (2020b)), and housing markets (Giglio et al. (2020)). Moreover, various central banks are already emphasizing the importance of climate risks in the financial system and the economy more broadly (see Bolton et al. (2020) for an excellent description of the challenges climate risks pose to society).[2]

Even though climate risk is increasingly becoming more recognized as a financial risk by both individual investors and central bankers, there are still few research papers focussing on one particular component of climate risk, namely climate regulatory exposure (CRE). A recent example of climate regulatory risk motivating our analysis is the case brought agaisnt Exxon-Mobil by New York's Attorney General.[3] The lawsuit charged that ExxonMobil was hiding from shareholders risks to the business emanating from climate change regulation. ExxonMobil eventually won this USD 1.6bn lawsuit in December 2019 but the case indicates that climate regulatory risks cannot be neglected.

Our main contribution is to investigate empirically whether there exists a causal relationship between CRE and stock returns. We use the unexpected outcome of the 2016 U.S. Presidential

[1] A recent excellent literature review of ESG and CSR research in corporate finance can be found in Gillan et al. (2021).

[2] In a speech on 26$^{th}$ August 2020, SEC Commissioner Lee further emphasizes the importance of climate risk for asset pricing pointing out that "By some estimates, over 90% of U.S. equities by market capitalization are exposed to material financial impact from climate change"; https://www.sec.gov/news/public-statement/lee-regulation-s-k-2020-08-26.

[3] Discussed in https://www.nytimes.com/2018/10/24/climate/exxon-lawsuit-climate-change.html.

election as an exogenous shock to investor perceptions about the stringency of future climate regulatory legislation and enforcement. Meeuwis et al. (2018) find that perceptions (beliefs) can change after the outcome of a public event like the Trump election; we have a similar narrative in mind with regards to climate issues.[4] The idea is that if we have a good measure of CRE at the time of the election, given the pre-electoral commitments, investors would react in the subsequent few months to internalize those promises. To the extent that climate regulatory risks would be perceived to be less important after the election, firms with higher CRE at the time of the election would stand to gain in value relative to firms with lower CRE.[5] Wagner et al. (2018), Meeuwis et al. (2018) and Child et al. (2020) further discuss in detail the evidence supporting the hypothesis that the 2016 Trump election was a surprise to the markets, a key identification assumption in our analysis.

To empirically test such a hypothesis, we first need to construct a firm-specific measure of CRE. We do this based on text mining 10-K disclosure statements submitted to the Securities and Exchange Commission (SEC) by individual corporates. The SEC provides, since 2010, guidelines with regards to firms' disclosures on climate regulatory risks as follows: "when assessing potential disclosure obligations, a company should consider whether the impact of certain existing laws and regulations regarding climate change is material. In certain circumstances, a company should also evaluate the potential impact of pending legislation and regulation related to this topic."[6] We follow the text-mining literature to construct a firm-specific CRE measure based on these disclosures. We do this using an algorithm similar to Hassan et al. (2019a) implementing the classical bag-of-words method to quantify the unstructured text information. We use two different dictionaries, one from Koijen et al. (2016) on government policy regulatory uncertainty and the second from Chou and Kimbrough (2019a) on climate risks. By combining these two dictionaries, we construct a firm-level CRE measure.

Our measure is based on the climate regulatory risks individual firms describe in reports required by the SEC. Given that these risks are self-reported, one could argue that firms might downplay such risks, but as the 2018 ExxonMobil case indicates, downplaying such risks can result in lawsuits to defraud shareholders. Firms therefore have a relatively strong incentive to report the relevant risks accurately. Nevertheless, we also validate the constructed measure in two different ways. First, we show that the aggregate CRE measure has risen since 2007, when the relevant requirements on firms to discuss climate change risks were introduced by the SEC. Second, the measure ranks industries in expected ways, with oil, coal and gas industries being by far at the top of the CRE ranking over the whole period (2006-2018).[7]

[4]Baldauf et al. (2020) provide evidence for house price effects emanating from heterogeneity in beliefs about long-run climate risks. In our case, the heterogeneity in beliefs is about the future evolution of climate regulatory risks in the U.S. after the 2016 Trump election.

[5]Gerarden et al. (2020) discuss how the Trump administration made certain climate-related laws less stringent.

[6]See https://www.sec.gov/news/press/2010/2010-15.htm.

[7]Large firms also issue sustainability reports, where climate risks are perhaps discussed more explicitly; we

Given this CRE measure is at the individual firm level, we can explore its relationship with individual firm stock returns and investigate whether there exists a robust negative relationship between these two variables. To identify the causality of this relationship, we use the Donald Trump election as an exogenous shock to CRE. A number of recent papers have used the Trump election as an exogenous shock. For instance, Child et al. (2020) use this identification strategy to investigate the value of political connections, while Ilhan et al. (2020) use the election to investigate carbon risk. Our identification strategy is similar as a way to provide an exogenous variation in investor perceptions on climate regulatory exposure risks and their effect on valuation.

Using the difference-in-differences (diff-in-diff) approach across different groups based on their CRE at the time of the election, we find a statistically significant and economically important effect from this perceived decrease in the risks faced by high CRE firms. The cumulative stock returns over two months are between 240 and 410 basis points higher, depending on how the treatment and control groups are defined. Moreover, the empirical results are robust. First, when conducting the same analysis using cumulative abnormal returns (CARs) using either the Capital Asset Pricing Model (CAPM) or the Fama-French three-factor model to define abnormal returns, the results do not change, either statistically, or in terms of economic magnitude. Second, the results do not change when we switch to different training samples for these factor models. Third, using a placebo test with the Obama presidential election in 2012 replacing the Trump 2016 presidential election, confirms that our findings are not driven by other competing political confounding factors.

Importantly, we also construct a climate change exposure (CCE) measure that contains instances where climate is mentioned without an explicit reference to regulation. We use this measure to guard against the possibility that the CRE measure is capturing alternative climate risks. We show that the constructed CCE is positively correlated in a statistically significant way with the CRE measure. This positive correlation indicates that our empirical results might be driven by investors' perceptions about changes in CCE rather than in CRE. Nevertheless, our results remain robust and statistically and economically significant after we control for CCE, indicating that CCE changes cannot rule out our findings.

One key question is the exact event date that should be used in determining the change in investor perceptions. Recent evidence in Ramelli et al. (2018) questions whether the actual event is the election itself (November 8th), and emphasizes the importance of two other major events in affecting investor opinions about the plans of President Trump's administration. President Trump gave a widely followed interview in the *New York Times* on November 23rd, and subsequently surprised the markets by announcing the appointment of EPA Administrator Scott Pruitt (a climate change sceptic) on December 7th. Ramelli et al. (2018) focus on the

leave that analysis to future research.

election and the announcement of the appointment to create their event window; we argue that the interview can also be used in such an analysis, in the sense that investors continually form expectations about future regulatory changes during this period. To avoid taking a precise stand on when the actual change in investor perceptions with regards to climate regulatory issues solidifies, we use the end of month returns in our baseline regression specifications to capture the potentially slow change in these perceptions.[8] The slow reaction of stock returns to news related to climate risks is also found in Hong et al. (2019), Ardia et al. (2020) and Pastor et al. (2021b).

Another possible identification strategy would be to exploit the December 2015 Paris agreement on climate change. Bolton and Kacperczyk (2020b) use this strategy in an international data set and find evidence that domestic regulations can be more important than international agreements in affecting carbon transition risk. Given that we use a domestic U.S. data set and the Paris agreement was never enforced in the U.S. because of the 2016 Trump election, we do not expect to have results that are as statistically and economically significant as when using the Trump election. Indeed, we find that using the Paris agreement as an exogenous shock generates lower returns, but only for the top decile CRE portfolio.

There are other potential hypotheses that could explain our findings; for instance, Wagner et al. (2018) argue that perceptions about international trade and corporate taxation policies also changed after the Trump election. Following Wagner et al. (2018), we construct the international trade exposure using pre-tax income from foreign operations as a percent of total assets, and the corporate taxation exposure using the corporate taxation expense as a share of pre-tax income. We show that the empirical conclusions on the role of CRE on individual stock returns remain unchanged after controlling for these variables.

Another explanation for our results involves the change in commodity (oil) prices after the election (see Gao et al. (2021)), given that such changes might be correlated with the firms selected for treatment according to our CRE measure. To control for this possibility we construct the oil-risk exposure for each firm, extending a CAPM model for individual stock returns to include the effect of oil price changes, in addition to the effect of market returns (see Ilhan et al. (2020) for a similar approach). We do that for each firm and use the individual firm exposures to oil price changes as an additional control to our baseline CRE diff-in-diff regressions. Our empirical conclusions are again unaffected, either in terms of statistical significance or economic magnitude.

Our baseline narrative is that investor perceptions change after the 2016 election based on the electoral promises of the incoming President and therefore firms with high CRE at that point in time are suddenly perceived to be less risky and experience a rise in their valuation over

[8] We also provide robustness tests using daily returns based on the interview date in Online Appendix Table A5, and they are consistent with our monthly return results.

the subsequent two to three months. We now offer evidence independent from stock market valuation to support this narrative. First, we show that carbon emissions from high-CRE firms rise more relative to low-CRE firms after the election, when compared to the period before the election.[9] This evidence is consistent with Fowlie et al. (2014), Fried (2018c) and Fried et al. (2020), who show that firms tend to take into account higher climate regulatory risks, and in turn reduce their carbon emissions. In our findings, perceived climate regulatory risks are reduced, and are associated with higher carbon emissions.

As a second piece of supporting evidence, we use different measures of investor attention to illustrate that investor interest for high-CRE firms falls after the Trump election. The idea shares similarities with recent evidence that experiencing warmer than usual temperatures can affect retail investor attention to climate change (Choi et al. (2020)). Lower investor attention to high-CRE firms after the Trump election is consistent with the interpretation that the negative exogenous shock to perceived climate regulatory exposures made investors care less about such risks after the election.

A number of recent papers are related to our work. Kölbel et al. (2020) measure climate risk disclosures using a natural language processing (NLP) model and show the connection with financial risks in the CDS market. We differ by focussing on a measure linked to climate regulatory risks, rather than climate change risks. Moreover, we emphasize the causal effect of CRE on stock returns using a diff-in-diff setting before and after the U.S. Presidential election. Berkman et al. (2019) and Sautner et al. (2020) focus on climate risks, while Ilhan et al. (2020) use the Trump election for identification purposes but focus on carbon risks rather than CRE. Bolton and Kacperczyk (2020a) explore the persistent relationship between carbon emissions and stock returns. Our results support empirically the interpretation that regulatory climate risk exposure can rationalise the positive relationship between carbon emissions and stock returns documented by Bolton and Kacperczyk (2020a).

Other recent papers tend either not to find any effect of climate shocks on firm sales (Addoum et al. (2020)), or find evidence that climate risks are not efficiently discounted (Hong et al. (2019)). In both of these cases, climate risks are not sufficiently appreciated by firms or investors, respectively. Our results show that the Trump election was an important event that changed investor perceptions about climate regulatory risks sufficiently to have an effect on stock returns. Ramadorai and Zeni (2019) do provide evidence that climate regulatory risks affect firms' beliefs about future investment activities, while Seltzer et al. (2020) make a link between corporate bond ratings, environmental scores and corporate bond yields using the Paris agreement as an exogenous shock.

The paper is structured in the following way. Section 2 builds the hypotheses and discusses the identification strategy. Section 3 describes the main algorithm to measure CRE and CCE.

[9] We compare one and two year windows before and after the election.

Section 4 describes the data, and Section 5 provides evidence to show that the proposed CRE has intuitive properties across industries and over time. Section 6 contains the empirical results and Section 7 offers supporting evidence through carbon emissions and investor attention. Section 8 concludes.

## 2.2 Empirical Strategy

### 2.2.1 Hypothesis Development

Assuming that we can accurately measure firm-specific climate regulatory exposures, what happens when investors perceive a higher probability of expected regulatory costs? To the extent that such risks are undiversifiable, we expect that the greater the exposure to climate regulatory risk, the lower the value of the specific firm.[10] Giglio et al. (2020) provide evidence for a negative link between physical climate exposure and real estate values; our hypothesis is similar but relates climate regulatory exposure and firm stock returns. Similarly, Hsu et al. (2020) provide evidence that a pollution premium exists in the market: investors require a premium from those firms with high exposure to environmental pollution risks, while Van der Ploeg and Rezai (2020) look at how government regulations can affect stock market valuations negatively through the impact of stranded assets.[11] Furthermore, the enforceability of rules and regulations can be as important to investors as their existence. For example, Bhattacharya and Daouk (2002) show that it is the enforceability of insider trading laws through prosecutions, rather than their mere existence, which is priced in the market. Overall, the theories confirm a negative premium between high and low CRE exposed stock returns. More important, as a source of undiversified risk in the market, there are negative abnormal premiums.

Our first hypothesis is therefore:

*Hypothesis 1: An exogenous perceived increase in firm-level CRE should imply higher expected climate regulation costs and, therefore, decrease firm value (realized abnormal return).*

A second hypothesis involves the information reaction through which this response can arise. An event that reduces the regulatory risks (either in terms of more lenient legislation or looser enforcement) may reduce investor attention to this risk exposure. If an event reduces the importance of CRE for whatever reason, then it is only natural to expect investors to start paying less attention to CRE news or disclosures. Therefore, to the extent that an event

[10] There might be some exceptions like utility firms under rate of return regulations; these firms can pass the costs of greater climate regulation to their customers. To the extent that these examples are important, they would work towards not finding any relationship between climate risk exposure and valuation.

[11] Barnett (2020) presents a structural model of climate policy affecting stock market valuation through stranded assets in the oil market.

reduces CRE risks, we would expect investor attention to those risks to be reduced. In such a situation, instead of investigating the effect of a potentially exogenous change in CRE on stock returns, we can instead link CRE to investor attention.

We therefore reach a second hypothesis:

*Hypothesis 2: If investors perceive lower future CRE risks, investor attention to such risks should decrease.*

To the extent that investors and regulators pay less attention to CRE risks, firms will also have a weaker incentive to abide by existing regulatory standards. Simultaneously, following the literature on climate policy uncertainty and carbon emissions, Fried (2018c) documents using a structural model that higher climate regulatory uncertainty incentivizes firms to exploit clean innovation and reduce carbon emissions. Fried et al. (2020) use firms' internal price data for carbon and show that higher climate policy uncertainty decreases carbon emissions. Similarly, Fowlie et al. (2014) also argue that with more stringent climate regulation, firm carbon emissions should fall in the next 20 years, while Gerarden et al. (2020) also show how environmental policies can affect carbon emissions. Overall, we interpret this literature to be consistent with the view that firms with higher CRE, and facing more stringent regulation, will tend to reduce their carbon emissions.

We therefore reach a third hypothesis:

*Hypothesis 3: If firms are less concerned about CRE risks, firms' climate-related behavior should change (for example, carbon emissions could rise).*

### 2.2.2 Identification

To identify an exogenous variation in perceptions of CRE intensity, we use the Donald Trump presidential election as a particular event providing such a variation. We focus on this event for three reasons. First, compared with other climate events like the Paris agreement, the Trump election is a domestic U.S. event and therefore, it can be more important for the U.S. stock market. Second, climate change was an important topic during the 2016 Presidential election. Comparing Hillary Clinton's with Donald Trump's political positions during the election, we can infer a more deregulatory approach with respect to climate risk concerns from the Trump campaign. Finally, our approach is also consistent with Ilhan et al. (2020), who consider the Trump election as a policy surprise shock to carbon emissions risks in the options market, and with Wagner et al. (2018), Meeuwis et al. (2018) and Child et al. (2020) who also use the election as an unexpected shock. We interpret this election as less likely to be endogenous and correlated with pre-trend confounding factors like business cycle indicators.

### 2.2.3 Roadmap

To test the hypotheses causally linking CRE with stock returns, we first construct a measure of CRE across individual firms. We then validate the measure and use it as a way to sort firms before the surprise Trump election to investigate the effect of any post-election change in investor perceptions on stock returns.

## 2.3 A Measure of Climate Regulatory Exposure

We use a dictionary approach (e.g. Baker et al. (2016a), Koijen et al. (2016) and Hassan et al. (2019a)) to map documents (the 10-K SEC filings) into a document-term matrix.[12] We propose a measure of CRE based on n-gram searching whose purpose is to capture the firm's evaluation of risks arising from climate-related regulations, and not simply of risks related to climate change per se. Specifically, we use two word lists or dictionaries: (a) the climate change vocabulary (CCV) proposed by Chou and Kimbrough (2019a) to capture climate-related terms, and (b) the government policy vocabulary (GPV) used by Koijen et al. (2016) to capture firm-level risks related to government policy. As far as we know, the CCV created by Chou and Kimbrough (2019a) is the most comprehensive dictionary used in the literature to explore climate-related topics in SEC filings.[13]

For every SEC filing for firm $i$ in year $t$, we start by removing all numbers and punctuation (excluding '-', since the CCV contains n-grams with this character), as is standard in the literature. However, we keep all the stop words, because the CCV includes some stop words (for example, the word "of"). Since we expand the dictionaries for plurals, tenses and inflections, we do not need to stem the words in the document. Furthermore, when searching for terms in either dictionary, we do not consider capitalizations.

In order to construct a measure capturing CRE, we use one dictionary to match a regulatory, or government policy word, and another dictionary to match the references to climate risks.[14]

[12] We use Item 1 (Business) and Item 1A (Risk Factors) of each 10-K filing only, ignoring the rest of the document.

[13] While it is impossible for any given dictionary to fully capture all terms and expressions related to climate change, Chou and Kimbrough (2019a) develop an expanded dictionary to guarantee a search for climate-related terms in SEC filings that is as exhaustive as possible. For example, the authors include not only single terms (like "climate"), but also bigrams (such as "climate risk"), trigrams ("climate change risk") and even 4-grams ("climate-related financial risks").

[14] Similar to Hassan et al. (2019a) and Malmendier et al. (2020)). Our proposed measure shares similarities with Koijen et al. (2016) who use two dictionaries, one for regulatory policy and one for health risks to create an index for firms' concerns about health regulations.

The two dictionaries are combined as follows:

$$CRE_{it} = \frac{1}{B_{it}} \left( \sum_{b=1}^{B_{it}} \mathbb{I}_{\{b \in CCV\}} \times S_b(r) \right). \tag{2.1}$$

In Equation (2.1), $B_{it}$ represents the number of words in firm's $i$ 10-K document in month $t$; $\mathbb{I}_{\{b \in CCV\}}$ is an indicator variable equal to 1 if the b$^{\text{th}}$ term in the document is in the CCV, and 0 otherwise. The innovative contribution of the algorithm comes from the $S_b(\cdot)$ function; it is the sum of the number of GPV terms that are within a distance of $r$ words from word $b$. Formally, $S_b(r)$ is defined as:

$$S_b(r) = \sum_{c=b-r}^{b+r} \mathbb{I}_{\{c \in GPV\}}, \tag{2.2}$$

where $\mathbb{I}_{\{c \in GPV\}}$ is an indicator variable equal to 1, if the c$^{\text{th}}$ term in the document is in GPV, and zero otherwise. Summing all the terms in each document in equation (1) generates a climate regulatory measure based on term frequency; the idea is that a higher frequency should be given higher weight.

Intuitively, the algorithm looks for government regulatory policy terms that fall within a distance of $r$ words from climate risk terms, the idea being that if climate and government policy are discussed in close proximity, then this intersection can provide a good proxy for climate regulatory exposure (CRE) for the particular firm.[15] For example, given a specific word distance $r$, the algorithm will search around the key phrase representing climate risk and record the count of government policy phrases that are within this word distance. The underlying assumption is that if climate risk is being discussed, and GPV terms are also mentioned near this climate risk discussion, it is likely that the firm is discussing regulations relating to climate change.

Let's consider an example as an illustration. This extract is taken from the 10-K filing of Fieldpoint Petroleum Corp in March 2015:

1cm *[...] Climate change* ***legislation*** *or* ***regulations regulating*** *emissions of GHGs and VOCs could result in increased operating costs and reduced demand for the oil and gas we produce. Both houses of* ***Congress*** *have actively considered* ***legislation*** *to reduce emissions of GHGs [...] These findings by the EPA allow the agency to proceed with the adoption and implementation of* ***regulations*** *that would restrict emissions of GHGs under existing provisions of the* ***federal*** *Clean Air Act.*

This is a specific example of our search-and-match algorithm. We search the entire paragraph for terms in both the CCV (underlined) and GPV (in bold) dictionaries. We find 9 CCV n-

[15] In what follows, we use a value of $r$ equal to 10 following Malmendier et al. (2020).

grams and 7 GPV n-grams. We then match each of the nine CCV terms with any GPV term within a ten-word distance. In this example, we find 19 pairs.[16] Since this example contains 72 words in total, the CRE measure is $\frac{19}{72}$ (19 matched pairs divided by 72 words). Further details of the matching algorithm and n-grams used are in Online Appendices A and C.

One concern with our approach is that we are not only capturing climate regulatory risks but instead we are also measuring alternative climate risk exposure aggregately. We address this concern by creating a measure that focusses solely on general climate risks that are distinct from firms' concerns with regards to climate regulatory risks. We do so by measuring the climate risk variation conditional on no regulation words mentioned near the climate change ones (ten words before and ten words after the climate words).[17] To summarize, the main difference between the two climate exposure measures is that for the CRE measure we rely on regulation vocabularies to identify exposure to regulation risk, while the climate change exposure (CCE) calculation only includes exposure to physical climate risks, when no regulation words are nearby.

The CCE measure is therefore constructed as follows (where 'c' refers to the presence of any words near the climate change ones):

$$CCE_{it} = \frac{1}{B_{it}} \left( \sum_{b=1}^{B_{it}} \mathbb{I}_{\{b \in CCV\}} \times \mathbb{I}_{\{c \notin GPV\}} \right). \tag{2.3}$$

## 2.4 Data

Our data primarily consist of SEC filings from the SEC's Electronic Data Gathering, Analysis, and Retrieval system (EDGAR), as well as firm-specific data from COMPUSTAT and the Center for Research in Security Prices (CRSP). We provide further details below.

### 2.4.1 SEC Filings

We obtain the SEC filings from the EDGAR database using the web scraping open source code provided by Bill McDonald.[18] EDGAR is the online filing system created by the SEC and is the official database that stores the compulsory filings of U.S. firms. According to SEC

[16] They are ('climate change', 'legislation'), ('climate change', 'regulations'), ('climate change', 'regulating'), ('emissions', 'regulating'), ('GHGs', 'regulating'), ('VOCs', 'regulating'), ('emission', 'regulation'), ('emission', 'legislation'), ('GHGs', 'regulation'), ('GHGs', 'legislation'), ('VOCs', 'regulation'), ('VOCs', 'legislation'), ('oil', 'Congress'), ,('emission', 'legislation'), ('GHGs', 'legislation'), ('emission', 'Congress'), ('GHGs', 'Congress'), ('emission', 'federal') and ('GHGs', 'federal').

[17] We use ten words to be consistent with the CRE calculations but changing that number (say to twenty) does not affect our results.

[18] Available at `https://sraf.nd.edu/`.

disclosure rules, all U.S. public firms have to release internal information to the public on a regular basis. These filings are the firm's quarterly report (10-Q form), annual report (10-K form), and immediate information (8-K form).

We disregard the 10-Q and 8-K forms, choosing to focus solely on the content of 10-K filings. We do this because, in contrast to other filings, a firm's 10-K form provides the most comprehensive description about the firm's annual performance, including risks from climate change and climate-related regulations, as well as other expected risk factors. In contrast, the 10-Q form typically presents quarterly financial statements, with no accompanying information pertaining to business risks. This form is therefore not suited to our study as it does not mention the firm's evaluations of climate-related risks. Moreover, 10-K filings are regularly scheduled, compulsory reports. They are therefore more likely to provide firm assessments about forward-looking economic conditions; 8-K forms, on the other hand, are unscheduled reports through which the company communicates important matters to shareholders that mainly pertain to irregular and/or unexpected events.

We focus on Item 1 (Business Economics), which discusses firm outlooks on current business conditions, and Item 1A (Risk Factors), which discusses the main risk factors relevant to the firm. While the structure of 10-K filings has changed through time, Item 1 and Item 1A have consistently reflected the firm's evaluations of the economic outlook and risk factors that may affect the firm. According to SEC rules, public firms are required to disclose any risks to all potential investors, and climate regulatory exposure is one such risk discussed in these reports.

Our sample period ranges from January 2006 to December 2018. There are 104,032 10-K filings in total, as reported in Table 2.1. On average, every year in our sample includes 8,021 10-K filings for 7,930 individual firms. In addition, the average length of the document increases year by year, ranging from 4,819 to 8,164 words. We use the industry classification proposed by Fama and French (1992), and for the baseline results we exclude financial services firms. The examples of 10-K filings on climate regulatory exposure are provided in Appendix B.

### 2.4.2 Text Mining Data

We use the climate vocabularies (CCV) from Chou and Kimbrough (2019a) to capture climate related terms in 10-K filings. We should note here that the original CCV of Chou and Kimbrough (2019a) captures every aspect of the potential topics on climate change, including the environment and health-damaging chemicals. Accordingly, some of the words included in this dictionary do not capture climate risks per se, the focus of our analysis. We therefore adjust the CCV by excluding any n-grams that do not directly pertain to climate change (for example, words relating to health hazards, or public health). Our dictionary – the trimmed-down CCV – covers topics on carbon issue, green technology, global warming, chemical emissions and

temperature change. It should therefore be interpreted as a climate dictionary list on global warming.

Additionally, we use the government policy-related word list from Koijen et al. (2016) to capture regulatory risk in 10-K filings. Koijen et al. (2016) provide a comprehensive list of terms relating to government policy concerns (e.g. "taxation" and "regulation") and use it to explore the relation between government policy and a financial health premium. Table A1 in Appendix A provides the list of the top 30 n-grams in the CCV (out of a total of 106 n-grams), as well as the top 30 n-grams in the GPV (out of a total of 44 n-grams).

### 2.4.3 Stock Returns, Firm Fundamentals and Carbon Emissions

We obtain monthly stock return ($Ret_{it}$) data for individual firms from CRSP and quarterly balance sheet data from COMPUSTAT to construct commonly used control variables. The investment rate ($Invest_{it}$) is the firm capital expenditure divided by the total assets at time $t$. $Leverage_{it}$ denotes the book leverage of the firm, $log(Size)_{it}$ is the log of the firm's market capitalization, $log(PPEGT)_{it}$ is the natural logarithm of property, plant and equipment expenditures, and $B/M_{it}$ represents the book-to-market ratio, with market value calculated at the end of last calendar year. $Sale_{it}$ represents the value of sales from COMPUSTAT (obtained at the quarterly frequency). $Operation(Foreign)_{it}$ refers to the pre-tax income from foreign operations as a percentage of the firm's total assets, $ETR(GAAP)_{it}$ is the tax expense at the firm level, computed as the ratio of the tax expense on pre-tax income.

Besides the firm accounting information, we use the carbon emission data from Trucost as in Bolton and Kacperczyk (2020a). The carbon emissions measure is the sum of direct carbon emissions from production (scope 1) and indirect carbon emissions from electricity and heat consumption (scope 2). The emissions data are merged with COMPUSTAT through the common identifier gvkey.

To control for the effect of commodity prices like oil, we follow Ilhan et al. (2020). Specifically, $\beta_{it}^{Oil}$ measures the sensitivity of the firm's stock return to the changes in oil prices. We use the WTI future price and we follow the Fama-Macbeth method to compute $\beta_{it}^{Oil}$ in a first stage regression over a 2-year window, as a proxy for firm exposure to oil risk.[19] The benchmark model we follow is $Ret_{it} = \beta_0 + \beta_1 R_{Market,t} + \beta_{it}^{Oil} \Delta P_{oil,t}$, where $\Delta P_{oil,t}$ is the growth rate in oil prices.

We report the summary statistics for various variables used in the analysis in Table 2.2, Panel A. The mean firm market capitalization in the sample is around 0.63 billion and the average

[19]We follow Ilhan et al. (2020) and use the oil prices available at the Energy Information Administration (EIA) at `https://www.eia.gov/dnav/pet/hist/rclc1D.htm`.

book-to-market ratio is 72.7%. The mean firm leverage is about 24.3%, while there is some variation in the $\beta_{it}^{Oil}$ that has a standard deviation of 0.47. Carbon emissions also have a substantial standard deviation.

### 2.4.4 EDGAR Attention Measure

To measure market attention to news of climate regulatory policy releases in SEC filings, we follow the literature on investor attention. Our main measure of investor attention comes from the publicly available EDGAR server log, which records U.S. investor downloads of SEC filings. Investor downloads are identified by the CIK number and, therefore, can be merged with COMPUSTAT and CRSP.

We strictly follow the three EDGAR attention measures to SEC filings proposed by Drake et al. (2015), Loughran and McDonald (2017) and Ryans (2017). These approaches exclude the commonly confounding factor that large institutions may impose robot-based searching, instead of manual searches, to access SEC filings. Arguably, robot-based searching does not represent investor attention to news in SEC filings. Intuitively, these methods exclude extreme high-frequency downloads (page reviews) within a limited time. The three different measures are defined as follows:

1. DRTpv represents the EDGAR page review identified in Drake et al. (2015). This measure excludes any users (identified by their IP addresses) who have conducted more than 1,000 queries during a single day, as well as users who performed more than 5 queries per minute.

2. LMpv represents the EDGAR page review identified in Loughran and McDonald (2017). This measure excludes any users (identified by their IP addresses) who have conducted more than 50 queries during a single day.

3. Rpv represents the EDGAR page review identified in Ryans (2017). This measure excludes any users (identified by their IP addresses) who have downloaded more than 25 items during one minute, any users who have downloaded documents related to more than 3 different companies in a single minute, and users who have downloaded more than 500 items in a single day.

Table 2.2, Panel B reports summary statistics for these three measures. Overall, Drake et al. (2015) has the highest mean with 54 reviews and the highest standard deviation (931). Loughran and McDonald (2017) has an average 38 page reviews with a standard deviation equal to 894, while the Ryans (2017) data set has an average of 44 page reviews, with the smallest standard

deviation (750). Importantly, these measures display a large standard deviation, implying that there is significant variability in the sample.

### 2.4.5 Bloomberg Readership

Alternatively, we obtain the readership data of news relating to S&P500 stocks from Bloomberg, following Ben-Rephael et al. (2017). For each individual firm, we obtain the number of users who used a Bloomberg terminal to search for news about a particular firm.[20] While the original attention measure is available at an intraday frequency, we aggregate it at the monthly frequency for our analysis. Table 2.2, Panel B also shows summary statistics for this investor attention measure.

## 2.5 Climate Regulatory Exposure

Before moving on to empirical results, we validate our CRE measure in three ways: Section 2.5.1 provides examples of the word cloud that is used to determine the CRE but also the CCE measure, Section 2.5.2 discusses the time trend of the aggregate CRE and also compares this to aggregate CCE, and Section 2.5.3 explores the measures in the cross-section of industries.

### 2.5.1 Word Clouds

We start by reporting the word clouds associated with climate regulatory exposures and climate change exposures. Figure 2.1 is a word cloud that gives a visual picture of climate-related n-grams conditional on a GPV word within a 10-word distance; this will be used to construct the CRE measure. In Figure 2.1, the terms are assigned a font size proportional to their frequency in the corpus of documents. An important risk topic in the firms' view is "carbon emissions" since the n-grams $CO_2$, LCFS (Low Carbon Fuel Standard), PFCs (Perfluorinated Compound) and $CH_4$ (methane) show up frequently. Our proposed measure is therefore consistent with Bolton and Kacperczyk (2020a) and Görgen et al. (2019) to the extent that firms' primary climate regulatory concerns are related to carbon emissions. Besides n-grams relating to carbon risk, phrases like "renewable technology" and "sustainability" are also prominent in climate regulatory exposure concerns.

The distinction between climate-related concerns versus regulatory-related concerns is more clearly seen when comparing the top 30 words matched with the proposed algorithm, which

[20] Bloomberg selects the firms with the highest values of user attention as a measure of firm news hotness at an intraday frequency

are listed in Online Appendix A (Table A1). The CCV list of words includes “sustainable", “ch4" and “ozone", while the GPV list includes words relating mostly to “politics", “regulatory", “legislative", and “federal".

Figure 2 is a word cloud that will be used to construct climate change exposure. The words are fewer and different from Figure 2.1. The most frequent word is “oil", followed by “natural gas", “reserves", “weather" and then “emissions". As seen from Equation (2.3), these words are in the climate dictionary but do not contain words relating to government regulation near them. As shown in more detail in the Online Appendix (Table A2) the CCE list of words is sufficiently different from the CRE one (in both name and frequency) to provide us comfort that the two measures are capturing different variation.

### 2.5.2 Time Series Variation

Figure 2.3 shows the aggregate CRE and CCE (adjusted to start from the same point in 2006). The vertical lines represent the periods when investors started petitioning the SEC to require firms to provide more disclosures on their climate risk exposures.[21] The Figure shows that firm-disclosed CREs rise even before the initial petitions, but the upward trend accelerates after the petitions, and remains elevated after June 2008. In addition, firms disclose climate-related information more widely after September 18$^{th}$, 2007, a reflection of the fact that climate risk information in Item 1 of 10-K filings was relatively scarce prior to that date. The CCE measure is also rising during this period but not at the same rate as CRE. Of course, there is a positive correlation between the two measures and that is why we need to control for CCE in our analysis. Interestingly, in 2016 when our analysis is focussed on, the two aggregate measures diverge, with the CRE measure increasing in 2016 but the CCE measure decreasing.

Table 2.3, Panel A reports selected summary statistics for CRE. There is significant time-series variation in firm-disclosed climate information. Furthermore, the aggregate CRE trend is positive, consistent with the anecdotal evidence that investors request more climate-related disclosures after 2007. We find that before 2008, the median CRE is zero, indicating that firms did not discuss their exposure to climate regulations that often. After 2009 the annual medians of CRE are larger than 1 word in a document with word length equal to 10,000. Therefore, we focus on the sample after January 2009 in our analysis.

Table 2.3, Panel B summarizes the statistics for the sample between January 2009 and December 2018, but also reports statistics for the top 3 CRE industries. Overall, there are different patterns between the top 3 CRE industries (petroleum and natural gas, coal and utilities) and the rest. The sample with the top 3 industries has the highest average CRE (about 135

[21] These investor petitions are available at `https://www.sec.gov/rules/petitions.shtml`.

bps), whereas the CRE average is much lower (12.15 bps) for the sample excluding the top 3 industries. The conclusions are even starker when comparing medians (119.2 versus 1.56).[22]

### 2.5.3 Cross-Industry Variation

We also explore CRE and CCE heterogeneity across industries as a simple validation metric. Figure 2.4 reports the CRE measure aggregated by industry. We first note that the industry with the highest CRE is "Mining (Oil, Coal)", exceeding by far all other industries. This is internally consistent with our intuition that oil, coal and energy industries have the highest climate regulatory exposure. The "Transportation, Gas and Electricity" industry comes second over most of the period (from 2008 onwards). The CRE measure for both industries rises fast between 2007 and 2010. Interestingly, the finance and insurance industry has a high CRE measure, presumably because the industry partly trades in commodities.[23] We conclude that the CRE measure is consistent with our ex ante expectations in terms of its cross-sectional heterogeneity.

Table A4 displays the statistics using the 49 Fama-French industry categorization. Consistent with our intuition, "Petroleum and Natural Gas" and "Coal" (industries 30 and 29, respectively) exhibit the highest CRE (160.92 and 124.25, respectively). Industries with low CRE values are the medical equipment, pharmaceutical products and computer software industries. Table A3 in the Online Appendix reports the same statistics for CCE and the two highest CCE industries are again "Petroleum and Natural Gas" and "Coal", indicating the need to control for CCE in the analysis that follows.

### 2.5.4 CRE External Validation

Sautner et al. (2020) provide multiple measures of climate-related exposures (physical risk, opportunity risk and climate regulatory risk) from firm quarterly earning conference calls. We now investigate how our measures Another concern might be that the CRE measure is similar to the existing measures in the literature and how much other measure explain the SEC-based CRE quantitatively. One benchmark that has been widely accepted is Figure 2.6 displays the time-varying cross-sectional correlation between SEC-based climate regulatory exposure with annual-based measure from Sautner et al. (2020). Panel A shows that the dynamic cross-sectional correlation between CRE and climate change exposure is at least around 15% historically. This relatively high correlation implies that earning conference call based climate change exposure

[22] We report a similar table for CCE in the Online Appendix.

[23] We have repeated the analysis including the finance and insurance industry with similar results that we do not report for brevity; the results are available upon request.

is partially affected by climate regulatory concern. Importantly, Panel B confirms that our proposed CRE has low correlation with physical climate exposure with merely 2% to 7% cross sectional correlation dynamically. In contrast, Panel C shows a 9% to 14% correlation with green innovation (opportunity exposure). Panel D displays a comparably high (10% to 16%) correlation with SEC-based climate regulatory exposure (CRE). These high correlations in Panel C and Panel D are partially due to the endogeneity between earning conference call based exposure and SEC based exposure.

In addition, Table 2.5 shows the explanatory power of pre-existent variables including climate regulatory exposure, climate change exposure from earning conference call in explaining the SEC climate regulatory variation. In column (1), it shows that there is significant relationship between earning conference call based climate change exposure and CRE. This positiveness implies that the high exposure to climate risk from earning conference call is largely due to the climate regulatory concerns. However, the confounding variables including physical risk exposure and opportunity risk exposure from earning conference call have no significant relationship with CRE in (2) and (3). Importantly, consistent with cross sectional dynamic correlation, our proposed CRE is strongly and positively correlated with the transition risk and the variable is to some extent validated in this way.

## 2.6 Empirical Results

In the previous section we argued that SEC filings provide sufficient information to quantify firm-disclosed climate regulatory exposure. In this section, we investigate the asset pricing implications of such CRE disclosures.

### 2.6.1 Donald Trump Election

To address endogeneity concerns, we consider a diff-in-diff around the 2016 U.S. presidential election, as outlined in our identification strategy (refer to Section 2.2.2 set out above). We use the election of Donald Trump as an event that provides an exogenous variation in perceptions of CRE, which would allow us to establish a causal link between climate regulatory exposure and stock returns.

The main mechanism of our identification is therefore a diff-in-diff regression that uses the election of Donald Trump as an exogenous shock and separates firms between treatment and control based on their CRE index. Specifically, we classify firms into treatment groups with higher $CRE_{it}$, $Treat_{it} = 1$ if their CRE is greater than either the 80th, or the 60th or the 50th percentile, and use the firms with a CRE measure lower than the respective 20th, 40th, or 50th

percentile as control groups at time t. We denote $After_t$ as an indicator variable equal to 1 if $t$ occurs after the 2016 U.S. election (November 6$^{\text{th}}$, 2016), and 0 otherwise. In addition, we control for the common variation in firm characteristics including the size, book-to-market ratio, leverage, revenue and their interactions.

In line with our previous analysis, we consider different measures of stock returns as a dependent variable including cumulative raw returns or cumulative abnormal returns (CAPM and Fama-French-3 factor models). Moreover, we perform robustness tests with regard to the number of months used before and after the 2016 U.S. election, with different CRE percentiles when constructing the treatment and control groups, and with different training samples for abnormal returns.

**Cumulative Raw Returns**

The main regression is as follows:

$$CR_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 After_t + \beta_3 Treat_{it} \times After_t + \beta_4 Control_{it} + \epsilon_{it}. \tag{2.4}$$

$CR_{it}$ denotes the two (three) month stock returns immediately before, and after, the Trump election in 2016.[24] Table 2.6 shows the cumulative stock returns diff-in-diff regression on CRE for this specification. To emphasize the robustness of our results, the treated (control) group is defined as the firms with CRE larger (smaller) than the 50$^{\text{th}}$, 60$^{\text{th}}$ and 80$^{\text{th}}$ (50$^{\text{th}}$, 40$^{\text{th}}$ and 20$^{\text{th}}$) percentiles.

Table 2.6, Panel A shows the regression results with 2-month cumulative stock returns. Overall, the cumulative stock return in the treated group is at least 2.4% higher than the control group, depending on the specification of treated and control groups. The positive effect is statistically significant across all regressions and has the highest economic magnitude, 3.3%, when we define the treated (control) group with the CRE greater (smaller) than the 80$^{\text{th}}$ (20$^{\text{th}}$) percentile. As expected, the economic magnitudes of the diff-in-diff for the 60$^{\text{th}}$ percentile and the median regressions (columns (4) and (6)) are smaller than the 80$^{\text{th}}$ percentile (columns (1) and (2)). The magnitude can reach 4.1% (column (1)).

Table 2.6, Panel B documents a similar statistically significant and positive causal relationship between 3-month cumulative stock returns and CRE. The economic magnitudes essentially include the effects of an additional month, and therefore each coefficient is slightly larger than the respective responses in Panel A. We conclude that these results provide preliminary evidence that a perceived reduction in CRE risk after the 2016 Trump election resulted in a positive effect on stock returns for firms with higher CRE exposure.

[24]Thus, in the 2-month specification, we are comparing the returns between end of November and end of December with the returns between end of September and end of October.

Given our main result in Table 2.6 that CRE causally affects stock returns, and given our findings in Table 5 that CRE and CCE are positively correlated, one plausible explanation is that our findings are instead driven by CCE. To investigate this interpretation, we include the $CCE_{i,t}$ measure as an additional variable to control for this variation. Table 2.7 shows that our main results still hold with a similar economic magnitude after controlling for CCE. Both the economic magnitude and the statistical significance of the coefficient of interest (exposure to CRE) remain almost identical to the ones reported in Table 6.

**Cumulative Abnormal Returns**

To control for any possible risk adjustment concerns, we perform a similar diff-in-diff analysis using cumulative abnormal returns (CARs). To construct CARs, we first compute CAPM and FF3 betas based on a training sample from January 2006 to January 2015. We then compute the abnormal returns as the difference between realized and expected returns based on these two models, respectively. Finally, we estimate the following diff-in-diff regression:

$$CAR_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 After_t + \beta_3 Treat_{it} \times After_t + \beta_4 Control_{it} + \epsilon_{it}. \tag{2.5}$$

$CAR_{it}$ represents the CARs based either on the CAPM model or the FF3 models for firm i for the two (three) months immediately before, and after, the Trump election in 2016.

Table 2.8, Panel A reports the CARs based on the CAPM. Consistent with our previous work, we classify firms with a CRE measure greater than the $80^{\text{th}}$ ($60^{\text{th}}$) percentile as the treated group. The firms with CRE lower than the $20^{\text{th}}$ ($40^{\text{th}}$) percentiles are the control group. The coefficient of the interaction term is positive, indicating that CARs of high-CRE firms increase more than their low-CRE counterparts after the election of Donald Trump. This finding is robust and even stronger when we switch to the $80^{\text{th}}/20^{\text{th}}$ percentile treatment and control construction, and holds regardless of whether the additional control variables are included in the specification, or not. The magnitude is also economically significant: if we include all control variables and implement the treatment defined by the $60^{\text{th}}$ percentile CRE, we find a 2.3% (3.2%) impact on 2-month (3-month) CARs, as can be seen in columns (4) and (8), respectively. Moreover, the economic magnitude for the $80^{\text{th}}$ percentile treatment definition is higher than the $60^{\text{th}}$ percentile ones (for example, columns (2) and (4), or columns (6) and (8)). This identifies a consistent pattern as initially seen in Table 2.6, namely that the effect is positive for the climate-sensitive industries with higher CRE like oil, gas, utilities and coal.

Table 2.8, Panel B reports the results when we compute abnormal returns based on the FF3 model. The results found in Panel A using CAPM abnormal returns hold: we find a positive difference in differences of 2.7% and 3.6% for 2-month (column (4)) and 3-month (column (8)) CARs when we use the $60^{\text{th}}$ percentiles to define the treated group. Additionally, all coefficients

remain statistically and economically significant. As previously found, the economic magnitude becomes larger when we use the 80$^{th}$ percentile CRE definition for the treated group.[25]

Overall, we conclude that the results pointing to a negative association between CRE and stock returns are robust whether we use cumulative stock returns, or cumulative abnormal stock returns with different models to calculate expected stock returns. This evidence is consistent with the first hypothesis. The change in investor perceptions that future climate regulatory risks would be lower (either through less stringent legislation or more lax enforceability) meant that firms with higher CRE at the time of the election could increase in value more relative to firms with lower CRE.

**Different Event Dates**

Recent evidence in Ramelli et al. (2018) questions whether the actual event is the election itself, and points out the importance of two other major events in affecting investor opinions about President Trump's plans. President Trump gave a widely followed interview in the *New York Times* on November 23rd, and subsequently surprised the markets by announcing the appointment of EPA Administrator Scott Pruitt on December 7th. This is one main reason we have used end-of-month returns in the previous section. We use either of these dates as the event date in our analysis, and also switch to using daily returns data.

We start with the November 23$^{rd}$ event and repeat the analysis. We use both cumulative raw returns and cumulative abnormal returns as in Section 6.1 for 15 and 30 days before, and after, this event. We present the results in the Online Appendix for brevity (Table A5); the conclusions remain unchanged. Specifically, comparing cumulative raw returns before and after the interview, the ex-post cumulative return in the highest CRE-treated group increases about 3.7% (4.4%) more than the control group over 15 (30) days. The cumulative abnormal returns, depending on either the CAPM or the Fama-French 3-factor model, increase by about 4.2% or 3.8% within 15 days. The economic magnitude is more significant and larger over 30 days (4.7% or 6.6%, depending on whether we use the CAPM or Fama-French 3-factor model, respectively).

Repeating this experiment with the December 7th announcement does not generate significant results. These findings indicate that the exact event date when investor perceptions were revised is uncertain, but also illustrate that information takes time to be absorbed by investors.

[25] As in the previous subsection we also control for CCE and the results remain unchanged. We do not report them here for brevity.

**Different Training Sample**

One of the concerns with constructing CARs is that the results are sensitive to the training sample change. We therefore estimate our results using three alternative training samples: a 9-year sample (2006 to 2015), a 6-year sample (2009 to 2015) and a 3-year sample (2012 to 2015).[26] For brevity, we only report the robustness results for the 2009-2015 sample; the results for the training samples starting in 2006 and 2012 are not materially different from those reported for the 2012-2015 sample and are available upon request. Table 2.9 shows that our results are robust to this alternative training sample specification, for both the cumulative abnormal returns based on the CAPM (Panel A) and the FF3 (Panel B). The coefficients remain very similar to the baseline results, are statistically significant and therefore have similar economic magnitudes.

**Alternative Explanations: Foreign Trade, Taxation and Commodity Shocks**

Our main explanation for these results involves a change in the perception of investors on the regulatory risks associated with higher climate regulatory exposure after the Trump election. However, our results might have alternative explanations. Many policy changes beyond climate regulation could have been perceived to have changed after the 2016 presidential election. Therefore, our proposed CRE measure could be correlated with these policy changes or these policy changes might have had their own effects on valuation, therefore making our empirical findings spurious.

What are these alternative explanations? First, international trade policy could be perceived to change in ways that would arguably have a greater impact on U.S. firms with substantial foreign operations, than on U.S. firms with little, or no such operations, (Wagner et al., 2018). This is therefore potentially an alternative explanation that may account for our empirical results. To control for this explanation, we therefore include $Operation(Foreign)_{it}$, which represents firm $i$'s foreign operations as a percentage of its total assets, as an additional control variable in our baseline regressions.

A second alternative potential explanation involves taxation, since Donald Trump had plans to implement a series of taxation policies to stimulate corporate investment. To exclude this as a potentially confounding factor, we include a measure of firms' exposure to corporate taxation ($ETR(GAAP)_{it}$) in our baseline regression. Following (Wagner et al., 2018), this measure captures corporate taxation exposure faced by different firms, and is defined as the tax expense ($TXT$) as a share of pretax income ($PI$); $ETR(GAAP) = \frac{TXT}{PI} \times 100$.

[26] We ensure that the training sample spans at least three years of data to guarantee a sufficiently long sample.

Finally, a third confounding factor could be a variable potentially correlated with our CRE measure. A firm's exposure to oil shocks, for example, could be a confounding factor. Firms with high CRE might be more exposed to oil shocks, and perceptions about oil prices might have changed after the election. Therefore, it could be the case that we are capturing changes in oil market perceptions rather than regulatory exposure concerns. To control for this possibility, we construct $\beta^{Oil}$ with $Ret_{it} = \beta_0 + \beta_1 R_{Market,t} + \beta^{Oil}\Delta P_{oil,t}$ following Ilhan et al. (2020). We then include $\beta^{Oil}$ as a final control variable in our baseline specification to empirically test whether our results hold after controlling for this possibility.

Table 2.10, Panel A (CAPM-based CARs) and Table 2.10, Panel B (FF3-based CARs) show the results for the baseline CRE diff-in-diff specification when we control for the three measures described above. Specifically, we control for heterogeneity in the level of foreign trade, exposure to corporate tax regulations and sensitivity to oil prices.

For the CAPM-based CARs, Table 2.10 Panel A shows that the difference in the 2-month CAR between the treatment and control groups after the election of Trump ranges from 5.1% (Column (2)) to 3.1% (Column (4)). Results for CARs show that the 3-month (2-month) difference between the CARs of firms in the treatment and control groups is 4.8% (3.0%) after the election of Trump (columns (4) and (8)). All results are economically and statistically significant. Compared to the results in Table 2.8, Panel A, we conclude that the alternative competing interpretations are not sufficient to overturn the economic and statistical significance of the CRE exposure on stock returns.

The results reported in Table 2.10, Panel B for abnormal returns based on the FF3 model are in line with their CAPM counterparts in terms of economic and statistical significance. For example, comparing the respective coefficients in column (2), Table 2.10, Panel B and the same coefficient in Table 2.8, Panel B, we find that they are 0.042 and 0.044, respectively, with the same statistical significance.

We conclude that stock returns with higher CRE exposures when the election takes place, react more positively relative to the firms with lower CRE exposures. These results cannot be explained by trade effects, taxation changes or oil commodity shocks.

### Placebo Test

Another concern relating to the validity of our results is whether some unobserved risk specifically related to presidential elections is driving our results in Table 2.8. To address this concern, we conduct a placebo test focusing on the 2012 U.S. presidential election, and present evidence both graphically but also by replicating our baseline regressions. To do so, we divide firms in two groups according to the median value of CRE each month, before and after each of the 2012

and 2016 U.S. presidential elections. We then construct the cumulative return from August to June in each election year, 2012 (the election of Barack Obama) and 2016 (the election of Donald Trump).

Figures 2.7 and 2.8 show the comparison of the cumulative raw returns around these two consecutive presidential elections. We observe following the 2016 election a marked divergence between the high- and low-CRE groups, while there is no such divergence after the 2012 election. These results imply that the high-CRE firms had higher stock returns than low-CRE firms after the election of Donald Trump, but not after that of Barack Obama.

More formally, Table 2.11 presents the diff-in-diff results using the 2012 U.S. presidential election; the results presented here are equivalent to those reported in Table 2.8 for the 2016 U.S. presidential election. In contrast to the baseline results, the results reported in Table 2.11 are never statistically significant. This placebo test provides further evidence consistent with our hypothesis, namely the idea that the Trump election provides a useful identification strategy to test our hypothesis.

### 2.6.2 Paris Agreement

Another major event that might have affected investor perceptions on CRE is the 2015 Paris Agreement (see, for example, Seltzer et al. (2020)). Figure 2.9 shows the divergence of the cross-sectional stock returns around the Paris Agreement between the treated group (higher CRE) and the control group (lower CRE) after this event. A long-short trading strategy (long the bottom-CRE decile and short the top-CRE decile) during this event could realize a cumulative stock return as high as 10.48%.

This is a different regulatory shock from the 2016 Trump election. The conference was held from Nov 30th, 2015 to Dec 12th, 2015 and it reached an international consent towards carbon abatement by 175 parties including the U.S. (174 states and the European Union). Given its global impact and more stringent climate regulatory impact, we expect a negative relationship between higher CRE and individual stock returns. Therefore, parallel to equations (2.4) and (2.5), we identify the causal relationship with the Paris Agreement event:

$$CR_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 After_t^{Paris} + \beta_3 Treat_{it} \times After_t^{Paris} + \beta_4 Control_{it} + \epsilon_{it} \quad (2.6)$$

$CR_{i,t}$ refers to cumulative stock returns (or cumulative abnormal returns) with the same controls as in (2.4) and (2.5). Table 2.12 (Table A7 and Table A8) shows the 2-month and 3-month cumulative raw (abnormal) returns before and after the Paris Agreement. We compare different treatment and control definitions with different quantiles, and we also include the CCE measure as a control variable.

Table 2.12 shows that the results for the treated group with top 20% CRE quantile are trivially statistically significant. However, when we use top 10% (bottom 10%) for the treated (control) definition, the relationship is significantly negative. The treated group decreases by 7.1% or 7.0% (5.1% or 4.7%) more than the control group over 2-month (3-month) window, depending on how we specify the controls in the different models.[27]

Overall, the Paris Agreement offers some evidence that a perception of tighter climate regulation reduces firm-level stock returns in the U.S.. Nevertheless, the evidence is weaker than the reverse change offered by the Trump election, perhaps because the Paris agreement was an international agreement that was eventually never enforced in the U.S..

### 2.6.3 Climate Regulatory Risk vs Alternative Climate Risks

A general concern to our finding is that whether alternative climate risks could drive the regulatory or deregulatory shocks. To resolve this issue, we implement the benchmark measures from Sautner et al. (2020) as control to exclude these alternative confounding risks. Sautner et al. (2020) constructs different firm-level climate risk exposure including the physical risk exposure, opportunity risk exposure and regulatory exposure from earning conference call. Table 2.13 displays the main results after fully controlling these variation around the Donald Trump election in 2016. Overall, the treated group has changed at least 3.6% (3.4%) higher than the control group after the climate deregulatory shock for two month (three month) cumulative equity return. The innovation of these regression is to include alternative climate risk exposure including the physical risk, opportunity risk and climate regulatory risk. Conditional on these control variation, our main result still holds implying that the cross sectional CRE exposure is less likely due to other climate exposure like physical risk impact than climate regulatory exposure.
Table 2.14 shows symmetric regression result with regulatory shock as in Bolton et al. (2020). Consistently, firms with higher CRE tend to have lower valuation after the policy shock materializes. This negative effect is robust against whether we include alternative risk exposure from Sautner et al. (2020) as confounding variables. In sum, these regressions show that firms with higher identified climate regulatory exposure are more negatively affected with regulatory risk.

[27]Table A7 and Table A8 in the Online Appendix repeats this analysis for cumulative abnormal returns with similar findings.

## 2.7 Supportive Evidence

Our second set of hypotheses involves the idea that investors care less about climate regulatory exposures after the Trump election, presumably because they expect that the new administration will be less vigilant in introducing new climate change regulations, or be less stringent in enforcing existing ones, or even roll some of them back.

If the firm's perception on CRE risks matters for the firm's real economic decisions, we should expect an impact on the firm's real activity through the carbon emission channel, as in Bolton and Kacperczyk (2020a). If firms are less worried about potential legal liabilities due to an expected climate "deregulation" from the Trump administration, we can expect a higher carbon emission for those firms with higher CRE exposures.

If this narrative is correct, then we would also expect that investors' attention to climate regulatory exposure will also decrease. Moreover, to the extent that we have accurate measures of investor attention, we can use the same identification strategy to test the hypothesis that investors' attention to climate regulatory exposures decreases after the 2016 U.S. presidential election.

We first test whether the firm's CRE impacts real carbon emissions using data from the Trucost dataset (Section 2.7.1). We then explore the relation between firms' exposure to climate regulations and investors' attention to such exposures (Section 2.7.2).

### 2.7.1 Carbon Emissions

Bolton and Kacperczyk (2020a) document a positive relationship between scope 1 and 2 carbon emissions and stock returns. One of the potential economic channels that can be consistent with these empirical findings is through a lower regulation cost. Firms with a higher CRE should realize that their regulatory exposure becomes less risky after the 2016 election and therefore might be less likely to curb the firm's carbon emissions. Therefore, a perceived "deregulation" of climate regulatory policy should predict higher carbon emissions for the higher CRE firms, as well as higher stock returns.

In line with our previous identification strategy, we use the following diff-in-diff regression to investigate the hypothesis.

$$CE_{it} = \beta_0 + \beta_1 After_t + \beta_2 Treat_{it} + \beta_3 Treat_{it} \times After_t + \beta_4 Control_{it} + \epsilon_{it} \qquad (2.7)$$

where $CE_{it}$ is the sum of scope 1 carbon emissions (Trucost data (Section 4.3)).

Table 2.15 reports the results. We use a two-year and a four-year sample centered around the 2016 U.S. presidential election. For each sample size, we estimate model (2.7) for firms from the top 90% (80%) as treated and bottom 10% (20%) as control. The treatment and control groups are defined based on the quantiles of CRE cutoff values at election time.

We find that high CRE firms have higher carbon emissions after the election, as all diff-in-diff results are statistically and economically significant. When considering the full sample in the November 2014 to November 2018 window (column (1)), and for the 90% treatment definition, the treatment group's emissions increase by 3.826 ($10^5$) tCO2 more than the control group. When we consider a one-year window centered around November 2016 (column (2)), the results are similar, albeit less statistically significant but with a larger economic magnitude (5.065 ($10^5$) tCO2). The mean of carbon emissions ($CO_2$) is 29.25 across all firms. Given a coefficient in column (1) equal to 3.826, the treated group tends to increase carbon emissions by 13% (3.826/29.25) over the two-year specification. The equivalent increase for the 1-year window is 17.3% (5.065/29.25). Similarly, when we switch to regressions with 80% as treatment and bottom 20% as control, we obtain similarly statistically and economically significant results (2.006 ($10^5$) $tCO_2$ and 2.728 ($10^5$) $tCO_2$, respectively, in columns (3) and (4)).

Overall, this analysis presents evidence consistent with the hypothesis that firms with larger CRE at the time of the Trump election, tend to increase carbon emissions more relative to their lower CRE counterparts, after the election. These findings support the narrative that these firms are less concerned about climate litigation risks after that election. Moreover, this provides an economic mechanism through regulatory climate risk exposure may rationalise the positive relationship between carbon emissions and stock returns documented by Bolton and Kacperczyk (2020a).

### 2.7.2 Investor Attention

We use two different sources of data to measure attention: the first is constructed based on the intensity of downloads of 10-K files from the EDGAR server, while the second uses the readership of Bloomberg news. These are not explicit measures of attention to climate regulatory risks at the firm level, which would have been the ideal measure. We cannot construct this topic-specific measure of attention and we therefore construct a proxy that should be correlated with the ideal measure. The different proxies we use conflate attention to climate regulatory risks with attention to other firm fundamentals.

We use three different measures to identify manual searches of the EDGAR database (as opposed to automated searches) to capture a potential effect through investor attention: the measure constructed by Drake et al. (2015) (DRTpv), the measure constructed by Loughran

and McDonald (2017) (LMpv), and the measure constructed by Ryans (2017) (Rpv).[28] We follow a diff-in-diff approach with the treatment and control groups defined as in our baseline specification. The regression is as follows:

$$Attention_{it} = \beta_0 + \beta_1 Treat_{it} + \beta_2 After_t + \beta_3 Treat_{it} \times After_t + \beta_4 Control_{it} + \epsilon_{it}. \quad (2.8)$$

Table 2.16 (Columns (1) to (6)) reports the results. We find that all three measures of EDGAR attention yield similar estimates pointing towards a decrease in investor attention for high-CRE firms after the Trump election. We find a 3.1% to 4% reduction in monthly EDGAR manual searches for high-CRE firms, when compared with low-CRE firms. Since the average of the log EDGAR searches is between 3.67 to 3.9 (depending on which EDGAR attention measure we use), this implies a 70- to 80-basis point reduction in monthly EDGAR searches in the treated group, relative to the control group, after the election. This is an economically significant magnitude and is consistent with the hypothesis that investors are less concerned with risks associated with the high-CRE firms, always relative to the low-CRE ones.

The second measure of investor attention is from Bloomberg readership and is based on Ben-Rephael et al. (2017). This measure (whose descriptive statistics are presented in more detail in Table 2.2, Panel B) is considered to be a proxy for institutional investor attention because it is based on Bloomberg, and is provided directly by Bloomberg.

We re-estimate model (2.8) using the Bloomberg measure of investor attention; Table 2.16 (column (7) and column (8)) presents the results of this estimation. We obtain qualitatively similar results as with the EDGAR attention measures: we find that after the election, high-CRE firms receive less institutional investor attention. In terms of economic magnitude, a diff-in-diff of 10.2% implies that the treated group has 0.102 units lower institution investor attention than the control group. Given that the average institutional investor attention with Bloomberg readership is about 2.8, this estimate is economically significant, indicating a non-trivial decrease in the institutional investor attention.

The finding that, after the election, investors pay less attention to high-CRE firms (relative to low-CRE firms) is consistent with the narrative of a decrease in the perceived climate regulatory exposure.

[28]The difference between these measures is the algorithm through which the respective authors categorize EDGAR searches as manual or automated. Refer to Section 2.4.4 for a detailed discussion of the assumptions underlying these measures.

## 2.8 Conclusion

We explore the text information in 10-K forms submitted to the SEC and measure CRE in the U.S. stock market. Using the Trump presidential election as an exogenous surprise shock to perceived climate regulatory exposures, we find that our proposed CRE has an economically and statistically important effect on U.S. firm-level stock returns. Stock returns for firms with high CRE at the time of the election increase by around 240 to 410 basis points relative to firms with low CRE over a two month period, even after controlling for CCE exposure. Our results cannot be explained by changes in trade and taxation policies or oil commodity shocks, but future work can explore whether the valuation changes are arising from expected cash flow shocks or expected discount rate changes. Consistent with the hypothesis that investors are less worried about climate regulatory risks, we find that firms with higher CRE at election time, have higher carbon emissions and gather lower investor attention after the election. We interpret these findings as evidence that climate regulatory risks are increasingly important in valuation.

# Tables

Table 2.1: Summary Statistics of 10-K Filings

| Year | Firm | Doc | Doc Len |
|---|---|---|---|
| 2006 | 8,291 | 8,335 | 4,819 |
| 2007 | 8,097 | 8,205 | 5,011 |
| 2008 | 8,264 | 8,320 | 5,928 |
| 2009 | 9,346 | 9,463 | 5,501 |
| 2010 | 8,802 | 8,938 | 5,702 |
| 2011 | 8,481 | 8,604 | 6,038 |
| 2012 | 8,090 | 8,173 | 6,264 |
| 2013 | 7,748 | 7,845 | 6,564 |
| 2014 | 7,692 | 7,788 | 6,929 |
| 2015 | 7,590 | 7,662 | 7,300 |
| 2016 | 7,204 | 7,282 | 7,603 |
| 2017 | 6,863 | 6,943 | 7,766 |
| 2018 | 6,627 | 6,720 | 8,164 |

This table reports summary statistics for the documents (Items 1 and 1A in 10-K Filings) in our sample by year. **Firm** represents the number of unique firms in our sample identified by their Central Index Key (CIK) number, **Doc** is the number of 10-K filings, **Doc Len** is the average number of words in a document.

Table 2.2: Summary Statistics of Firm Characteristics

Panel A: Firm Characteristics from 2009 to 2018

| Statistic | N | Mean | St. Dev. | 25% | 75% |
|---|---|---|---|---|---|
| *Returns* | 115,323 | 0.017 | 0.168 | −0.054 | 0.072 |
| $log(Size)$ | 115,321 | 13.361 | 2.135 | 11.857 | 14.824 |
| $B/M$ | 109,232 | 0.727 | 0.983 | 0.286 | 0.864 |
| *Leverage* | 111,295 | 0.243 | 0.301 | 0.015 | 0.367 |
| $log(Sales)$ | 111,359 | −1.750 | 1.084 | −2.157 | −1.103 |
| $log(PPEGT)$ | 115,323 | 5.656 | 2.575 | 3.902 | 7.467 |
| $Operation(Foreign)$ | 31,234 | 3.635 | 6.634 | 0.437 | 4.365 |
| $ETR(GAAP)$ | 31,234 | 24.600 | 16.825 | 9.553 | 35.666 |
| $\beta^{Oil}$ | 115,323 | -0.0204 | 0.466 | -0.199 | 0.168 |
| *Carbon Emissions Scope*1 ($10^5$) | 47,096 | 29.258 | 118.170 | 0.498 | 10.754 |
| *Carbon Emissions Scope*2 ($10^5$) | 47,096 | 3.173 | 10.426 | 0.110 | 1.920 |

Panel B: Investor Attention and Information Environment Measure

| Statistic | N | Mean | St. Dev. | 25% | 75% | Source |
|---|---|---|---|---|---|---|
| DRTpv | 14,827,137 | 54.361 | 931.096 | 1 | 12 | EDGAR |
| LMpv | 14,827,137 | 38.420 | 893.724 | 1 | 8 | EDGAR |
| Rpv | 14,827,137 | 44.289 | 750.062 | 1 | 10 | EDGAR |
| Read | 47,634 | 2.82 | 0.836 | 2.303 | 3.434 | Bloomberg |

This table presents summary statistics for firm variables related to firm characteristics (Panel A) and investor attention (Panel B). For all variables, the table reports the mean, standard deviation, as well as the bottom and top quartiles. *Return* represents the raw monthly return. $log(Size)$ is the natural logarithm of a firm's market capitalization and $(B/M)$ is the book-to-market ratio. *Leverage* is the book leverage of the firm, $Log(Sales)$ the natural logarithm of a firm's sales, $log(PPEGT)$ is the natural logarithm of a firm's property, plants and equipment expenditure, $Operation(Foreign)$ the pre-tax income from foreign operations as a percentage of foreign assets, $ETR(GAAP)$ the ratio of tax expense on pre-tax income, $Beta^{Oil}$ the firm's exposure to oil risk; it is defined as the coefficient, $\beta_{2,i}$, in the regression $Ret_{it} = \beta_0 + \beta_{1,i}Market + \beta_{2,i}Oil\ Return$, estimated over a 24-month rolling window. DRTpv, LMpv and Rpv represent the investor attention to 10-K filings in EDGAR as defined in Drake et al. (2015), Loughran and McDonald (2017), and Ryans (2017), respectively. *Read* represents investors' attention defined as the number of Bloomberg users who used a terminal to search for news about the firm. The sample spans January 2009 to December 2018.

Table 2.3: CRE Summary Statistics by Year

Panel A: CRE statistics from 2006 to 2018

| Year | Mean | St. Dev. | 25% | Median | 75% |
|---|---|---|---|---|---|
| 2006 | 10.941 | 33.330 | 0.000 | 0.000 | 6.257 |
| 2007 | 12.652 | 35.415 | 0.000 | 0.000 | 7.192 |
| 2008 | 15.010 | 40.852 | 0.000 | 0.619 | 7.672 |
| 2009 | 17.624 | 45.166 | 0.000 | 1.080 | 9.482 |
| 2010 | 24.928 | 55.245 | 0.000 | 2.199 | 17.905 |
| 2011 | 26.375 | 55.826 | 0.000 | 2.784 | 22.970 |
| 2012 | 26.209 | 55.514 | 0.000 | 2.814 | 22.336 |
| 2013 | 26.003 | 53.158 | 0.000 | 2.983 | 22.571 |
| 2014 | 26.602 | 54.933 | 0.000 | 2.945 | 22.022 |
| 2015 | 25.696 | 53.939 | 0.000 | 2.732 | 21.370 |
| 2016 | 24.107 | 51.052 | 0.000 | 2.688 | 19.958 |
| 2017 | 24.671 | 52.867 | 0.000 | 2.577 | 19.770 |
| 2018 | 24.566 | 53.311 | 0.000 | 2.588 | 18.944 |

Panel B: CRE Statistics from 2009 to 2018

| | Mean | St. Dev. | 25% | Median | 75% |
|---|---|---|---|---|---|
| All Industries | 24.621 | 53.193 | 0.000 | 2.503 | 24.621 |
| All Ind. Ex. Top 3 | 12.150 | 27.799 | 0.000 | 1.564 | 12.150 |
| Top 3 | 135.035 | 86.436 | 76.803 | 119.188 | 135.035 |

This table reports summary statistics for Climate Regulatory Exposure ($CRE$) for every year in our sample. Specifically, the table reports the mean, standard deviation, as well as the bottom quartile, median and top quartile. Units are in basis points; 1 word in a document with word size equal to 10,000 is 1 basis point.

Table 2.4: CRE Summary Statistics by Industry from 2006 to 2018

| Industry No | Industry Name | Mean | S.D. | Firm Count |
|---|---|---|---|---|
| 1 | Agriculture | 10.11 | 15.90 | 19 |
| 2 | Food Products | 15.80 | 25.46 | 90 |
| 3 | Candy Soda | 10.34 | 15.60 | 7 |
| 4 | Beer Liquor | 24.07 | 80.70 | 20 |
| 5 | Tobacco Products | 5.16 | 6.75 | 10 |
| 6 | Recreation | 3.40 | 6.96 | 35 |
| 7 | Entertainment | 2.85 | 4.65 | 78 |
| 8 | Printing and Publishing | 2.65 | 7.60 | 51 |
| 9 | Consumer Goods | 10.34 | 19.50 | 77 |
| 10 | Apparel | 2.83 | 5.37 | 68 |
| 11 | Healthcare | 2.16 | 4.94 | 128 |
| 12 | Medical Equipment | 2.14 | 10.34 | 275 |
| 13 | Pharmaceutical Products | 1.25 | 3.18 | 700 |
| 14 | Chemicals | 31.87 | 37.61 | 143 |
| 15 | Rubber and Plastic Products | 7.61 | 12.53 | 40 |
| 16 | Textiles | 12.15 | 13.15 | 13 |
| 17 | Construction Materials | 21.97 | 37.80 | 90 |
| 18 | Construction | 19.37 | 28.10 | 75 |
| 19 | Steel Works Etc | 27.66 | 31.90 | 82 |
| 20 | Fabricated Products | 15.97 | 19.77 | 15 |
| 21 | Machinery | 21.84 | 38.58 | 164 |
| 22 | Electrical Equipment | 14.21 | 22.25 | 76 |
| 23 | Automobiles and Trucks | 24.11 | 31.56 | 87 |
| 24 | Aircraft | 9.48 | 16.56 | 27 |
| 25 | Shipbuilding, Railroad Equipment | 15.02 | 20.57 | 13 |
| 26 | Defense | 8.24 | 11.53 | 14 |
| 27 | Precious Metals | 32.78 | 33.99 | 21 |
| 28 | Non-Metallic and Industrial Metal Mining | 54.68 | 64.78 | 36 |
| 29 | Coal | 124.25 | 49.23 | 26 |
| 30 | Petroleum and Natural Gas | 160.92 | 92.12 | 308 |
| 31 | Utilities | 78.98 | 53.94 | 201 |
| 32 | Communication | 3.23 | 7.80 | 196 |
| 33 | Personal Services | 5.93 | 15.77 | 79 |
| 34 | Business Services | 5.61 | 16.02 | 534 |
| 35 | Computers | 2.95 | 6.53 | 121 |
| 36 | Computer Software | 1.78 | 4.30 | 465 |
| 37 | Electronic Equipment | 5.02 | 9.90 | 409 |
| 38 | Measuring and Control Equipment | 4.71 | 10.10 | 131 |
| 39 | Business Supplies | 27.01 | 39.36 | 43 |
| 40 | Shipping Containers | 25.39 | 44.02 | 17 |
| 41 | Transportation | 38.57 | 39.62 | 166 |
| 42 | Wholesale | 16.37 | 40.07 | 194 |
| 43 | Retail | 8.64 | 21.96 | 294 |
| 44 | Restaurants, Hotels, Motels | 6.41 | 20.06 | 140 |
| 49 | Almost Nothing | 54.53 | 44.89 | 25 |

This table reports summary statistics for the Climate Regulatory Exposure ($CRE$) by industry. The industry classification follows the Fama-French website based on 6-digit SIC codes. For each industry, the table presents the mean and standard deviation of $CRE$ as well as the number of firms per the CRSP database classified in that industry. The sample spans January 2006 to December 2018.

Table 2.5: Determinants of SEC based CRE disclosure

| | *Dependent variable: CRE* | | | |
|---|---|---|---|---|
| | $CRE_{i,t}$% | | | |
| | (1) | (2) | (3) | (4) |
| Climate Change$_{i,t}$ | 3.962* | | | |
| | (2.154) | | | |
| Physical Risk$_{i,t}$ | | −17.725 | | |
| | | (50.709) | | |
| Opportunity$_{i,t}$ | | | 3.100 | |
| | | | (3.568) | |
| Transition$_{i,t}$ | | | | 50.317*** |
| | | | | (13.113) |
| $\beta^{Oil}_{i,t}$ | 0.044*** | 0.044*** | 0.044*** | 0.043*** |
| | (0.007) | (0.007) | (0.007) | (0.007) |
| $log(PPEGT)_{i,t}$ | 0.036*** | 0.036*** | 0.036*** | 0.036*** |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| $log(Size)_{i,t}$ | −0.036*** | −0.036*** | −0.036*** | −0.036*** |
| | (0.007) | (0.007) | (0.007) | (0.007) |
| $B/M_{i,t}$ | −0.029** | −0.030** | −0.030** | −0.030** |
| | (0.012) | (0.012) | (0.012) | (0.012) |
| $Leverage_{i,t}$ | −0.0002** | −0.0002** | −0.0002** | −0.0002** |
| | (0.0001) | (0.0001) | (0.0001) | (0.0001) |
| $log(Sale)_{i,t}$ | 0.059 | 0.057 | 0.058 | 0.056 |
| | (0.044) | (0.044) | (0.044) | (0.044) |
| $log(Sale)_{i,t} \times log(Size)_{i,t}$ | −0.006* | −0.006* | −0.006* | −0.006* |
| | (0.003) | (0.003) | (0.003) | (0.003) |
| $log(Sale)_{i,t} \times B/M_{i,t}$ | −0.007 | −0.007 | −0.007 | −0.007 |
| | (0.005) | (0.005) | (0.005) | (0.005) |
| $log(Sale)_{i,t} \times Leverage_{i,t}$ | −0.088** | −0.088** | −0.088** | −0.089** |
| | (0.036) | (0.036) | (0.036) | (0.036) |
| Industry FE | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES |
| Observations | 59,965 | 59,965 | 59,965 | 59,965 |
| Adjusted $R^2$ | 0.601 | 0.601 | 0.601 | 0.602 |

Note: This table reports the relationship between SEC-based CRE and climate exposure from earning conference call. The dependent variable, CRE, is in percentage. Climate Change$_{i,t}$, Physical Risk$_{i,t}$, $Opportunity_{i,t}$ and $Transition_{i,t}$ refer to the climate change exposure, physical risk exposure, opportunity risk exposure and climate regulatory exposure from Sautner et al. (2020). We include controls with $log(Size)$, B/M, $log(PPEGT)_{i,t}$, $log(Sale)$, Leverage, $log(Sale) \times log(Size)$, $log(Sale) \times B/M$ and $log(Sale) \times Leverage$. $log(Size)$ represents the natural logarithm of a firm's market capitalization. $B/M$ represents the firm's book-to-market ratio. $Leverage$ represents the book leverage of the firm. $log(Sale)$ refers to the logarithmic of sale for firm i. $\beta^{Oil}$ denotes the oil exposure as Table 2.2. $log(PPEGT)_{i,t}$ refers to the plant property and equipment expenditure. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.6: Difference-in-Difference around Trump Election with CRE (Cumulative Raw Return)

Panel A: Dependent variable: Cumulative Raw Returns (2 Months)

| | 80% | | 60% | | 50% | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | 0.041*** | 0.033*** | 0.034*** | 0.027*** | 0.031*** | 0.024*** |
| | (0.012) | (0.012) | (0.010) | (0.010) | (0.009) | (0.009) |
| Controls | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 11,671 | 10,321 | 16,691 | 14,892 | 20,850 | 18,664 |
| Adjusted $R^2$ | 0.639 | 0.673 | 0.597 | 0.619 | 0.557 | 0.586 |

Panel B: Dependent variable: Cumulative Raw Returns (3 Months)

| | 80% | | 60% | | 50% | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | 0.042*** | 0.048*** | 0.032*** | 0.036*** | 0.027** | 0.027*** |
| | (0.014) | (0.013) | (0.012) | (0.011) | (0.011) | (0.010) |
| Controls | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 11,708 | 10,356 | 16,730 | 14,931 | 20,902 | 18,712 |
| Adjusted $R^2$ | 0.631 | 0.672 | 0.593 | 0.621 | 0.560 | 0.588 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.4), with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $80^{\text{th}}$ ($20^{\text{th}}$), $60^{\text{th}}$ ($40^{\text{th}}$) and $50^{\text{th}}$ percentiles. $After$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2016 U.S. presidential election (November 2016), and 0 for the two (three) months before the November election. We include controls with $log(Size)$, B/M, $log(Sale)$, Leverage, $log(Sale) \times log(Size)$, $log(Sale) \times B/M$ and $log(Sale) \times Leverage$. $log(Size)$ represents the natural logarithm of a firm's market capitalization. $B/M$ represents the firm's book-to-market ratio. $Leverage$ represents the book leverage of the firm. $log(Sale)$ refers to the logarithmic of sale for firm i. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 2.7: Difference-in-Difference around Trump Election (Cumulative Raw Return)

Panel A: Dependent variable: Cumulative Raw Returns (2 Months)

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | 80% | | 60% | | 50% | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | 0.042*** | 0.035*** | 0.035*** | 0.028*** | 0.032*** | 0.024** |
| | (0.012) | (0.012) | (0.010) | (0.010) | (0.010) | (0.010) |
| $CCE_{i,t}$ | 6.330 | 12.934 | 7.252 | 11.081 | 0.722 | 1.486 |
| | (9.863) | (8.886) | (9.096) | (8.453) | (1.019) | (1.005) |
| Controls | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 11,691 | 10,340 | 16,711 | 14,909 | 20,882 | 18,689 |
| Adjusted $R^2$ | 0.641 | 0.676 | 0.598 | 0.620 | 0.111 | 0.132 |

Panel B: Dependent variable: Cumulative Raw Returns (3 Months)

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | 80% | | 60% | | 50% | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | 0.047*** | 0.048*** | 0.029** | 0.029** | 0.028** | 0.023** |
| | (0.016) | (0.015) | (0.013) | (0.013) | (0.011) | (0.011) |
| $CCE_{i,t}$ | 1.711 | 2.407* | 1.571 | 2.231* | 1.016 | 1.818 |
| | (1.407) | (1.330) | (1.293) | (1.227) | (1.307) | (1.252) |
| Controls | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 11,724 | 10,371 | 16,754 | 14,948 | 20,934 | 18,737 |
| Adjusted $R^2$ | 0.095 | 0.114 | 0.093 | 0.109 | 0.097 | 0.113 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.4), with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $80^{\text{th}}$ ($20^{\text{th}}$), $60^{\text{th}}$ ($40^{\text{th}}$) and $50^{\text{th}}$ percentiles. $After$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2016 U.S. presidential election (November 2016), and 0 for the two (three) months before the November election. We include controls with $log(Size)$, B/M, $log(Sale)$, Leverage, $log(Sale) \times log(Size)$, $log(Sale) \times B/M$ and $log(Sale) \times Leverage$. $log(Size)$ represents the natural logarithm of a firm's market capitalization. $B/M$ represents the firm's book-to-market ratio. $Leverage$ represents the book leverage of the firm. $log(Sale)$ refers to the logarithmic of sale for firm i. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.8: Difference-in-Differences (Cumulative Abnormal Returns)

Panel A: Dependent Variable: CAPM Cumulative Abnormal Returns

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.044*** | 0.037*** | 0.029** | 0.023** | 0.046*** | 0.054*** | 0.027** | 0.032*** |
| | (0.013) | (0.012) | (0.011) | (0.011) | (0.016) | (0.014) | (0.014) | (0.012) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 11,624 | 10,295 | 16,620 | 14,845 | 11,657 | 10,326 | 16,663 | 14,884 |
| Adjusted $R^2$ | 0.566 | 0.637 | 0.524 | 0.571 | 0.559 | 0.657 | 0.526 | 0.594 |

Panel B: Dependent Variable: FF3 Cumulative Abnormal Returns

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.056*** | 0.044*** | 0.039*** | 0.027** | 0.053*** | 0.057*** | 0.035*** | 0.036*** |
| | (0.014) | (0.014) | (0.012) | (0.012) | (0.015) | (0.014) | (0.013) | (0.012) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 11,389 | 10,092 | 16,304 | 14,572 | 11,422 | 10,123 | 16,347 | 14,611 |
| Adjusted $R^2$ | 0.628 | 0.653 | 0.593 | 0.601 | 0.673 | 0.695 | 0.635 | 0.646 |

This table reports estimation results for the difference-in-differences regression defined in Equation (2.4), with the cumulative abnormal returns as the dependent variable. Panel A (Panel B) presents results for CAPM (FF3) abnormal returns, which are defined as per Equation (2.5). The training sample spans January 2006 to January 2015. In each Panel, Columns (1) to (4) present results for a horizon h = 2 months, and Columns (5) to (8) for a horizon h = 3 months. $After$ is an indicator variable equal to 1 after the 2016 U.S. presidential election (November 2016) and 0 before it. For each horizon, we present two alternative specifications for the treatment and control groups ($Treat$ is an indicator variable equal to 1 for firms in the treatment group and 0 otherwise): (a) treatment (control) consists of firms greater than the $80^{\text{th}}$ (lower than the $20^{\text{th}}$) CRE percentile; (b) treatment (control) consists of firms greater than the $60^{\text{th}}$ (lower than the $40^{\text{th}}$) CRE percentile. Controls are the same as in Table 2.6. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.9: CAR Regression on CRE for a Different Training Sample (2009-2015)

Panel A: Dependent Variable: CAPM Cumulative Abnormal Returns

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.042*** | 0.035*** | 0.026** | 0.020* | 0.044*** | 0.051*** | 0.023* | 0.028** |
| | (0.013) | (0.012) | (0.011) | (0.011) | (0.016) | (0.014) | (0.014) | (0.012) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 11,600 | 10,272 | 16,596 | 14,824 | 11,633 | 10,303 | 16,639 | 14,863 |
| Adjusted $R^2$ | 0.567 | 0.638 | 0.526 | 0.572 | 0.561 | 0.658 | 0.527 | 0.595 |

Panel B: Dependent Variable: FF3 Cumulative Abnormal Returns

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.050*** | 0.039*** | 0.034*** | 0.023* | 0.050*** | 0.054*** | 0.032** | 0.033*** |
| | (0.014) | (0.014) | (0.013) | (0.012) | (0.015) | (0.014) | (0.013) | (0.012) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 11,377 | 10,081 | 16,292 | 14,563 | 11,410 | 10,112 | 16,335 | 14,602 |
| Adjusted $R^2$ | 0.627 | 0.651 | 0.592 | 0.599 | 0.675 | 0.697 | 0.638 | 0.648 |

This table reports estimation results for the difference-in-differences regression defined in Equation (2.4), with the cumulative abnormal returns as the dependent variable. Panel A (Panel B) presents results for CAPM (FF3) abnormal returns, which are defined as per Equation (2.5). The training sample spans January 2009 to January 2015. In each Panel, Columns (1) to (4) present results for a horizon h = 2 months, and Columns (5) to (8) for a horizon h = 3 months. *After* is an indicator variable equal to 1 for the two or three months after the 2016 U.S. presidential election (November 2016), and 0 for the two (three) months before the November election. For each horizon, we present two alternative specifications for the treatment and control groups (*Treat* is an indicator variable equal to 1 for firms in the treatment group and 0 otherwise): (a) treatment (control) consists of firms greater than the $80^{\text{th}}$ (lower than the $20^{\text{th}}$) CRE percentile; (b) treatment (control) consists of firms greater than the $60^{\text{th}}$ (lower than the $40^{\text{th}}$) CRE percentile. Controls are the same as in Table 2.6. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.10: Controlling for Alternative Explanations

Panel A: Difference-in-Differences (CAPM CARs) with Additional Control Variables

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.052*** | 0.050*** | 0.034** | 0.030** | 0.052** | 0.047** | 0.034** | 0.029* |
| | (0.020) | (0.018) | (0.015) | (0.014) | (0.021) | (0.020) | (0.016) | (0.015) |
| $Operation(Foreign)_{it}$ | | 0.001 | | 0.0002 | | 0.002 | | 0.001 |
| | | (0.002) | | (0.002) | | (0.001) | | (0.001) |
| $ETR(GAAP)_{it}$ | | −0.0001 | | 0.00003 | | −0.0004 | | 0.00002 |
| | | (0.0003) | | (0.0002) | | (0.0003) | | (0.0002) |
| $\beta_{it}^{Oil}$ | | −0.002 | | −0.0002 | | 0.008 | | 0.009 |
| | | (0.030) | | (0.024) | | (0.033) | | (0.026) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Observations | 4,584 | 4,309 | 7,002 | 6,569 | 4,584 | 4,309 | 7,006 | 6,573 |
| Adjusted $R^2$ | 0.082 | 0.097 | 0.063 | 0.072 | 0.063 | 0.078 | 0.049 | 0.059 |

Panel B: Difference-in-Differences (FF3 CARs) with Additional Control Variables

| | 2-Month | | | | 3-Month | | | |
|---|---|---|---|---|---|---|---|---|
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | 0.040* | 0.042** | 0.029* | 0.026* | 0.047** | 0.044** | 0.034** | 0.030** |
| | (0.021) | (0.020) | (0.016) | (0.015) | (0.021) | (0.021) | (0.016) | (0.015) |
| $Operation(Foreign)_{it}$ | | 0.002 | | 0.0005 | | 0.002 | | 0.001 |
| | | (0.002) | | (0.002) | | (0.001) | | (0.001) |
| $ETR(GAAP)_{it}$ | | −0.0001 | | −0.0001 | | −0.0003 | | −0.0001 |
| | | (0.0003) | | (0.0002) | | (0.0003) | | (0.0003) |
| $\beta_{it}^{Oil}$ | | −0.015 | | −0.008 | | 0.001 | | 0.006 |
| | | (0.037) | | (0.029) | | (0.039) | | (0.032) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Observations | 4,500 | 4,242 | 6,898 | 6,482 | 4,500 | 4,242 | 6,902 | 6,486 |
| Adjusted $R^2$ | 0.054 | 0.062 | 0.041 | 0.045 | 0.054 | 0.068 | 0.041 | 0.049 |

Estimation results for the diff-in-diff Equation (2.4), with cumulative abnormal returns as the dependent variable. Panel A (Panel B) presents results for CAPM (FF3) abnormal returns, which are defined as per Equation (2.5). The training sample spans January 2006 to January 2015. In each Panel, Columns (1) to (4) present results for a horizon h = 2 months, and Columns (5) to (8) for a horizon h = 3 months. $After$ is an indicator variable equal to 1 for the two (three) months after the 2016 U.S. presidential election (November 2016), and 0 for the two (three) months before the election. For each horizon, we present two alternative specifications for the treatment and control groups ($Treat$ is an indicator variable equal to 1 for firms in the treatment group and 0 otherwise): (a) treatment (control) consists of firms greater than the $80^{\text{th}}$ (lower than the $20^{\text{th}}$) CRE percentile (Columns (1), (2), (5), and (6)); alternatively, (b) treatment (control) consists of firms greater than the $60^{\text{th}}$ (lower than the $40^{\text{th}}$) CRE percentile (Columns (3), (4), (7) and (8)). Controls are the same as in Table 2.6, with the addition of pre-tax income from foreign operations as a percentage of foreign assets ($Operation(Foreign)$), the ratio of tax expense on pre-tax income ($ETR(GAAP)$) and the firm's exposure to oil risk ($Beta^{Oil}$), which is defined as the coefficient, $\beta_{2,i}$, in the regression $Ret_{it} = \beta_0 + \beta_{1,i}Market + \beta_{2,i}Oil\ Return$, estimated over a 24-month rolling window. Industry fixed effects (FE) follow the Fama-French definitions, excluding the financial service industries. The main result sill holds after interacting the confounding variables with the indicator, $After_t$. We put industry FE and Year FE as previous tablesStandard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.11: Placebo Test: 2012 U.S. Presidential Election

| | *Dependent variable: CAR (3 Months)* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | CAPM | | | | FF3 | | | |
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_{it} \times After_t$ | −0.012 | −0.015 | −0.007 | −0.012 | −0.013 | −0.015 | −0.005 | −0.008 |
| | (0.011) | (0.011) | (0.010) | (0.010) | (0.012) | (0.011) | (0.010) | (0.010) |
| Controls | NO | YES | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 12,632 | 11,767 | 17,318 | 16,247 | 12,556 | 11,694 | 17,206 | 16,146 |
| Adjusted $R^2$ | 0.572 | 0.583 | 0.556 | 0.556 | 0.540 | 0.558 | 0.535 | 0.541 |

This table reports estimation results for the difference-in-differences regression defined in Equation (2.4), with the cumulative abnormal returns as the dependent variable. Columns (1) and (2) (Columns (3) and (4)) present results for CAPM (FF3) 3-month cumulative abnormal returns, which are defined as per Equation (2.5). The training sample spans January 2009 to January 2012). We present two alternative specifications for the treatment and control groups ($Treat$ is an indicator variable equal to 1 for firms in the treatment group and 0 otherwise): (a) treatment (control) consists of firms greater than the $80^{\text{th}}$ (lower than the $20^{\text{th}}$) CRE percentile (Columns (1), (2), (5) and (6)); alternatively, (b) treatment (control) consists of firms greater than the $60^{\text{th}}$ (lower than the $40^{\text{th}}$) CRE percentile (Columns (3), (4), (7) and (8)). $After$ is an indicator variable equal to 1 for the two months after the 2012 U.S. presidential election (November 2012), and 0 for the two months before the election. Controls are the same as in Table 2.6. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^*p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 2.12: Difference-in-Difference around Paris Agreement (Cumulative Raw Return)

Panel A: Dependent Variable: 2-Month Cumulative Raw Returns

| | *Dependent variable:* $CR^{2M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | −0.071*** | −0.070*** | −0.070*** | −0.024 | −0.029* | −0.030* |
| | (0.022) | (0.020) | (0.020) | (0.019) | (0.017) | (0.017) |
| $CCE_{i,t}$ | | | 22.111 | | | 54.888 |
| | | | (28.342) | | | (38.425) |
| Control | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,360 | 5,706 | 5,706 | 8,998 | 8,110 | 8,110 |
| Adjusted $R^2$ | 0.827 | 0.566 | 0.566 | 0.769 | 0.517 | 0.519 |

Panel B: Dependent Variable: 3-Month Cumulative Raw Returns

| | *Dependent variable:* $CR^{3M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | −0.051** | −0.047** | −0.047** | −0.001 | −0.004 | −0.004 |
| | (0.025) | (0.023) | (0.023) | (0.021) | (0.019) | (0.019) |
| $CCE_{i,t}$ | | | 26.331 | | | 62.192 |
| | | | (32.548) | | | (44.317) |
| Control | NO | YES | NO | YES | NO | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,377 | 5,719 | 5,719 | 9,023 | 8,130 | 8,130 |
| Adjusted $R^2$ | 0.806 | 0.603 | 0.603 | 0.741 | 0.537 | 0.539 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.6), with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $90^{\text{th}}$ ($10^{\text{th}}$) and $80^{\text{th}}$ ($20^{\text{th}}$) percentiles. $After^{Paris}$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2015 Paris Agreement (December 2015), and 0 for the two (three) months before the Paris Agreement. Controls are the same as in Table 2.6. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.13: Difference-in-Difference around Donald Trump Election with Confounding Climate Variation

| | *Dependent variable: Cumulative Raw Return (2M/3M)* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $CR^{2M}$ | | | | $CR^{3M}$ | | | |
| | 80% | | 60% | | 80% | | 60% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_i \times After_t$ | 0.061*** | 0.051*** | 0.046*** | 0.036*** | 0.050*** | 0.051*** | 0.034*** | 0.034*** |
| | (0.014) | (0.012) | (0.010) | (0.010) | (0.016) | (0.015) | (0.012) | (0.011) |
| Physical $\text{Risk}_{i,t} \times After_t$ | 5.160 | 18.511 | −27.465 | −1.809 | −55.146 | −31.260 | −53.026 | −41.127 |
| | (29.343) | (25.431) | (30.960) | (23.980) | (48.327) | (34.791) | (40.409) | (31.148) |
| $Transition \times After_t$ | −33.182*** | −39.448*** | −30.606*** | −32.400** | −25.149 | −29.785 | −27.677* | −28.728 |
| | (10.620) | (13.676) | (11.068) | (13.734) | (21.222) | (26.947) | (16.724) | (22.803) |
| $Opportunity_{i,t} \times After_t$ | −13.263*** | −12.892*** | −7.912*** | −8.264*** | −9.744*** | −9.426*** | −3.818 | −4.210 |
| | (2.588) | (2.515) | (2.371) | (2.284) | (3.112) | (3.163) | (3.261) | (3.126) |
| $\text{Physical Risk}_{i,t}$ | −59.147 | 36.722 | −17.430 | 50.614 | 66.324 | 175.044 | −20.695 | 80.863 |
| | (109.339) | (84.236) | (83.538) | (71.693) | (172.807) | (131.816) | (115.704) | (95.150) |
| $Transition \times After_t$ | −4.347 | 5.973 | 39.134** | 45.417** | 25.784 | 48.809 | 78.699*** | 91.436*** |
| | (19.007) | (20.959) | (18.585) | (21.787) | (29.969) | (35.938) | (27.392) | (33.166) |
| $Opportunity_{i,t}$ | −5.942 | −1.908 | −12.784 | −11.855 | −14.136 | −7.237 | −15.786 | −11.971 |
| | (7.191) | (6.654) | (8.015) | (7.425) | (10.336) | (9.529) | (9.618) | (8.782) |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 9,247 | 8,380 | 13,780 | 12,551 | 9,274 | 8,405 | 13,813 | 12,582 |
| Adjusted $R^2$ | 0.707 | 0.742 | 0.690 | 0.726 | 0.699 | 0.737 | 0.682 | 0.722 |

Note: This table reports estimation results for the difference-in-differences regression for stock market reaction around the Donald Trump election in 2016, with the climate measures from Sautner et al. (2020) as controls. Columns (1)-(4) present the results for two-month cumulative raw return Difference-in-Difference around Paris Agreement. Columns (5)-(8) present the results for three-month cumulative raw return Difference-in-Difference around Paris Agreement. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment group includes firms with a $CRE$ measure greater than the 80% (60%) for columns (1), (2), (5) and (6) ((3), (4), (7) and (8)). The control group includes firms with a $CRE$ measure smaller than the 20% (40%) for columns (1), (2), (5) and (6) ((3), (4), (7) and (8)). $After$ is an indicator variable equal to 1 for all months after the 2016 U.S. presidential election (November 2016), and 0 otherwise. Climate Change, Physical Risk, $Transition_{i,t}$ and $Opportunity_{i,t}$ are the same as Table 2.5. We include the following as control variables: the natural logarithm of a firm's market capitalization ($log(Size)$), its book-to-market ratio ($B/M$)), its book leverage ($Leverage$), and the natural logarithm of the property, plant and equipment expenditures ($log(PPEGT)$. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Columns (1) and (3) present the results for all industries, and Columns (2) and (4) present the results for industries excluding the three industries with the highest average CRE. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.14: Difference-in-Difference around Paris Agreement with Confounding Climate Variation

| | *Dependent variable: Cumulative Raw Return (2M/3M)* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | $CR^{2M}$ | | | | $CR^{3M}$ | | | |
| | 90% | | 80% | | 90% | | 80% | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_i \times After_t$ | −0.111*** | −0.108*** | −0.047** | −0.052*** | −0.083*** | −0.080*** | −0.013 | −0.020 |
| | (0.024) | (0.023) | (0.020) | (0.018) | (0.027) | (0.026) | (0.023) | (0.021) |
| Physical $\text{Risk}_{i,t} \times After_t$ | 125.419 | 62.002 | 85.162 | 56.797 | 112.698 | 53.429 | 77.056 | 64.284 |
| | (96.588) | (92.109) | (58.050) | (57.222) | (108.201) | (98.611) | (62.807) | (57.091) |
| $Transition \times After_t$ | 24.597 | 22.769 | −7.320 | −3.919 | 56.535 | 49.999 | −4.585 | −0.681 |
| | (58.488) | (55.700) | (15.094) | (16.693) | (52.550) | (49.989) | (20.085) | (20.430) |
| $Opportunity_{i,t} \times After_t$ | 25.267*** | 23.339*** | 25.208*** | 23.240*** | 16.734** | 14.465** | 17.327*** | 15.764*** |
| | (6.340) | (6.041) | (5.557) | (5.367) | (7.108) | (6.593) | (5.239) | (5.127) |
| Physical $\text{Risk}_{i,t}$ | 70.711 | −8.957 | −2.528 | −43.127 | 74.369 | −15.262 | 46.628 | 5.388 |
| | (136.327) | (121.462) | (104.482) | (94.989) | (157.980) | (132.893) | (116.993) | (98.154) |
| $Transition_{i,t}$ | −29.977 | −59.815 | −12.106 | −23.799 | −43.178 | −59.529 | −5.663 | −24.723 |
| | (67.737) | (66.285) | (44.524) | (44.104) | (67.158) | (65.691) | (49.319) | (45.320) |
| $Opportunity_{i,t}$ | 17.025 | 14.139 | 1.340 | 1.328 | −3.267 | 2.111 | −1.267 | 0.242 |
| | (18.368) | (18.857) | (16.261) | (16.960) | (22.269) | (23.115) | (19.633) | (20.611) |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 4,773 | 4,338 | 7,299 | 6,663 | 4,781 | 4,346 | 7,313 | 6,676 |
| Adjusted $R^2$ | 0.587 | 0.615 | 0.547 | 0.565 | 0.601 | 0.640 | 0.535 | 0.569 |

Note: This table reports estimation results for the difference-in-differences regression for stock market reaction around the Paris Agreement, with the climate measures from Sautner et al. (2020) as controls. Columns (1)-(4) present the results for two-month cumulative raw return Difference-in-Difference around Paris Agreement. Columns (5)-(8) present the results for three-month cumulative raw return Difference-in-Difference around Paris Agreement. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment group includes firms with a $CRE$ measure greater than the 90% (80%) for columns (1), (2), (5) and (6) ((3), (4), (7) and (8)). The control group includes firms with a $CRE$ measure smaller than the 10% (20%) for columns (1), (2), (5) and (6) ((3), (4), (7) and (8)). $After$ is an indicator variable equal to 1 for all months after the 2015 Paris Agreement (Dec 2015), and 0 otherwise. Climate Change, Physical Risk, $Transition_{i,t}$ and $Opportunity_{i,t}$ are the same as Table 2.5. We include the following as control variables: the natural logarithm of a firm's market capitalization ($log(Size)$), its book-to-market ratio ($B/M$)), its book leverage ($Leverage$), and the natural logarithm of the property, plant and equipment expenditures ($log(PPEGT)$. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Columns (1) and (3) present the results for all industries, and Columns (2) and (4) present the results for industries excluding the three industries with the highest average CRE. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.15: Carbon Emission Difference-in-Differences around Trump Election

| | *Dependent variable: Carbon Emission Scope 1* | | | |
|---|---|---|---|---|
| | 90% | | 80% | |
| | [-2y,2y] | [-1y,1y] | [-2y,2y] | [-1y,1y] |
| | (1) | (2) | (3) | (4) |
| $Treat_i \times After_t$ | 3.826** | 5.065* | 2.006* | 2.728* |
| | (1.640) | (2.626) | (1.201) | (1.613) |
| $Treat_i$ | −1.142 | −2.063 | −1.695 | −2.486* |
| | (1.339) | (1.799) | (1.317) | (1.321) |
| $After_t$ | −1.029* | −1.329* | −1.185** | −1.438* |
| | (0.534) | (0.746) | (0.590) | (0.746) |
| $log(Size_{i,t})$ | 1.406* | 2.642*** | 3.399*** | 4.705*** |
| | (0.762) | (0.935) | (1.285) | (1.311) |
| B/M | 1.183 | 1.783 | −0.549 | 0.226 |
| | (0.896) | (1.088) | (1.361) | (0.911) |
| $Leverage_{i,t}$ | −0.001 | 0.0002 | 0.004 | 0.006 |
| | (0.003) | (0.003) | (0.008) | (0.005) |
| $log(PPEGT_{i,t})$ | 1.541*** | 3.022*** | 3.266*** | 4.451*** |
| | (0.551) | (0.958) | (1.171) | (1.202) |
| $log(Sale_{i,t})$ | −10.073* | −14.882*** | −19.792*** | −23.475*** |
| | (5.275) | (5.335) | (7.274) | (7.166) |
| $log(Size_{i,t}) \times log(Sale_{i,t})$ | 0.755* | 1.121*** | 1.546*** | 1.812*** |
| | (0.401) | (0.403) | (0.565) | (0.549) |
| $log(Sale)_{i,t} \times BM_{i,t}$ | 0.340 | 0.539* | −0.245 | 0.039 |
| | (0.238) | (0.309) | (0.501) | (0.324) |
| $log(Sale)_{i,t} \times Leverage$ | −0.0002 | 0.0002 | 0.003 | 0.003 |
| | (0.002) | (0.001) | (0.004) | (0.002) |
| Industry FE | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES |
| Observations | 4,014 | 7,145 | 5,278 | 9,456 |
| Adjusted $R^2$ | 0.991 | 0.985 | 0.989 | 0.986 |

This table reports estimation results for the difference-in-differences regression defined in Equation (2.7), with the total monthly carbon emissions as the dependent variable. Columns (1) and (2) (Columns (3) and (4)) present the results for a two-year (four-year) sample, centered around the 2016 U.S. Presidential Election (November 2016). $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment group includes firms with a $CRE$ measure greater than the median. $After$ is an indicator variable equal to 1 for all months after the 2016 U.S. presidential election (November 2016), and 0 otherwise. We include the following as control variables: the natural logarithm of a firm's market capitalization ($log(Size)$), its book-to-market ratio ($B/M$)), its book leverage ($Leverage$), and the natural logarithm of the property, plant and equipment expenditures ($log(PPEGT)$. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Columns (1) and (3) present the results for all industries, and Columns (2) and (4) present the results for industries excluding the three industries with the highest average CRE. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table 2.16: Information Attention Difference-in-Differences around Trump Election

| | *EDGAR Attention* | | | | | | *Bloomberg Attention* | |
|---|---|---|---|---|---|---|---|---|
| | log(Rpv) | | log(DRTpv) | | log(LMpv) | | log(Read + 1) | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $Treat_i \times After_t$ | $-0.033^{***}$ | $-0.031^{***}$ | $-0.029^{***}$ | $-0.026^{**}$ | $-0.037^{***}$ | $-0.040^{***}$ | $-0.102^{*}$ | $-0.113^{*}$ |
| | (0.011) | (0.010) | (0.011) | (0.010) | (0.012) | (0.012) | (0.060) | (0.058) |
| Controls | YES | No | YES | No | YES | No | YES | No |
| Industry FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES | YES | YES |
| Observations | 14,977 | 16,971 | 14,977 | 16,971 | 14,977 | 16,971 | 2,013 | 2,111 |
| Adjusted $R^2$ | 0.858 | 0.848 | 0.853 | 0.843 | 0.847 | 0.836 | 0.609 | 0.601 |

This table reports estimation results for the difference-in-differences regression defined in Equation (2.8), with four alternative investor attention measures as the dependent variable: (a) the natural logarithm of Rpv following Ryans (2017) (Columns (1) and (2)), (b) the natural logarithm of DRTpv following Drake et al. (2015) (Columns (3) and (4)), (c) the natural logarithm of LMpv following Loughran and McDonald (2017) (Columns (5) and (6)) and (d) the Bloomberg readership following Ben-Rephael et al. (2017) (Columns (7) and (8)). The latter is restricted to firms in the S&P500. *Treat* is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment group includes firms with a *CRE* measure greater than the median. *After* is an indicator variable equal to 1 after the 2016 U.S. presidential election (November 2016), and 0 before the election. We use a 16-month sample centered around the 2016 election. Controls are the same as in Table 2.6. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

## 2.9 Figures




Figure 2.1: Word Cloud of Climate Regulatory Exposure

This figure presents the word cloud of the climate change vocabulary (CCV) in the SEC filings from 2009 to 2018. Terms are assigned a font size proportional to their frequency in the corpus of documents.




Figure 2.2: Word Cloud of Physical Climate Exposure

This figure presents the word cloud of the physical climate change vocabulary (CCE) in the SEC filings from 2009 to 2018. Terms are assigned a font size proportional to their frequency in the corpus of documents.

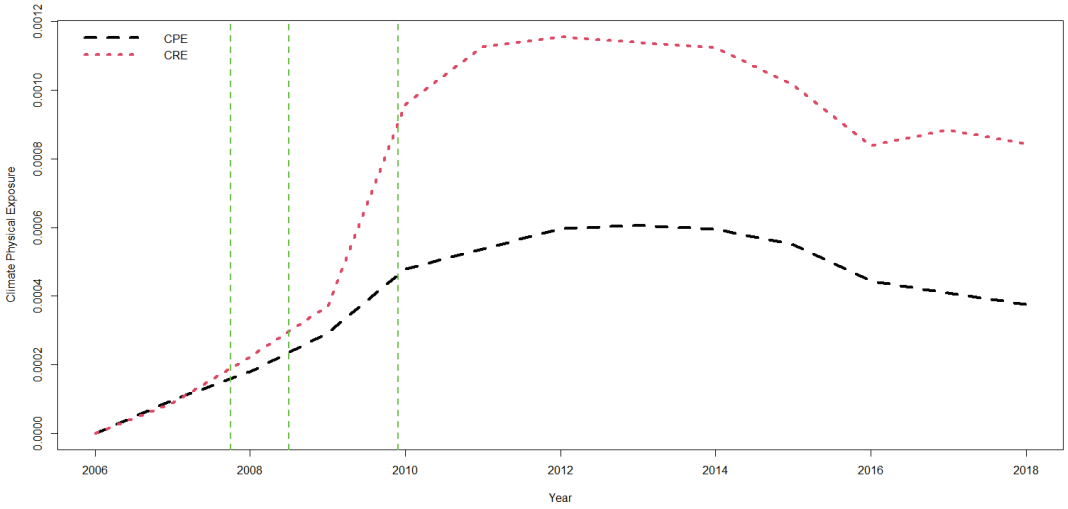



Figure 2.3: Aggregate Climate Regulatory and Physical Exposure

This figure presents the aggregate (average) climate regulatory exposure (CRE) and climate change exposure (CCE) based on SEC filings from 2006 to 2018. The vertical lines represent periods in which investors start to petition the SEC to require firms to provide more disclosures on climate change risks: `https://www.sec.gov/rules/petitions.shtml`.

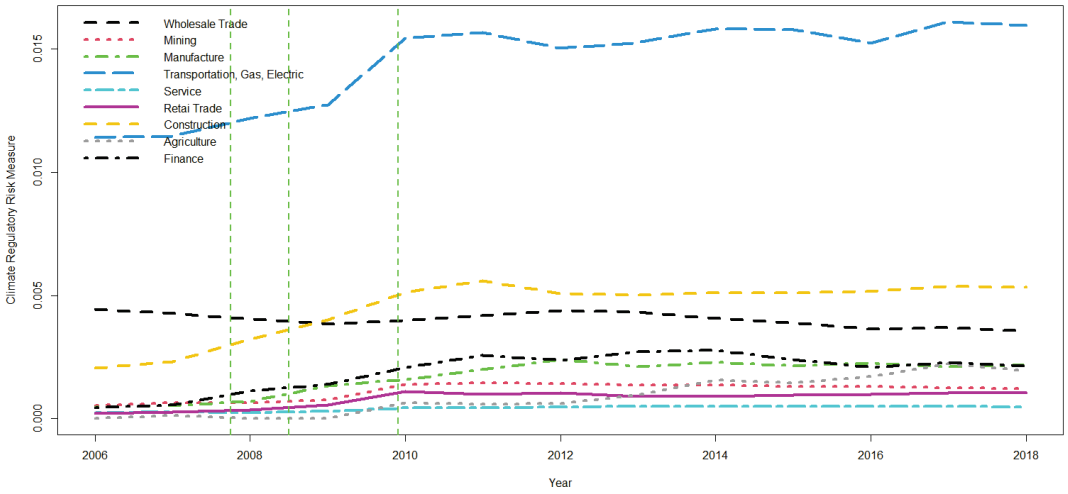



Figure 2.4: Climate Regulatory Exposure (CRE) by Industry

This figure plots the CRE aggregated at the industry level. The industry classification is based on the first 2 digits of the 6-digit Standard Industrial Classification (SIC). The vertical lines represent periods in which investors start to petition the SEC to require firms to provide more disclosures on climate change risks.

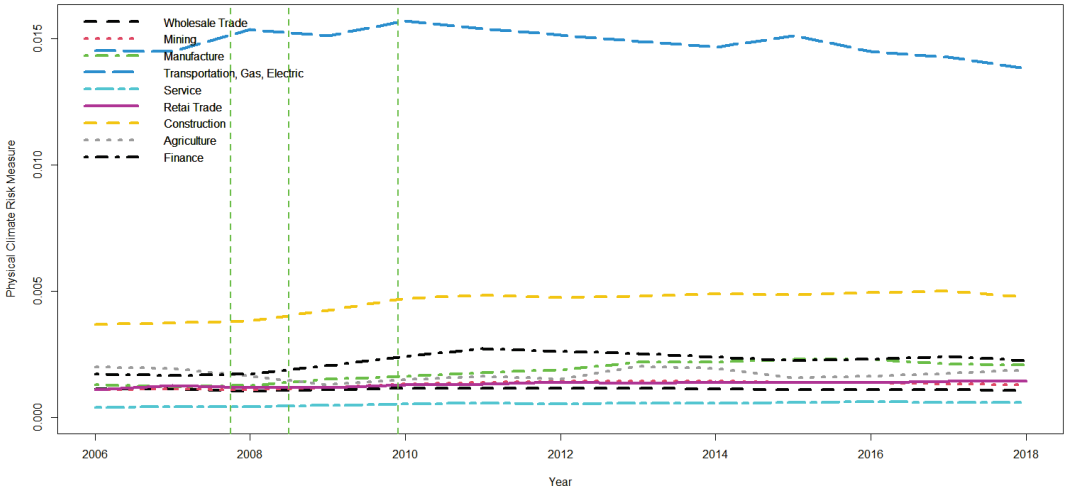



Figure 2.5: Climate Change Exposure (CCE) by Industry

This figure plots the CCE aggregated at the industry level. The industry classification is based on the first 2 digits of the 6-digit Standard Industrial Classification (SIC). The vertical lines represent periods in which investors start to petition the SEC to require firms to provide more disclosures on climate change risks.

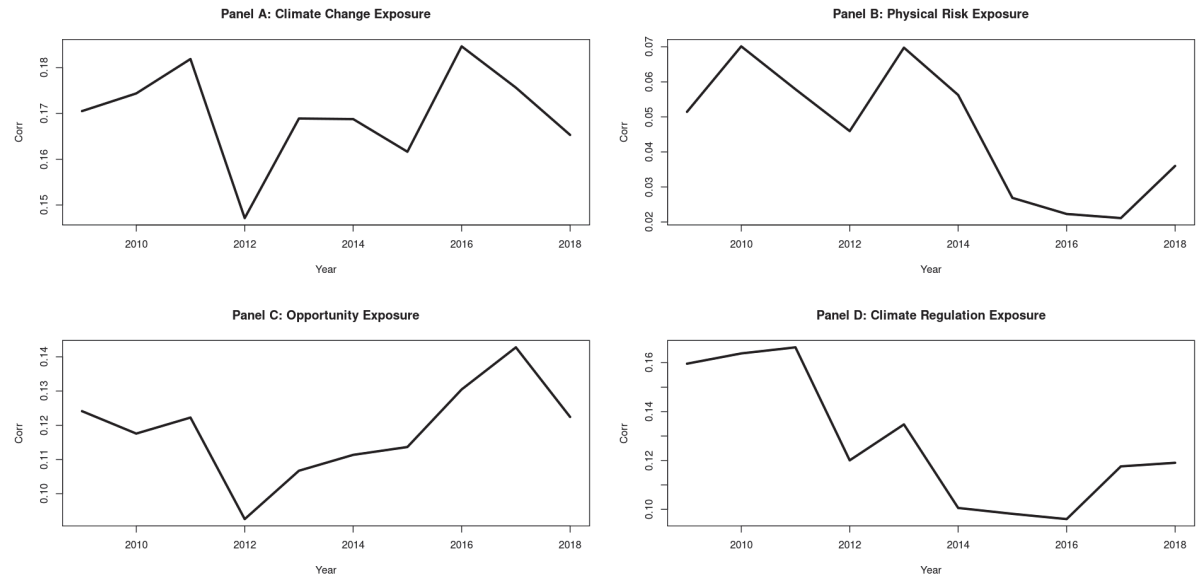



Figure 2.6: Cross Sectional Correlation between CRE and alterntive Climate Measures

This figure plots the cross sectional correlation between SEC-based CRE and benchmarks from Sautner et al. (2020). Panel A shows the annual correlation between CRE and climate change exposure. Panel B shows the dynamic correlation between physical climate risk exposure and CRE, Panel C the correlation between opportunity risk exposure and CRE and Panel D denotes the dynamic correlation between climate regulatory exposure and CRE.

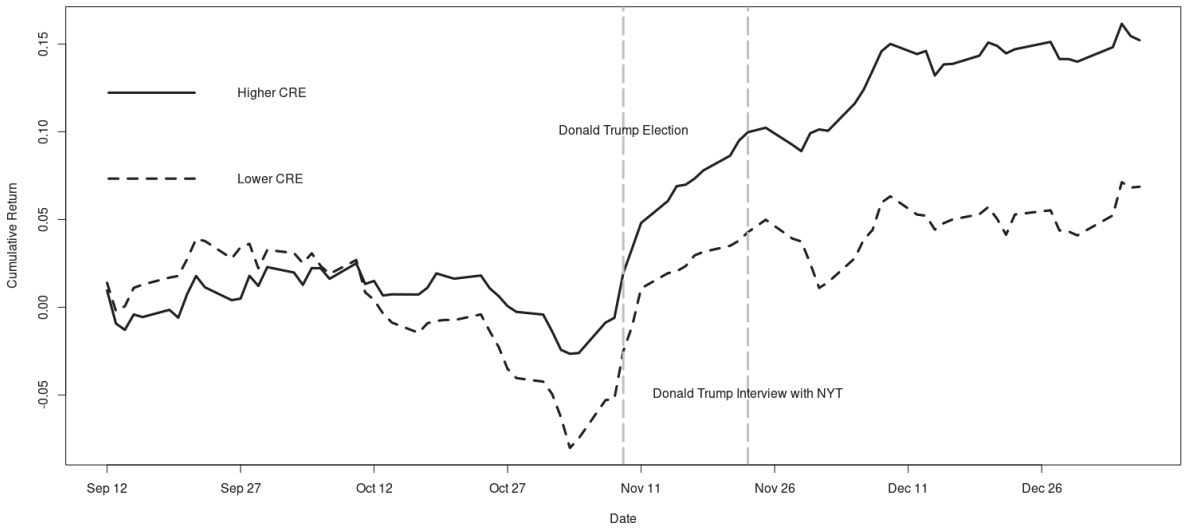



Figure 2.7: Cumulative Raw Returns (October 2016 to January 2017)

This figure plots the cumulative raw returns of holding a high-CRE portfolio (dashed line) and a low-CRE portfolio (dotted line) in the months around the 2016 U.S. presidential election (represented by a grey dashed line). Firms with a $CRE$ measure greater (less) than the median are classified as high-CRE (low-CRE).

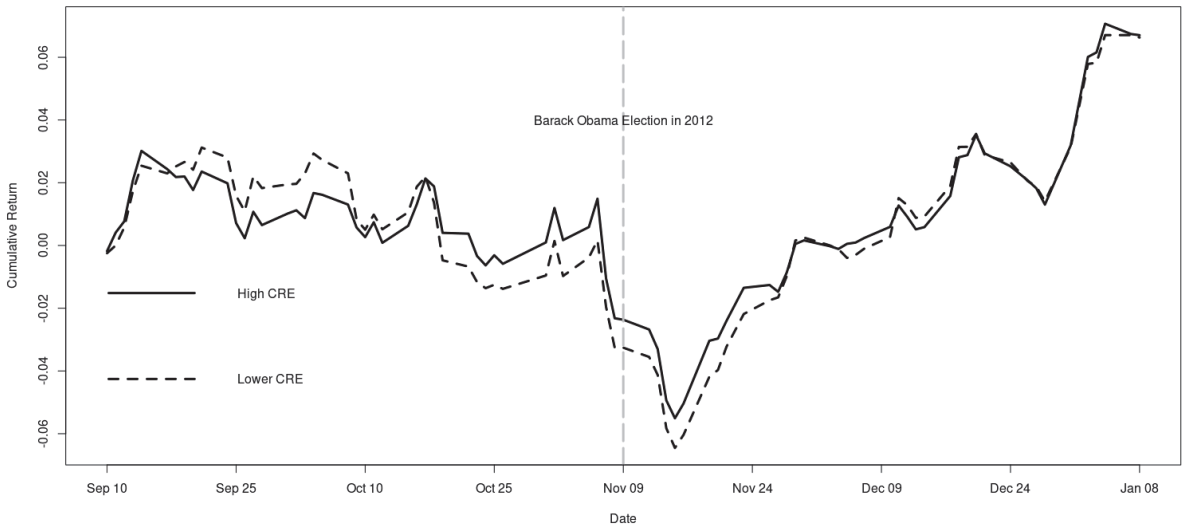



Figure 2.8: Cumulative Raw Returns (October 2012 to January 2013)

This figure plots the cumulative raw returns of holding a high-CRE portfolio (dashed line) and a low-CRE portfolio (dotted line) in the months around the 2012 U.S. presidential election (represented by a grey dashed line). Firms with a $CRE$ measure greater (less) than the median are classified as high-CRE (low-CRE).

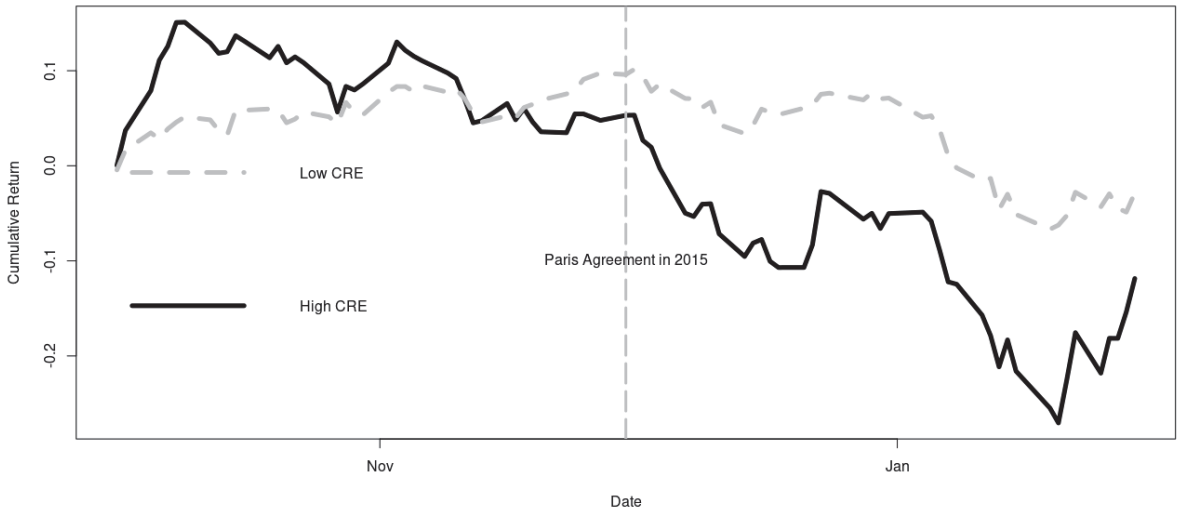



Figure 2.9: Cumulative Raw Returns (October 2015 to January 2016)

This figure plots the cumulative raw returns of holding a high-CRE portfolio (dashed line) and a low-CRE portfolio (dotted line) in the months around the 2015 Paris Agreement (represented by a vertical line). Firms with a $CRE$ measure in the top (bottom) decile are classified as high-CRE (low-CRE).

# Appendix A. Tables for CCV and GPV Dictionary

Table A1: Matched Top 30 GPV and CCV

| | CCV Top 30 | CCV Count | GPV Top 30 | CPV Count |
|---|---|---|---|---|
| 1 | sustainab | 66,935 | politic | 73,249 |
| 2 | ch4 | 55,537 | regulatory | 69,041 |
| 3 | ozone | 48,168 | act | 65,547 |
| 4 | renewable | 41,598 | Legislative | 64,822 |
| 5 | gwp | 38,085 | pederal | 61,560 |
| 6 | ghg | 36,319 | legislation | 57,066 |
| 7 | ods | 34,127 | accord | 56,450 |
| 8 | hfcs | 33,889 | reimbursement | 51,236 |
| 9 | co2 | 31,857 | policy | 50,351 |
| 10 | nox | 31,491 | subsidy | 48,770 |
| 11 | pfcs | 31,379 | congressional | 48,288 |
| 12 | ets | 30,750 | agreement | 47,109 |
| 13 | lcfs | 29,830 | government | 46,692 |
| 14 | methane | 29,677 | government approved | 46,668 |
| 15 | oils | 29,327 | legislatory | 46,192 |
| 16 | n2o | 28,051 | law | 45,052 |
| 17 | euets | 26,431 | regulation | 45,035 |
| 18 | green | 26,280 | legal | 43,097 |
| 19 | sf6 | 22,971 | government sponsored | 42,878 |
| 20 | temperatures | 21,939 | congress | 41,679 |
| 21 | nf3 | 21,558 | tax | 41,542 |
| 22 | climate | 21,255 | political | 35,461 |
| 23 | extinctions | 19,332 | governmental | 31,266 |
| 24 | rps | 18,348 | kyoto protocol | 2,083 |
| 25 | reserves | 16,424 | debt ceiling | 1,979 |
| 26 | emissions | 15,406 | price constraint | 1,800 |
| 27 | emit | 14,077 | government approval | 1,796 |
| 28 | low carbon fuel standards | 13,816 | governmental regulation | 1,751 |
| 29 | corporate average fuel economy | 13,720 | government deficit | 1,719 |
| 30 | nitrogen trifluoride | 12,845 | regulatory delay | 1,597 |

We rely on two dictionaries: CCV (climate change vocabulary) to capture climate-related terms, and GPV (government policy vocabulary) to capture firm-level risks related to government regulatory policy. Further details can be found in Section 3 (specifically Equations (1) and (2)).

Table A2: Matched Climate Change Exposure (CCE)

| Names | Freq |
|---|---|
| oil | 652776 |
| natural.gas | 541355 |
| reserves | 437840 |
| emissions | 156907 |
| weather | 127381 |
| natural.disasters | 69291 |
| renewable | 59635 |
| hurricanes | 42940 |
| earthquakes | 42489 |
| temperatures | 42157 |
| climate | 41586 |
| green | 30774 |
| ghg | 26906 |
| renewable.energy | 25302 |
| climate.change | 21375 |
| pollutants | 19509 |
| catastrophic.events | 19024 |
| ghg.emissions | 18956 |
| mitigations | 18371 |
| carbon.dioxides | 16730 |
| agriculture | 16006 |
| greenhouse.gases | 11091 |
| infectious.diseases | 9855 |
| atmosphere | 9835 |
| fossil.fuels | 9725 |
| droughts | 9589 |
| air.quality | 8594 |
| ozone | 8247 |
| extreme.weather | 7972 |
| flooding | 7265 |

We rely on two dictionaries: CCV (climate change vocabulary) to capture climate-related terms, and use the GPV (government policy vocabulary) to delete instances where government regulatory policy is also discussed. Further details can be found in Section 3 (specifically Equation (3)).

Table A3: CCE Summary Statistics by Year

Panel A: CCE statistics from 2006 to 2018

| Year | Mean | St. Dev. | 25% | Median | 75% |
|---|---|---|---|---|---|
| 2006 | 18.533 | 42.091 | 1.446 | 4.820 | 13.812 |
| 2007 | 19.844 | 43.618 | 1.581 | 5.208 | 14.471 |
| 2008 | 21.208 | 46.361 | 1.619 | 5.225 | 14.654 |
| 2009 | 22.519 | 47.484 | 1.958 | 5.768 | 16.308 |
| 2010 | 24.762 | 50.179 | 2.204 | 6.387 | 18.379 |
| 2011 | 25.525 | 49.759 | 2.507 | 7.273 | 20.272 |
| 2012 | 26.305 | 49.970 | 2.683 | 7.589 | 20.969 |
| 2013 | 26.506 | 49.137 | 2.949 | 7.766 | 21.380 |
| 2014 | 26.453 | 48.604 | 2.866 | 7.733 | 22.150 |
| 2015 | 25.944 | 48.683 | 2.689 | 7.303 | 21.384 |
| 2016 | 24.555 | 45.922 | 2.556 | 7.208 | 20.732 |
| 2017 | 24.070 | 45.125 | 2.469 | 6.924 | 20.077 |
| 2018 | 23.666 | 44.384 | 2.582 | 7.013 | 19.500 |

Panel B: CCE Statistics from 2009 to 2018

| | Mean | St. Dev. | 25% | Median | 75% |
|---|---|---|---|---|---|
| All Industries | 25.009 | 47.993 | 2.524 | 7.050 | 25.009 |
| All Ind. Ex. Top 3 | 12.424 | 19.487 | 2.240 | 5.867 | 12.424 |
| Top 3 | 136.438 | 74.264 | 80.299 | 119.589 | 136.438 |

This table reports summary statistics for Climate Change Exposure ($CCE$) for every year in our sample. Specifically, the table reports the mean, standard deviation, as well as the bottom quartile, median and top quartile. Units are in basis points; 1 word in a document with word size equal to 10,000 is 1 basis point.

Table A4: CCE Summary Statistics by Industry from 2006 to 2018

| Industry No | Industry Name | Mean | S.D. | Firm Count |
|---|---|---|---|---|
| 1 | Agriculture | 19.13 | 13.85 | 19 |
| 2 | Food Products | 20.93 | 18.86 | 90 |
| 3 | Candy Soda | 14.88 | 4.81 | 7 |
| 4 | Beer Liquor | 21.19 | 29.08 | 20 |
| 5 | Tobacco Products | 5.35 | 3.91 | 10 |
| 6 | Recreation | 5.44 | 5.24 | 35 |
| 7 | Entertainment | 8.28 | 8.69 | 78 |
| 8 | Printing and Publishing | 2.93 | 3.35 | 51 |
| 9 | Consumer Goods | 8.69 | 10.62 | 77 |
| 10 | Apparel | 7.78 | 7.50 | 68 |
| 11 | Healthcare | 4.01 | 5.31 | 128 |
| 12 | Medical Equipment | 4.05 | 4.91 | 275 |
| 13 | Pharmaceutical Products | 3.94 | 8.04 | 696 |
| 14 | Chemicals | 33.63 | 29.33 | 143 |
| 15 | Rubber and Plastic Products | 12.79 | 12.39 | 40 |
| 16 | Textiles | 10.71 | 11.24 | 13 |
| 17 | Construction Materials | 21.04 | 22.26 | 90 |
| 18 | Construction | 22.43 | 17.65 | 75 |
| 19 | Steel Works Etc | 19.83 | 16.94 | 82 |
| 20 | Fabricated Products | 24.11 | 21.06 | 15 |
| 21 | Machinery | 26.89 | 33.25 | 164 |
| 22 | Electrical Equipment | 17.54 | 17.63 | 76 |
| 23 | Automobiles and Trucks | 15.72 | 16.36 | 87 |
| 24 | Aircraft | 8.31 | 8.67 | 27 |
| 25 | Shipbuilding, Railroad Equipment | 9.61 | 8.04 | 13 |
| 26 | Defense | 6.70 | 6.55 | 14 |
| 27 | Precious Metals | 37.24 | 19.02 | 21 |
| 28 | Non-Metallic and Industrial Metal Mining | 57.61 | 48.50 | 36 |
| 29 | Coal | 89.56 | 28.56 | 26 |
| 30 | Petroleum and Natural Gas | 172.45 | 75.02 | 308 |
| 31 | Utilities | 89.29 | 47.36 | 201 |
| 32 | Communication | 4.28 | 6.20 | 196 |
| 33 | Personal Services | 6.17 | 7.18 | 79 |
| 34 | Business Services | 6.04 | 12.36 | 531 |
| 35 | Computers | 4.98 | 4.15 | 121 |
| 36 | Computer Software | 3.08 | 3.99 | 465 |
| 37 | Electronic Equipment | 6.94 | 7.56 | 409 |
| 38 | Measuring and Control Equipment | 8.65 | 10.44 | 131 |
| 39 | Business Supplies | 15.12 | 13.23 | 43 |
| 40 | Shipping Containers | 18.39 | 14.99 | 17 |
| 41 | Transportation | 30.05 | 30.20 | 166 |
| 42 | Wholesale | 17.98 | 29.04 | 194 |
| 43 | Retail | 13.53 | 18.61 | 294 |
| 44 | Restaurants, Hotels, Motels | 12.44 | 11.93 | 140 |
| 49 | Almost Nothing | 37.42 | 31.76 | 25 |

This table reports summary statistics for the Climate Change Exposure ($CCE$) by industry. The industry classification follows the Fama-French website based on 6-digit SIC codes. For each industry, the table presents the mean and standard deviation of $CCE$ as well as the number of firms per the CRSP database classified in that industry. The sample spans January 2006 to December 2018.

# Appendix B. Example of SEC Filings on Climate Issues

In this section, we list some examples from 10-K filings to illustrate how the CCV and GPV dictionaries capture concerns related to climate regulations. Terms related to climate (CCV) are colored in blue, while terms pertaining to regulations (GPV) are colored in red.

**Example 1: Gulpfort Energy Corporation (2017)**

*Our operations are subject to various governmental laws and regulations which require compliance that can be burdensome and expensive and could expose us to significant liabilities [...] From time to time, regulatory agencies have imposed price controls and limitations on production by restricting the rate of flow of oil and natural gas wells below actual production capacity to conserve supplies of oil and gas. In addition, the production, handling, storage, transportation, remediation, emission and disposal of oil and natural gas, by-products thereof and other substances and materials produced or used in connection with oil and natural gas operations are subject to regulation under federal, state and local laws and regulations, including those relating to protection of human health and the environment [...] Federal and state legislative and regulatory initiatives relating to hydraulic fracturing could result in increased costs and additional operating restrictions or delays [...] The process, which involves the injection of water, sand and chemicals under pressure into formations to fracture the surrounding rock and stimulate production, is typically regulated by state oil and natural gas commissions.*

**Example 2: Atlas Energy L.P. (2015)**

*Climate change laws and regulations restricting emissions of "greenhouse gases" could result in increased operating costs and reduced demand for the natural gas, while potential physical effects of climate change could disrupt our operations and cause us to incur significant costs in preparing for or responding to those effects. In response to findings that emissions of carbon dioxide, methane and other greenhouse gases, or GHGs, present an endangerment to public health and the environment, the EPA has adopted regulations under existing provisions of the federal Clean Air Act that, among other things, establish Prevention of Significant Deterioration, or PSD, construction and Title V operating permit reviews for certain large stationary sources that are potential major sources of GHG emissions. Facilities required to obtain PSD permits because of their potential criteria pollutant emissions will be required to comply with "best available control technology" standards for GHGs. These regulations could adversely affect our operations and restrict or delay our ability to obtain air permits for new or modified sources.*

**Example 3: Greenbrier Companies Inc (2016)**

*Business, regulatory, and legal developments regarding climate change may affect the demand for our products or the ability of our critical suppliers to meet our needs. We have followed the current debate over climate change in general, and the related science, policy discussion, and prospective legislation. Some scientific studies have suggested that emissions of certain gases, commonly referred to as greenhouse gases (GHGs) and including carbon dioxide and methane, may be contributing to warming of the Earth s atmosphere and other climate changes. Additionally, the potential challenges and opportunities for our company that climate change policy and legislation may pose are reviewed. However, any such challenges or opportunities are heavily dependent on the nature and degree of climate change legislation and the extent to which it applies to our industries. Table of Contents In response to an emerging scientific and political consensus, legislation and new rules to regulate emission of GHGs has been introduced in numerous state legislatures, the U.S. Congress, and by the EPA. Some of these proposals would require industries to meet stringent new standards that may require substantial reductions in carbon emissions. While we cannot assess the direct impact of these or other potential regulations, we recognize that new climate change protocols could affect the demand for our products and/or affect the price of materials, input factors and manufactured components which could impact our margins. Potential opportunities could include greater demand for certain types of railcars, while potential challenges could include decreased demand for certain types of railcars or other products and higher energy costs. Other adverse consequences of climate change could include an increased frequency of severe weather events and rising sea levels that could affect operations at our manufacturing facilities, the price of insuring company assets, or other unforeseen disruptions of our operations, systems, property or equipment. Ultimately, when or if these impacts may occur cannot be assessed until scientific analysis and legislative policy are more developed and specific legislative proposals begin to take shape. Repercussions from terrorist activities or armed conflict could harm our business.*

**Example 4: Phx Minerals Inc (2016)**

*Climate Change Certain studies have suggested that emission of certain gases, commonly referred to as "greenhouse gases," may be impacting the earth's climate. Methane, the primary component of natural gas, and carbon dioxide, a by-product of burning oil and natural gas, are examples of greenhouse gases. Various state governments and regional organizations are considering enacting new legislation and promulgating new regulations governing or restricting the emission of greenhouse gases from stationary sources such as oil and gas production equipment and operations. At the federal level, the EPA has already made findings and issued regulations that require operators to establish and report an inventory of greenhouse gas emissions . Legislative and regulatory proposals for restricting greenhouse gas emissions or otherwise addressing climate change could require us to incur additional operating costs and could adversely affect*

*demand for the sale of oil and natural gas. The potential increase in our operating costs could include new or increased costs to obtain permits, operate and maintain equipment and facilities, install new emission controls on equipment and facilities, acquire allowances to authorize greenhouse gas emissions and pay taxes related to greenhouse gas emissions. Even without federal legislation or regulation of greenhouse gas emissions, states may pursue the issue either directly or indirectly . In addition, the United States was involved in the United Nations Conference on Climate Change in Paris, which led to the creation of the Paris Agreement. The Paris Agreement will require countries to review and represent a progression in their (14) intended nationally determined contributions, which set emissions reduction goals, every five years, beginning in 2020. If adopted, the Paris Agreement could further drive regulation in the United States. Restrictions on emissions of methane or carbon dioxide that may be imposed in various states , or at a federal level, could adversely affect the oil and gas industry.*

# Appendix C. Details of the Pre-Processing Algorithm for 10-K Filings

We collect all 10-K filings from the EDGAR database and parse them using the open-source code provided by Bill McDonald.[29] For every 10-K document, we only keep the information that precedes item 1B (Unresolved Staff Comments). Specifically, we focus on Item 1 (Business) and Item 1A (Risk Factors). This information includes management's expectations of future economic conditions and risk factors likely to impact the company. To guarantee the accuracy of this cutoff point, we use the following expression to capture the title of Item 1B:

```
re.compile(r'(\s.|.)\s*(Item|Items|\s*)[:.\s-]*(1|1B|IB|\s*)[:.\s-]\s
↪  *U\s*n\s*r\s*e\s*s\s*o\s*l\s*v\s*e\s*d\s*(S\s*t\s*a\s*f\s*f|\s*)\
↪  s*C\s*o\s*m\s*m\s*e\s*n\s*t\s*s',
↪  flags=re.IGNORECASE)
```

Then, at the stage of pre-processing the document for term matching, we adjust the CCV and GPV with regular expressions to ensure that there are no problems arising from using capitalization, and/or plural form. The example for the climate change vocabularies from the CCV is:

```
\bcarbon asset risks{0,1}\b,
\b(carbon intensity|carbon intensities)\b
```

The example above shows that both the singular and plural forms of "carbon asset risk" are considered, and this is also the case for "carbon intensity". This approach minimizes the number of climate-related terms that our bag-of-words method may miss.

[29] The code is available from `https://sraf.nd.edu/data/stage-one-10-x-parse-data/`.

# Appendix D. Merging the Data

The SEC assigns a unique identifier, the Central Index Key (CIK), to identify a corporation's or an individual's filing with the SEC. The CIK is the common identifier that we use to merge the 10-K files with data from CRSP and COMPUSTAT.

To do so, we follow the merging process in Fama and French (1992) implementing the firm identifiers, permno and permco, from the CCM link table on CRSP from 2006 to 2019. We include all U.S. public firms (including firms listed on AMEX, NASDAQ and NYSE) and exclude the firms in the financial services industry.

For the data of Bloomberg readership and MSCI ESG data, first, we follow Ben-Rephael et al. (2017) and download the data from Bloomberg. Then, we merge the data based on the joint identifiers of both year and ticker.

# Appendix E. Robustness Tables

Table A5: Difference-in-Differences (Cumulative Abnormal Returns) at Daily Frequency Around Trump Interview, Nov 23rd, 2016

| | *Dependent variable: Cumulative (Abnormal) Returns* | | | | | |
|---|---|---|---|---|---|---|
| | 15-Day | | | 30-Day | | |
| | $CRR$ | $CAR_{CAPM}$ | $CAR_{FF3}$ | $CRR$ | $CAR_{CAPM}$ | $CAR_{FF3}$ |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | 0.037*** | 0.042*** | 0.038*** | 0.044*** | 0.047*** | 0.066*** |
| | (0.007) | (0.006) | (0.007) | (0.009) | (0.009) | (0.010) |
| Observations | 11,002 | 11,002 | 11,002 | 11,016 | 11,016 | 11,016 |
| Controls | YES | YES | YES | YES | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Adjusted $R^2$ | 0.756 | 0.748 | 0.761 | 0.653 | 0.655 | 0.679 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.4) in the main paper, with the cumulative daily stock returns as the dependent variable. $CRR$ refers to the cumulative raw return; $CAR$ refers to the cumulative abnormal returns. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $80^{\text{th}}$ ($20^{\text{th}}$). $After$ in Panel A (Panel B) is an indicator variable equal to 1 for the 15 (30) days after the New York Times interview with Donald Trump (November 23rd 2016), and 0 for the 15 (30) days before the interview. Controls are the same as in Table 2.6 in the main paper. Industry fixed effects (FE) follow the Fama-French defined 49 industries, including financial services. Standard errors are robust and are clustered at the firm level. *$p$ <0.1; **$p$ <0.05; ***$p$ <0.01.

Table A6: Difference-in-Difference around Paris Agreement (CAPM Cumulative Abnormal Return)

Panel A: Dependent Variable: 2-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{2M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | $-0.079^{***}$ | $-0.091^{***}$ | $-0.091^{***}$ | $-0.021$ | $-0.036^{*}$ | $-0.036^{*}$ |
| | (0.024) | (0.023) | (0.023) | (0.021) | (0.020) | (0.020) |
| $CCE_{i,t}$ | | | 35.209 | | | 63.997 |
| | | | (34.509) | | | (40.836) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,317 | 5,669 | 5,669 | 8,939 | 8,058 | 8,058 |
| Adjusted $R^2$ | 0.771 | 0.384 | 0.384 | 0.697 | 0.339 | 0.341 |

Panel B: Dependent Variable: 3-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{3M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | $-0.058^{**}$ | $-0.065^{***}$ | $-0.065^{***}$ | 0.001 | $-0.010$ | $-0.011$ |
| | (0.026) | (0.024) | (0.024) | (0.022) | (0.021) | (0.021) |
| $CCE_{i,t}$ | | | 36.449 | | | 69.199 |
| | | | (35.906) | | | (45.496) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,334 | 5,682 | 5,682 | 8,964 | 8,078 | 8,078 |
| Adjusted $R^2$ | 0.774 | 0.523 | 0.524 | 0.704 | 0.471 | 0.473 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.6) in the main paper, with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $90^{\text{th}}$ ($10^{\text{th}}$) and $80^{\text{th}}$ ($20^{\text{th}}$) percentiles. $After^{Paris}$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2015 Paris Agreement (December 2015), and 0 for the two (three) months before the Paris Agreement. Controls are the same as in Table 2.6 in the main paper. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^{*}p$ <0.1; $^{**}p$ <0.05; $^{***}p$ <0.01.

Table A7: Difference-in-Difference around Paris Agreement (CAPM Cumulative Abnormal Return)

Panel A: Dependent Variable: 2-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{2M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{it} \times After_t$ | $-0.079^{***}$ | $-0.091^{***}$ | $-0.091^{***}$ | $-0.021$ | $-0.036^{*}$ | $-0.036^{*}$ |
| | (0.024) | (0.023) | (0.023) | (0.021) | (0.020) | (0.020) |
| $CCE_{i,t}$ | | | 35.209 | | | 63.997 |
| | | | (34.509) | | | (40.836) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,317 | 5,669 | 5,669 | 8,939 | 8,058 | 8,058 |
| Adjusted $R^2$ | 0.771 | 0.384 | 0.384 | 0.697 | 0.339 | 0.341 |

Panel B: Dependent Variable: 3-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{3M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | $-0.058^{**}$ | $-0.065^{***}$ | $-0.065^{***}$ | 0.001 | $-0.010$ | $-0.011$ |
| | (0.026) | (0.024) | (0.024) | (0.022) | (0.021) | (0.021) |
| $CCE_{i,t}$ | | | 36.449 | | | 69.199 |
| | | | (35.906) | | | (45.496) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,334 | 5,682 | 5,682 | 8,964 | 8,078 | 8,078 |
| Adjusted $R^2$ | 0.774 | 0.523 | 0.524 | 0.704 | 0.471 | 0.473 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.6) in the main paper, with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $90^{\text{th}}$ ($10^{\text{th}}$) and $80^{\text{th}}$ ($20^{\text{th}}$) percentiles. $After^{Paris}$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2015 Paris Agreement (December 2015), and 0 for the two (three) months before the Paris Agreement. Controls are the same as in Table 2.6 in the main paper. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^{*}p$ <0.1; $^{**}p$ <0.05; $^{***}p$ <0.01.

Table A8: Difference-in-Difference around Paris Agreement (FF3 Cumulative Abnormal Return)

Panel A: Dependent Variable: 2-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{2M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | $-0.081^{***}$ | $-0.088^{***}$ | $-0.088^{***}$ | $-0.032^{*}$ | $-0.045^{***}$ | $-0.045^{***}$ |
| | (0.023) | (0.021) | (0.021) | (0.019) | (0.017) | (0.017) |
| $CCE_{i,t}$ | | | 27.768 | | | 58.322 |
| | | | (29.010) | | | (40.104) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,233 | 5,585 | 5,585 | 8,811 | 7,943 | 7,943 |
| Adjusted $R^2$ | 0.800 | 0.465 | 0.466 | 0.738 | 0.422 | 0.424 |

Panel B: Dependent Variable: 3-Month Cumulative Abnormal Returns

| | *Dependent variable:* $CAR^{3M}_{i,t}$ | | | | | |
|---|---|---|---|---|---|---|
| | 90 % | | | 80 % | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $Treat_{i,t} \times After_t$ | $-0.053^{**}$ | $-0.055^{**}$ | $-0.055^{**}$ | $-0.0003$ | $-0.010$ | $-0.010$ |
| | (0.025) | (0.024) | (0.024) | (0.021) | (0.019) | (0.019) |
| $CCE_{i,t}$ | | | 21.220 | | | 57.679 |
| | | | (32.032) | | | (45.591) |
| Control | NO | YES | YES | NO | YES | YES |
| Industry FE | YES | YES | YES | YES | YES | YES |
| Firm FE | YES | YES | YES | YES | YES | YES |
| Year FE | YES | YES | YES | YES | YES | YES |
| Observations | 6,334 | 5,682 | 5,682 | 8,964 | 8,078 | 8,078 |
| Adjusted $R^2$ | 0.774 | 0.523 | 0.524 | 0.704 | 0.471 | 0.473 |

Note: This table reports estimation results for the difference-in-differences regression defined in Equation (2.6) in the main paper, with the cumulative monthly stock returns as the dependent variable. $Treat$ is an indicator variable equal to 1 for firms in the treatment group, and 0 otherwise. The treatment (control) group includes firms with a $CRE$ measure greater (smaller) than the $90^{\text{th}}$ ($10^{\text{th}}$) and $80^{\text{th}}$ ($20^{\text{th}}$) percentiles. $After^{Paris}$ in Panel A (Panel B) is an indicator variable equal to 1 for the two (three) months after the 2015 Paris Agreement (December 2015), and 0 for the two (three) months before the Paris Agreement. Controls are the same as in Table 2.6 in the main paper. Industry fixed effects (FE) follow the Fama-French defined 49 industries, excluding financial services. Standard errors are robust and are clustered at the firm level. $^{*}p$ <0.1; $^{**}p$ <0.05; $^{***}p$ <0.01.

# 3. Pricing Climate Risk: A Machine Learning Approach

Zhenhao Tan and Yi Zhang

## 3.1 Introduction

With the advent of the digital age, the amount of stored digital data has increased significantly, especially in unstructured data. These data have no predefined format (e.g., textual data), making them more difficult to process and analyze; however, they contain much qualitative information. There is a huge amount of literature recently contributing to this field. For instance, Florackis et al. (2023) documents the ex-ante cyber-security risk in SEC 10K filings and shows how the risk could impact firm performance. Liu and Matthies (2022) tracks the consumption risk in news media and shows the consistency between long-run consumption risk and long-run risk asset pricing theory. Engle et al. (2020b) uses social media data to show that aggregate climate risk exists and prices the cross-sectional equities over the market. All these papers are trying to construct text-implied risk from textual analysis.

However, conducting text data for empirical asset pricing tests takes much work. On the one hand, unstructured data is notoriously famous for its low signal-to-noise ratio (Ke et al. (2019)). There is a huge amount of uninformative data within unstructured data. On the other hand, there are alternative ways to construct text-implied risk exposure. Finding the most potential methodology to generate reliable exposure takes time and effort. For instance, researchers may have different definitions for IDF (inverse document frequency), which could generate different measures even for the same text topic.

This paper introduces a machine learning framework incorporating natural language processing (NLP) to resolve the problem with two-stage regression. Like Fama-Macbeth regression, the first econometric framework assumes a linear structure between the stock returns and the odds of associated climate vocabulary occurrence. The count of the climate vocabularies in a given 10-K follows a multinomial distribution. This multinomial distribution generates realized sufficient cross-sectional score as Taddy (2013) documents. The second stage regression assumes a

linear relationship between expected return and premium component given the sufficient estimator as firm exposure. The structure of the econometric model is similar to Fama-Macbeth, but we make innovative revisions by replacing the first-stage linear inverse regression with NLP inverse regression.

Given the specified econometric framework, we introduce a two-step analysis for the price of text risk. In the first step, we rely on cosine similarity in King et al. (2017) to filter out the irrelevant noisy terms. The algorithm produces a comprehensive dictionary including both unigrams and bigrams based on domain knowledge benchmark (Chou and Kimbrough (2019b), and Sautner et al. (2020)). In the second step, an essential procedure for asset pricing, we introduce the previous two-stage econometric framework for asset pricing tests. The method's strength is to generate a sufficient estimator for firm topic-specific exposure for cross-sectional stock returns and compress the high dimensionality of the text data into a univariate variable. The estimator's sufficient property can be interpreted as a ceiling possible methodology between undiversified risk on the given topic and stock returns for pricing efficiency. Following Gentzkow et al. (2019), we also show the consistency of the price of the risk estimator and provide a method for econometric inference of the structure.

Concerning the application of the methodology, we implement a case study on climate risk exploration. The reasons for us to focus on this application are the following. First, climate change has been widely discussed and attracted much attention recently due to its immediate impacts (e.g., sea-level rise and global warming). There are lots of anecdotes and events to support the validation of the experimentation. For instance, many agreements have been signed during climate-related events (e.g., G8 Climate Agreement and Paris Agreement) to determine blueprints for stabilizing Earth's climate and improving humans' adaptability to climate change. Second, according to recent literature like Engle et al. (2020b), a long-term climate risk should exist. These climate-related news items may have important impacts on various industries. However, the climate risk at the firm level is generally hard to capture since there has yet to be a federal-level policy with measurable indices. Instead of relying on an alternative measure like carbon emission in (Bolton and Kacperczyk (2021a) and Bolton and Kacperczyk (2021c)) or third-party rating (Pastor et al. (2021a)), it is not easy to quantify how sensitive a firm is to such news, even though soft information disclosure is one reliable data source. Therefore, this paper tests whether the qualitative textual data in 10-Ks contain useful climate information in pricing cross-sectional stock returns.

Based on the econometric model, our findings in this paper are consistent and externally valid with recent literature. The firm-level climate risk index, CRS, shows internal validity at the aggregate industry level. First, industries like utilities, oil, and transportation with high carbon exposure in Bolton and Kacperczyk (2021a) have higher CRS. Second, the CRS shows sensible t-SNE statistics under lower dimensional visualization especially focusing on renewable, agriculture, emission, and weather issues. In addition, the word cloud correctly shows that the most relevant vocabularies refer to greenhouse gas, extreme weather, and climate regulatory

concern. In addition to validation over the full sample period, we provide evidence for the external validation by checking the rationality around key events causally. Following Bolton and Kacperczyk (2021a), we conduct a Diff-in-Diff causal inference around Paris Agreement. The empirical evidence shows that firms with high CRS tend to have higher expected returns after the regulatory event. Moreover, a balanced test on sale, leverage and sale growth rejects the alternative explanation for the finding.

Given the CRS, our paper explores the potential asset pricing test for climate risk in 10-Ks. Firms are sorted based on CRS and are split into five quintile portfolios. Our *CRF*, factor-mimicking portfolio based on CRS, is constructed by longing the quintile five portfolio and shorting the one portfolio. We considered both equal- and value-weighted portfolios. Since the text-implied climate information is preserved in the CRS, the excess returns of *CRF* can be regarded as the premium of taking the climate risks implied in the textual data. The *CRF* is regressed with the capital asset pricing model (CAPM) (Sharpe, 1964), the Fama-French 3-factor model (Fama and French, 1993), the Carhart 4-factor model (Carhart, 1997), and the Fama-French 5-factor model (Fama and French, 2015) to test whether existing factor models can explain the abnormal returns of the portfolio. The result reports that an abnormal return from 1.2% to 2.1%, depending on the specification of the benchmark model, can be generated through the long-short portfolio. Bivariate portfolio sorts are also conducted. We utilize the same sort as the univariate portfolio-level analysis. Then we independently sort stocks into ascending order according to some firm-level characteristics. The results show that firms with high CRS outperformance persist, ensuring we are not capturing proxies of other well-known risks. We also examine the cross-sectional relation between CRS and stock returns by running stock-level Fama and MacBeth (1973) regressions. A positive and significant relationship exists between climate risk and stock returns, implying a general risk compensation story for holding climate risk-intensive stocks. The CRS predicts cross-sectional variation in stock returns up to 6 months into the future, showing that predictability is a relatively long-term phenomenon. In addition, we provide evidence to verify that our projected measure is sensible at the firm level, including the relationship with some important climate indicators (e.g., carbon intensity in Bolton and Kacperczyk (2021b) and WSJ index in Engle et al. (2020b)).

Our paper is organized as follows. We show our incentive of analyzing climate risk in 3.2. In 3.3, we present our data and demonstrate how we conduct the data collection and data cleaning. 3.4 presents a machine learning framework to extract the text-implied climate risk and how this can be incorporated into classical Fama-French analysis. We provide supporting mathematical proofs in the appendix. 3.5.1 validates the relation between our climate risk index and events related to climate change. 3.5.2 reports an empirical analysis of CRS and stock returns, and 3.5.3 provides various robustness tests. Finally, 3.6 concludes.

## 3.2 Related Literature

In this paper, we use the sufficient dimension-reduction property of multinomial inverse regression to obtain risk scores on the climate-related information provided in firms' 10-K files and test whether the factor-mimicking portfolio based on these scores is priced. Our paper is related to some frontiers in textual analysis for asset pricing or climate finance literature.

### 3.2.1 Textual Analysis

Our paper relates to this literature using textual analysis to retrieve financial information. For example, unlike some macro models dealing with the first moment, measuring the second moment (i.e., uncertainty) of the economic environment (e.g., the uncertainty of Brexit) sometimes needs to be quantitatively accessed. The uncertainty can be constructed from alternative unstructured data sources (Baker et al., 2016b; Hassan et al., 2019b). Some monetary measures (e.g., GDP) can only be released quarterly. Therefore, the unstructured data sources that come from social media can be integrated into the traditional macro-forecasting model, and the variation and content of the media, which vary daily, can be used to generate signals about the macro economies on a more frequent basis (Larsen and Thorsrud, 2019; Mueller and Rauh, 2018). Another application is determining what skills are in demand in certain economies by analyzing job descriptions using text-mining techniques. Hershbein and Kahn (2018) proposed powerful methods for detecting demand changes, which can further implicate unemployment and wage situations. There have been many other recent applications for text mining techniques in areas such as polarization (Gentzkow and Shapiro, 2010; Gentzkow et al., 2019), transparency and monetary policy (Hansen et al., 2018), idea exchange (Iaria et al., 2018) and product market competition (Hoberg and Phillips, 2016).

From methodological perspective, a popular quantitative representation of text for analyzing unstructured information is the document-term matrix, which calculates the frequency of each word in each document. The rows of a document-term matrix represent different documents, and the columns represent terms in the vocabulary. However, the document-term matrix is usually highly dimensional and sparse, meaning it is a large matrix in which most elements are zero. Therefore, analyzing such a large matrix by some deep models (e.g., neural networks (Hirose et al., 1991; Van Ooyen and Nienhuis, 1992; LeCun et al., 2015)) is computationally intensive. Hence, some dimension-reduction techniques may be useful, which will be further discussed.

More importantly, our paper directly links to recent literature that uses textual analysis within the asset pricing framework, but we focus on different perspectives. Bybee et al. (2022) introduces an econometric framework incorporating the aggregate news flow as a state variable. Their method focuses more on modeling the news stream into the pricing framework but needs

to explain the cross-sectional premiums. Garcia et al. (2023) also documents multinomial inverse regression applications in the financial market, but they focus on neither empirical asset pricing framework nor topic-specific risk exposure. Instead, they care more about realized return correlation with market sentiment. (Ke et al., 2019) introduces a mixture multinomial framework for aggregate stock prediction with econometric properties on accuracy of forecasting.

### 3.2.2 Climate Risk

Climate change has been widely discussed in recent decades due to its immediate impacts, such as global warming and sea level rise (Ghoneem, 2016). The United Nations Framework Convention on Climate Change (UNFCCC, 2005) states that the rapid changes in Earth's climate are highly related to the development of human beings. Since the pre-industrial era, human activities have directly or indirectly changed the Earth's climate system on both global and regional scales (IPCC, 2001b). These activities increase the atmospheric concentrations of greenhouse gases (e.g., CO2), contributing to global warming and sea level rise.

Concerning the textual analysis for climate finance, we are one of many to construct the measure. Sautner et al. (2020) implements the earning conference call data and constructs a firm-level climate risk exposure. Compared to their methodology with term frequency, the difference of our paper is to use a sufficient projection (supervised machine learning) to form firm exposure. Given this sufficiency, we can document a strong climate risk factor in asset pricing literature. Li et al. (2020) also uses the earning conference call to recover the firm-level climate with tfidf.[1]; however, they focus on the corporate finance story through Tobin'q and investment. Baz et al. (2022b) also uses SEC 10-K filings to construct climate-related risk exposure, but their framework concentrates on transition risk and identification strategy. Engle et al. (2020b) also introduces a technique to construct factor mimicking portfolio to hedge against the climate news risk. However, their strategy focused on the aggregate news and did not provide empirical evidence for cross-sectional abnormal return findings. Ardia et al. (2022) also uses high frequent news and identifies daily return reactions to climate news shock, interpreting it as investor preference risk. Their finding is not directly linked to abnormal cross-sectional return. However, they show that their measure could explain the abnormal return over Green-minus-Brown portfolio return [2].

From an empirical perspective, our paper also complements the literature that uses alternative measures to show climate risk's impact on stock returns. Bolton and Kacperczyk (2021a) regress stock returns with carbon emission at the firm level and show that high carbon-intensive firms outperform the low carbon ones because of exposure to carbon risk. Bolton and Kacperczyk

[1] tfidf means term frequency and inverse document frequency.

[2] Green-minus-Brown follows the same definition as MSCI ESG rating

(2021c) exploits the international setting and shows that higher climate policy tightness in a country could amplify the carbon premiums between high and low-carbon firms. Even though carbon risk is a highly endogenous variable, it could be regarded as a measure of general climate risk like our CRS. Overall, our finding is consistent with this literature.

## 3.3 Data

We utilise several databases to construct the sample and conduct the analysis. For the textual data, preprocessing is an important step for natural language processing (NLP) tasks, since it transforms text into a more digestible form improving the performances of machine learning algorithms. We follow common steps from the NLP literature to clean the textual data we used in this paper. First, we normalise the text, including 1) converting all words in the documents to lower case letters; 2) removing digitals, punctuations, and special symbols; 3) removing email addresses; 4) deleting non-English words; 5) expanding contractions. Second, lemmatisation is conducted to group words with different forms and analyse them as a single root word (e.g. "energetic" to "energy", "emits" to "emit"). Third, we remove stop words (e.g. "a", "is", "the", "and"), which are commonly used in natural languages but do not add value to the meaning of the document.

**Data on IPCC Reports** The IPCC reports are downloaded from its website and the list of reports we used is attached in Table A1. We apply the keywords discovery algorithms to the IPCC reports since they contain many climate-related unigrams and bigrams. A climate-related vocabulary can be generated from these reports via several machine learning techniques. This vocabulary is used to generate the document-term matrix on the 10-K corpus, measuring the climate exposure of each firm.

**Data on Existing Climate-related Vocabulary** Similar to Chou and Kimbrough (2019b), a comprehensive vocabulary related to climate change is assembled by gathering 158 terms from reports pertaining to climate change and public policy more broadly. During the period from 2005 to 2017, 6 out of 158 terms had counts of zero (i.e., they were not mentioned) in all documents. Therefore, we only consider 152 terms in this paper.

The final vocabulary combines our own-discovered unigrams and bigrams with the vocabulary created by Chou and Kimbrough (2019b). We believe this vocabulary is comprehensive, but we do not claim it is exhaustive. A more exhaustive vocabulary can be used in the future for further exploration. The full vocabulary is attached in 3.9.4.

**Data on Returns and Textual Information of Firms** The monthly stock return data come from The Center for Research in Security Prices (CRSP), and the corporate fundamental data come from Compustat. Two datasets were merged following standard procedures and the keys for the matching are the 6-digit CUSIP and the date. We obtain the 10-K textual documents from the SEC EDGAR database. We then used CIK (i.e., Central Index Key) and filing date as keys to merge the 10-K files with the CRSP. We used the 10-K files from 2005 to 2017.

Each public company is required to provide a comprehensive annual report, namely 10-K, by the SEC. The operational and financial performances of a firm are summarised in this report. The 10-K also includes a risk factors section disclosing some significant risks that apply to the company. However, the summarised information is difficult to measure numerically. Therefore, by analysing these textual data, we can better understand the risks affecting the firms. Some data pre-processing procedures are conducted for the textual documents. Raw risk factor documents extracted from 10-K files were first tokenised and filtered through the dictionary as a sequence of words.

**Data on Returns of Factor-mimicking Portfolios** We downloaded the returns of other factor-mimicking portfolios and the risk-free rates from Kenneth French's data library. While conducting the MNIR, we used the excess returns, which are the returns beyond the associated risk-free rate. Since we obtained the projections based on firms' yearly returns and the returns from CRSP are on a monthly basis, we accumulated these monthly returns. We report the summary statistics in the Panel C of Table A5.

**Data on Carbon Intensity** we use the MSCI ESG data to evaluate our projected measure relationship with the carbon emission. Our data is from 2007 to 2017 at the monthly frequency. The data on carbon intensity from MSCI ESG is widely cited by researchers like Engle et al. (2020b) and the intensity index is calculated with firm-level carbon emission divided by the sales of the firms. Table A5 shows the details of the statistics.

## 3.4 Methodology

Our procedure to analyze the text-implied climate risk can be divided into two procedures. First, we implemented unsupervised machine learning and domain knowledge to define a climate-related vocabulary including both unigrams and bigrams. Secondly, we generate a document-term matrix for climate-related phrases in the 10K filings emphasizing the potential climate

risk variation. Thirdly, we incorporate a supervised linguistic model, MNIR by Taddy (2013) into classical Fama-French analysis. The general procedures are similar to Cong et al. (2019).

### 3.4.1 Feature Extraction

In the first step, we construct a dictionary reflecting the climate-related arguments in the 10K filings. Based on the preliminary dictionary documented in Chou and Kimbrough (2019b), we extend the dictionary via the word embedding in NLP. The intuition is that with likelihood estimation to parametrize the phrases in vectors, the climate-related ones should have high similarities with the benchmark dictionaries. The overall procedure should be divided into two categories: one is to extend the unigrams and the other is to find similar bigrams from the texts.

Our main technique relies on word embedding which has been developed in NLP for years. Word2vec is a group of models used to construct word embeddings. It takes contexts of words into account (considered location information) and uses a neural network model to learn the word association. Compared to representing a word with a one-hot encoded vector, the word embedding method maps words as continuous vectors in a low dimensional space. There are two widely used word2vec models, which are Continuous Bag-of-Words (CBOW) and Continuous Skip-gram (Skip-gram) models. The CBOW model can be regarded as a feedforward neural network model with one hidden layer, where the current word is predicted based on its contexts. The Skip-gram model is a "flipped" CBOW model. It uses each current word as an input and predicts the surrounding words (i.e. words within a certain range before and after the current word). The models can be trained via stochastic gradient descent and backpropagation (detailed derivations: see e.g. Goldberg and Levy, 2014; Rong, 2014). Details on dictionary construction are provided in Appendix 3.8.

**Algorithm 1** The Keyword Algorithm: Unigram

1: Input initial set of bigrams $\mathbb{C}^{B_0}$
2: Define a reference set $\mathbb{R}$ and a search set $\mathbb{S}$ ▷ $\mathbb{R}$ contains sentences including initial bigrams
3: **procedure** PARTITION DOCUMENTS IN $\mathbb{S}$ ▷ into two groups: $\mathbb{T}$ and $\mathbb{S} \setminus \mathbb{T}$

- $\mathbb{T}$ for target set, $\mathbb{S} \setminus \mathbb{T}$ for non-target set
- Define a training set by drawing a random sample from $\mathbb{R}$ and $\mathbb{S}$ (consistent with y train as 1 and 0)
- Fit one or more machine learning classifiers to the training set
- Use the estimated parameters from the classifiers to predict the probability of $\mathbb{R}$ membership for documents in $\mathbb{S}$
- Compute the percentage of appending the document into $\mathbb{T}$ by different models. If >80%, we assign it to $\mathbb{T}$, otherwise into $\mathbb{S} \setminus \mathbb{T}$

**Algorithm 2** The Keyword Algorithm: Bigram

1: **procedure** KEYWORDS DISCOVERY ▷ Find keywords that best classify documents into either $\mathbb{T}$ or $\mathbb{S} \setminus \mathbb{T}$

- Generate a set of bigrams from $\mathbb{S}$ (occur above a chosen frequency threshold e.g. 10), as $B_S$
- Compare the appearance frequency of each bigram $b \in B_S$ in $\mathbb{T}$ and $\mathbb{S} \setminus \mathbb{T}$ respectively (measured by the proportion of documents containing $b$ in $\mathbb{T}$ and the proportion of documents containing $b$ in $\mathbb{S} \setminus \mathbb{T}$)
- Rank bigrams in $\mathbb{T}$ by a modified statistical likelihood measuring how well this bigram discriminates $\mathbb{T}$ from $\mathbb{S} \setminus \mathbb{T}$ (same as the bigrams in $\mathbb{S} \setminus \mathbb{T}$)
- Select the top n% of bigrams in $\mathbb{T}$ as $\mathbb{C}^{Bs}$

2: Output the final set $\mathbb{C}^B = \mathbb{C}^{B_0} \bigcup \mathbb{C}^{Bs}$

We use the t-SNE Van der Maaten and Hinton (2008) to visualise the high-dimensional data (e.g. 200 in our case) by locating a data point in a low-dimensional map (e.g. two- or three-dimensional). Figure 3.1 shows the the results for unigrams in IPCC. As a preliminary test for the quality of Word2vec, the figure shows a clear classification of the potential unique unigrams at the lower dimension. Figure 3.2 displays the IPCC relevant word frequency. Overall, it seems that topics on climate and emission and greenhouse gas are important within their annual reports.

Figure 3.3 shows the word cloud for the relevant phrases in SEC documents with the expanded dictionaries from 2009 to 2012. It seems that the topics on global warming ('emission', 'lcfs', 'co2'), rare disaster ('wildfires', 'extreme weather', 'weather impacts') and pollution ('air quality', 'pollutants' etc) outweigh the others. 3.9.5 displays the results in the other years.

During this procedure, there is a tunning parameter for dictionary construction. How many words do we need to keep for each dictionaries ? To resolve this issue, we calculate the cumulative recall and precision measure for each dictionaries and select the totoal number 100 in Figure 3.4, where both cumulative precision get convergent. Finally, the Table A2 shows the initial word list with expanded dictionary for unigrams. Table A3 and A4 display the initial word list form (Sautner et al., 2020) and expanded dictionary respectively.

### 3.4.2 Econometric Specification

We implement a similar structure as Feng et al. (2020) and Shanken (2015) but revised the structure with an unobserved textual factor. The basic underlying SDF process follows:

$$m_t := \lambda_0^{-1}(1 - \lambda^* L_t^* - \lambda^T v_t) \tag{3.1}$$

$\lambda_0$ denotes the zero-beta risk rate for assets. $L^*$ refers to the latent text implied risk like the climate-related phrases. The opaqueness of the text implied risk contributes to the misspecification of SDF in the classical asset pricing model and then hurts the pricing efficiency of assets. The text-implied climate risk is one of the most concerns in the recent literature. $v_t$ denotes the other observed macroeconomic factors like value, size, momentum, and predictability.

Then, we assume that the asset pricing returns follow the procedure:

$$r_t^i = \gamma_i + \beta_i^* * L_t^* + B_i^T v_t + \epsilon_{i,t} \tag{3.2}$$

Overall, $L_t$ denotes the latent text-implied factor (one dimension) in the macroeconomic. $B_t^i$ refers to the other K dimensional risk factor exposures (B is a $K \times 1$). $\epsilon_{i,t}$ denotes the idiosyncratic risk for each asset.

Based on the econometric settings for asset pricing, our asset pricing test focuses on the cross-sectional test as Shanken (2015). We hold the basic econometric setting as:

$$E_t(R_i) = \lambda_0 + \alpha_{i,t}^* \lambda_t^* + \alpha_{i,t}^T \lambda_t \tag{3.3}$$

Given $E(L_t) = 0, E(v_t) = 0$ (normalized) and $\alpha_i^* = Cov(L_t^*, R_i)$ and $\alpha_i = Cov(v_t, R_i)$ where $\alpha_i$ is a $K \times 1$ vector.

**Assumption 1** $E(\epsilon|v_t) = 0, Var(\epsilon|v_t) = \Sigma$ *with full rank.*

The key difference of our framework compared with Shanken (2015) is that the text-implied macroeconomic risk is unobserved directly. Therefore, different from the classical two-pass Fama-MacBeth to reconstruct the observed factor exposures directly, we need to employ the linguistic model first to retrieve the text-implied exposure.

### 3.4.3 Multinomial Inverse Regression

We implement a MNIR (multinomial inverse regression) by Taddy (2013) to retrieve the text-implied climate exposure. Each document can be represented as $\mathbf{x}_i = [x_{i1}, x_{i2}, \ldots, x_{ip}]'$ under a given tokenisation. $x_{ip}$ is the counts for token $p$ in the vocabulary for document $i$. Since the document may not contain all tokens in the vocabulary, the vector is sparse. The word counts and the term frequencies $\mathbf{f}_i = \mathbf{x}_i / m_i$ are the basic data units for statistical text analysis, where $m_i = \sum_{j=1}^{p} x_{ij}$ is the total number of vocabulary words in document $i$. The MNIR is implemented under the token exchangeable assumption in the Bag-of-Words model and estimated with sparse parametrized assumption to deal with the curse of high dimensionality. Similar to the first pass in Fama-Macbeth to retrieve the independent variables for the second pass, the

model is suitable in capturing risk exposure, since it's feasible to compress the high dimensional textual information into a lower dimensionality.

In MNIR, each document $\mathbf{x}_y^i$ in year $y$ follows a multinomial distribution, which is widely used in the computational linguistic literature due to its natural interpretative ability with textual data (Blei et al., 2003; Mcauliffe and Blei, 2008),

$$\mathbf{x}_y^i \sim \text{Multinomial}(\mathbf{q}_y^i, m_y^i), \text{ with } q_{yj}^i = \frac{\exp(\eta_{yj}^i)}{\sum_{l=1}^{D} \exp(\eta_{yj}^i)}, \ j = 1, \ldots, D, \ \eta_{yj}^i = \phi_j + \gamma_j r_{y+1}^i, \quad (3.4)$$

where $\mathbf{q}_y^i$ contains the probability that each term appears in the document (i.e., a logistic link), and $m_y^i$ is the total counts of vocabulary words in document $i$ in year $y$. The excess return of stock $i$ can be represented as $r_{y+1}^i$ from $y$ to $y+1$, and $\eta_{yj}^i$ represents the log odds of the word $j$, which can be explained by a linear function of $r_{y+1}^i$. The loading of word $j$ in $r_{y+1}^i$ can be denoted as $\gamma_j$, and since the regression only contains one variable, it also represents the correlation between $\eta_{yj}^i$ and $r_{y+1}^i$. If the frequency of word $j$ in document $i$ in year $y$ increases (decreases), meaning the word is more (less) frequently discussed by the annual report of firm $i$, we would expect a positive (negative) correlation and higher (lower) return of stock $i$ during the period. Since a higher expected return implies a higher systematic risk, we can classify the words as high-risk if they have positive probability loadings $\gamma$ and as low-risk if they have negative probability loadings $\gamma$. $\phi$, which is the intercept of the linear relation between $\eta_{yj}^i$ and $r_{y+1}^i$, represents the fixed effect of $\eta_{yj}^i$.

There is a significant and attractive property of MNIR: sufficient dimension reduction. We can restate Proposition 3.1 and Proposition 3.2 in Taddy (2013) as follows:

**Theorem 1** *Under 3.4, given* $p(r_{y+1}^i \,|\, \mathbf{x}_y^i) = p(r_{y+1}^i \,|\, \mathbf{f}_y^i)$, $\mathbf{\Phi}'\mathbf{f}_y^i$ *is a sufficient statistic for the stock return* $r_{y+1}^i$, *meaning* $r_{y+1}^i \perp \mathbf{x}_y^i \,|\, \mathbf{\Phi}'\mathbf{f}_y^i$.

In proposition 1 [3], $\mathbf{\Phi}' = [\gamma_1, \ldots, \gamma_p]$, where $p$ denotes the total number of terms in the document-term matrix. $\mathbf{f}_y^i \equiv \mathbf{x}_y^i / m_y^i$, which is a vector containing the frequencies of words associated with document $i$ in year $y$.

Therefore, conditional on the projections, each document's distribution of words is independent of the distribution of a stock's excess returns. In other words, the projections summarise all of the information contained in the documents. Thus, we can measure the climate risks through these projections instead of the high-dimensional and sparse document-term matrix.

Based on the sufficient statistics for each document, we run cross-sectional analysis at each fixed month and stack over the sufficient projection by each document:

[3] The proof of proposition 1 is shown in 3.7.1.

**Definition 2** *Given the previous sufficient statistics, $r^i_{y+1}x^i_y$, we have the $\beta^*$ exposure on sparse document-term matrix: $\hat{\beta}^*_{i,y} = \mathbf{\Phi}'\mathbf{f}^i_y$.*

Given the text-implied risk exposure, $\hat{\beta}^*_{i,y}$, we might be able to use the factor mimicking portfolio to retrieve the latent factor, $L_t$, as long-short portfolio construction. We stack over assets and timing in 3.2:

$$E(R_t) = \left[\tau_N, \hat{\beta}^*_{i,y}, B\right]\Gamma \tag{3.5}$$

where $R_t = (R_{1,t}, R_{2,t}....R_{N,t})^T$ and $\tau_N = (1, 1, 1, 1, ....1)^T$, a $N \times 1$ vector. $\Gamma$ is the final price of risk that we care $\Gamma = (\lambda_0, \lambda^*, \lambda)^T$, an $(K+2) \times 1$ vector and $X = \left[\tau_N, \hat{\beta}^*_{i,y}, B\right]$ denotes a $N \times (K+2)$ matrix.

### 3.4.4 Three-Pass Estimation

Given the previous models for the text-implied climate factor, we provide a three-pass process to estimate an asset pricing model with a text-implied factor.

- First, we implement the previous MNIR model to retrieve the text-implied climate risk $\hat{\beta}^*$. Then, construct the factor mimicking portfolio to estimate for $L_t$.
- Second, employ the model in 3.2 to estimate the exposures on the confounding factors, $\hat{B}$. Then, combine with to form $X = \left[\tau_N, \hat{\beta}^*, \hat{B}\right]$
- Follow the last steps in two-pass of Fama-MacBeth and evaluate the significance of the factors with $\hat{\Gamma} = (X^T\Sigma^{-1}X)^{-1}(X^T\Sigma^{-1}\bar{R}_t)$,

The convergence of our estimator is feasible following the Gentzkow et al. (2019) under the assumption that $L_t$ could be captured with the cross-variation. The core idea is that the estimation of the MNIR in Taddy (2013) is estimated with penalized MLE method for parameter sparsity. This dimensional reduction process makes the asymptotic distribution of $\Gamma$ hard to derive. However, since Gentzkow et al. (2019) shows that the MNIR penalized estimators converge to a non-degenerate distribution, $\sqrt{N}(\hat{\Phi} - \Phi) \xrightarrow{d} \tilde{\Phi}$. we might regard the estimator $\Gamma$ as a function of $\hat{\Phi}$. Therefore, we could have a corollary that

**Corollary 2** *Under MNIR assumptions with MLE penalized estimation by Gentzkow et al. (2019), the estimated $\sqrt{N}(\hat{\Gamma}(\hat{\Phi}) - \Gamma(\Phi_0)) \xrightarrow{d} \tilde{\Gamma}(\tilde{\Phi})$, where $\hat{\Gamma}, \Gamma$ and $\tilde{\Gamma}$ are continuous functions of $\hat{\Phi}, \Phi$ and $\tilde{\Phi}$*

The corollary 2 is to be proved with the property that $\Gamma$ is a continuous function of $\Phi$. Given this proposition, we could follow a similar strategy as Gentzkow et al. (2019) and exploit the subsampling method to derive the bootstrap inference for the $\hat{\Gamma}$.

## 3.5 Application in Climate Finance

In this section, we follow the procedure introduced previously and apply it for climate benchmark dictionary from (Chou and Kimbrough, 2019b) and Sautner et al. (2020) to construct self-defined dictionary. Then, with the dictionary, we construct our proposed CRS, (Climate Risk Score) with sufficient projection, at either firm or industry level. In the first, part of the section, we provide evidence to validate the sensibility of the measure in climate literature. Then, we show its pricing implication with canonical asset pricing test.

### 3.5.1 Validation

#### Visualising the Word Frequency

In each year, an aggregate document-term matrix was formed by aggregating all firms' frequencies on the same word to capture the major climate-related topics discussed in the 10-K files during that year. The visualisations of the aggregate document-term matrix from 2005 to 2017 are attached in 3.9.5. We can see that the most discussed topics match the climate events that happened in the same year, which validates that our document-term matrix sufficiently contains information on climate change. For example, the Paris Agreement was assigned in 2016, aiming at substantially reducing global greenhouse gas emissions. Term "lcfs", which represents the Low Carbon Fuel Standard (LCFS), was mentioned a lot in the 10-K files in the same year. LCFS is one of the nine discrete early action measures to reduce California's greenhouse gas (GHG) emissions that cause climate change under the AB 32 Scoping Plan. And in 2015, the terms "extreme weather", "flooding", "droughts" and "temperatures" are mostly discussed in the 10-Ks. Chasing the news, we had many extreme weathers during this year (e.g. heatwaves across Europe, the Middle East and South Asia, and flooding in the US and Africa). We had a new record of high temperature in July 2015 according to the National Oceanic and Atmospheric Administration (NOAA). In the US, Oklahoma and Texas were hit by record flash floods in May and California was facing its most severe drought in a millennium.

#### Experiments on the 48 Kenneth French Industries

Different industries could have different exposures to climate risks due to their different levels of relevance to climate change. Some industries are directly related to the change of climate. For instance, burning oil releases carbon dioxide contributing to global warming. Climate strategies are adopted by companies to reduce emissions in the oil industry, supporting some climate agreements (e.g., the Kyoto Protocol) (Skjærseth and Skodvin, 2018). Therefore, the oil

industry could be largely affected by climate news related to each year's climate events. Natural gas is a fossil fuel, though it has less carbon dioxide emission than oil. It is mainly composed of methane, which is also a greenhouse gas with the potential to contribute to global warming. Moreover, transport accounts for 1/5 of global carbon dioxide emissions. We should expect relatively high projections for the transportation industry since it could be largely affected by the policies related to carbon emission. Therefore, we would expect these industries (i.e., industry labels 30, 31, 40) should have relatively high and positive projections compared to other less climate-related industries.

To validate this, we conduct industry-level projections. We firstly sort industries by the total frequency of terms in the industry document. Note that the term-frequency in the industry document aggregates the word counts of firms inside the sector. We then project the industry-level document-term matrix, which contains the climate-related information in that industry, on the sorted labels. Figure 3.5 visualises the projections of each sector during different periods. Industries have relatively stable exposures to climate risk across different periods, and we can visualise the difference in climate risk exposure between industries. Moreover, industries 30, 31 and 40, which represent oil, natural gas, and transportation respectively, indeed have relatively high and positive projections. The reasonable distribution of projections across industries validates that the projections well contain the climate information from the textual data at the industry-level.

**Factor-Mimicking Portfolio**

With the given method, we construct our firm-level projection for CRS in Figure 3.6. Specifically, we run the MNIR with cumulative annual return from fiscal year t to t+1 on the SEC based 10K documents during the same period and obtain the projections as CRS. The climate risk scores are the projections on the labels based on the sorted returns. Based on Proposition 1, given the CRS, the returns are orthogonal to the sparse document-term matrix. Therefore, the CRS contains sufficient information for the climate risk implied in the textual data. Table A5 reports the summary statistics of the CRS each year. The range of CRS (i.e., the difference between the maximum and minimum) increases significantly in 2007 and 2014, which may be due to some large climate-related events that occurred in those years (i.e., the G8 Climate Agreement and the 3rd National Climate Assessment, respectively). Some firms and industries, whose operations might be climate-related, were greatly affected (positively or negatively) by these climate events, which increases their exposure to climate risks.

In each year, we sort firms based on their CRS, and firms are split into five quintile portfolios. The excess return of each of the five portfolios is equally/valuably weighting the stock returns in the portfolio. A factor-mimicking portfolio *CRF* can be constructed by longing the quintile 5

(i.e., high CRS) and shorting the quintile 1 (i.e., low CRS). If the CRS can truly measure the risk exposures on the climate-related textual information, we would expect *CRF* to have significant positive returns. To further test whether the systematic risk captured in the climate information can be explained by other risk factors, we check the alpha generation from the regression with the five Fama-French factors and the momentum factor introduced in the Carhart 4-factor model in 3.5.2.

Figure 3.7 shows the volatility of our mimicking portfolio's monthly returns, in addition to that of the time series of the WSJ Climate Change News Index introduced by Engle et al. (2020b) and the Crimson Hexagon's (CH) Negative Climate Change News Index for comparisons. Note that we use squared return as a proxy for the volatility. After matching the dates with those of major climate-related news announcements, we found that similar to the WSJ index and the CH index, volatility spikes during salient climate events from 2012, such as the adoption of global climate treaties (e.g., the Paris Agreement) or important global conferences about combating climate change (e.g., the 2016 Marrakech UN Climate Change Conference). Interestingly, the volatility did not capture the spike of the 2009 Copenhagen UN Climate Change Conference, which might be due to the sparsity of the climate-related words in 10Ks in the previous years. Regressing the WSJ Climate Change News Index on returns of *CRF* rolling for every four years, we can visualise the increase in correlation and it turns significant from 2012 (in Figure 3.7 Panel (b)).

**Difference-in-difference**

We next use a difference-in-difference approach to verify that firms with high projections (i.e. CRS) have higher climate risk exposures than the others. Similar to Bolton and Kacperczyk (2021b), we explore the cross-sectional variation of the effect of the Paris Agreement. We should expect higher returns from the top-quintile portfolio after the event since they have higher exposures to climate risk. Our treatment sample is the subset of firms with the largest CRS in 2014. We match the treatment group with a control group having similar firm characteristics (same as the ones we used in our return regressions) via the nearest neighbour technique in the pre-period. Panel B of Table A6 reports the balance test for the matched samples reflecting the quality of the matching. Two samples (i.e. the treatment and control groups) are similar to each other on the utilised firm characteristics. There is no significant mean difference for all of these firm-level dimensions.

We then compare the returns of treatment and control groups in the one-year period around the Paris Agreement event. The difference-in-difference regression is as follow,

$$\text{Return}_{i,t} = \alpha + \beta_1 \text{TREAT} * \text{AFTER}_{i,t} + \beta_2 \text{Controls}_{i,t} + \gamma_i + \delta_t + \varepsilon_{i,t},$$

where TREAT equals one if firms are from the treatment group and zero for the firms in the control group. AFTER is an indicator variable equal to zero for the period 2015/05–2015/11 and equal to one for the period 2015/12–2016/05, representing the pre-ante and post-ante of the event. We include firm characteristics as controls. We also include firm and year-month fixed effects in the regression, where $\gamma_i$ and $\delta_t$ represent stock $i$ and year-month $t$ fixed effects, respectively. Panel A of Table A6 reports the $\beta_1$, which measures the differential effect of the change on firms with high CRS and firms with low CRS. We find a positive and statistically significant effect on returns, implying that there is an increase in the expected mean change from before to after the Paris Agreement.

**Carbon Intensity**

We also verify that the CRS (i.e. the sufficient projection) captures climate risk exposure by investigating whether the CRS is consistent with the carbon intensity. The carbon intensity is defined as the amount of carbon by weight emitted per unit of energy consumed, which partially reflects the influence on climate change of firms' operations (see e.g. Bolton and Kacperczyk, 2021b). Therefore, if our measure truly captures climate risk exposures, we should expect a positive relation between CRS and the carbon intensity scores. We run panel regressions controlling for several firm characteristics, and firm and year fixed effects. Standard errors are clustered at the firm level. Table A7 reports the performances, showing positive and significant associations between CRS and the carbon intensity.

### 3.5.2 Empirical Analysis

In this section, we examine whether the stock market prices the text-implied climate risk in the cross-section of returns. We first conduct univariate portfolio analyses to review the performance difference in returns of firms exposed to high and low climate risk. We conjecture a compensation should be given to an investor for bearing additional climate risks. We also test whether these abnormal returns are explained by other risk factors, including the market aggregate return *MKT*, the mimicking portfolios for size risk *SMB*, book-to-market risk *HML*, profitability risk *RMW* and investment risk *CMA*. These are the factors of the Fama-French five-factor model Fama and French (2015). We also considered a momentum factor *UMD* introduced by Carhart (1997). Second, we utilise bivariate portfolio sorts based on several firm characteristics (e.g. book-to-market, sales growth) to check whether some subsamples with different firm characteristics have more prevalent exposure to the climate risk. Third, we test this text-implied latent factor within the classical Fama-MacBeth cross-sectional regression. Fourth, we aggregate the document-term matrix within each of the 48 Kenneth French industries and run MNIR at the industry-level to investigate properties of industry-level projections.

### Time Variation of Climate Risk

The climate risk measurement may be time variate since the risks to the climate-related events of stocks can differ in different periods. Although different stocks may have different exposures to the risks associated with climate-related events, the majority of stocks could be affected by the same events in the same direction (i.e., positive or negative). Therefore, in each year, we conjecture that the majority of firm projections are in the same direction (positive or negative).

In Figure 3.6, we can visualise that the majority of the firms' projections are in the same direction during the same period. The time variation of the projections validates that the climate effect is time variate and can change every year (e.g., due to different events). Corresponding to Figure 3.7, in 2008, the majority of the stocks have positive projections, which could be due to the G8 Climate Agreement, which was struck that year. However, note that we still cannot eliminate the probability that the risks that the firms take stem from the global financial crisis (GFC). In 2011 and 2012, the projections are close to zero, which might be because there were fewer large climate-related events within this period. In 2014 and 2016, the positive projections might be due to the 3rd National Climate Assessment & EPA Climate Change Initiative and the Paris Agreement, respectively.

### Estimate Climate Risk *CRS*

**Average Monthly Excess Returns of Climate-Projection-Sorted Portfolios** Five portfolios are built based on the 20th, 40th, 60th, and 80th percentiles (or quintile) of the sorted CRS. The label '1' or 'L' represents the portfolio with the lowest climate projections, and the label '5' or 'H' represents the portfolio with the highest climate projections. Table A8 shows the average monthly excess returns of 5 portfolios during the period from 2005 to 2017. We can see a clear increasing pattern from label '1' to '5' in Table A8 (i.e. the higher the projections are, the higher the excess returns will be). A long-short portfolio created by longing the quintile 5 (i.e., high CRS) and shorting the quintile 1 (i.e., low CRS) generates significant positive excess returns.

**Factor-Mimicking Portfolio of Climate Risk** Given the CRS sufficiently preserved the climate-related information implied in the textual data, a mimicking portfolio, namely *CRF*, can be constructed based on the sorted CRS. Firms are split into five equally weighted portfolios based on the quintiles of CRS. The *CRF* longs the stocks in the top quintile and shorts the stocks in the bottom quintile.

The correlations between the *CRF* and the five Fama-French factors, which are the market return *MKT*, size factor *SMB*, value factor *HML* (based on the book-to-market ratio), prof-

itability factor *RMW* and investment factor *CMA*, are shown in Panel A of Table A9. We also include a momentum factor *UMD* introduced by Carhart (1997). Note that we used monthly returns to compute the correlations. The *CRF* is negatively correlated with the *MKT*, *SMB*, *HML* and *CMA* factors and is slightly positively correlated with *RMW* and *UMD*.

Panel B of Table A9 investigate the mean, standard deviation and auto-correlations of the monthly returns of different factor-mimicking portfolios. The auto-correlation of the *CRF* decays quickly towards zero (i.e., from the first lag), indicating that the *CRF* has no unit root and is a stationary process. Auto-correlations of other factors are also included for comparison.

Figure 3.9 shows the *CRF* and other factors' accumulated excess returns. Unlike other factors with sharp decreases during the Global Financial Crisis (GFC), the *CRF* was quite stable during this period, although it had some vibrations. The pattern of the *CRF* since the GFC was similar to that of the market but more stable. The *CRF* increased steadily during 2010-2012, which could be because there were no major climate-related events in this period. The sharp increase after the steady period in 2014 could be due to the 3rd National Climate Assessment.

**Univariate Portfolio-level Analysis**

We expect firms with greater CRS to have higher average returns. Therefore, in each period (i.e., each year, in this case), we can sort all of the stocks based on their projections and build five portfolios based on quintile cut points. A factor-mimicking portfolio *CRF* can be built up by longing the top quintile and shorting the bottom quintile. We compute both equal-weighted (ew) and value-weighted (vw) monthly portfolio returns. We report average excess returns of *CRF* as well as portfolio alphas adjusted for market risk (MKT) according to the CAPM or, alternatively, for market, size (SMB) and value (HML) factor exposures according to the Fama-French three-factor model (three-factor alphas) Fama and French (1993), as well as alphas adjusted for market, size, value, profitability (RMW) and investment (CMA) factor exposures according to the Fama-French five-factor model (five-factor alphas) Fama and French (2015). The results are presented in Table A8, which implies that firms with high climate risk have higher future excess returns and positive alphas net of well-known risk factors. An abnormal return from 1.2% to 2.1%, depending on the specification of the benchmark model, can be generated through the long-short portfolio. Note that the standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation.

**Bivariate Portfolio-level Analysis**

In this section, we utilise bivariate portfolio sorts. Specifically, we sort stocks at the end of each year in ascending order on the basis of their climate risk and allocate them into five groups based

on the quintile cut points. Then we independently sort stocks into ascending order according to some firm-level characteristics (e.g. market value). We assign firms into two portfolios (low and high) based on the median values for each of the firm-level characteristics. The intersection of the above classifications yields several double-sorted portfolios. The results are presented in Table A10. We report both equal-weighted and value-weighted average returns, Carhart four-factor alphas and five-factor alphas for the spread strategy (longing stocks with the highest CRS and shorting stocks with the lowest CRS) within each HIGH and LOW classification. Again, the standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. The results imply that the outperformance of the firms with high CRS (i.e. high climate risk) persist for both equal- and value-weighted portfolio returns and for different measures of excess returns (i.e., average excess return, Carhart four-factor alphas, and five-factor alpha). These results ensure we are not capturing proxies of other well-known risks.

**Individual-stock-level Fama-MacBeth Cross-sectional Regression**

For each month of a sample, we run cross-sectional regressions of excess stock returns on lagged CRS, which is the climate-risk exposure. Table A11 reports the average of the estimated coefficients of the monthly cross-sectional regressions as well as the Newey-West standard errors. We control for 10-K characteristics such as the book-to-market ratio, sales, sales growth, and leverage status. The definitions of these firm characteristics are included in 3.9.1. The results show that the coefficient estimate of climate risk is positive and significant. To further show the predictability of the climate risk index is not a short-term phenomenon, we regress it with the stock returns up to 12-month into the future. The results indicate that our climate risk index has relatively long-term predictive power on monthly cross-sectional variation in stock returns.

### 3.5.3 Portfolio Analysis: Robustness Tests

We present several robustness tests in this section. The results are reported in Table A12. In Panel A, we try to build up portfolios based on different cut points (i.e. sorting the climate risk measure and separating them into 1/3, 1/4, 1/5, 1/10). The results remain unaffected and we can obtain the outperformance of stocks exposed to high climate risk. Interestingly, when we focus on more "extreme" portfolios to calculate the spread (i.e., quartile, quintile and decile portfolios), the return spread increases. In Panel B, we assume a firm's exposure to climate risk is similar to other firms in the same industry, hence, we replace all zeros with the median of climate-risk measure of each industry. In Panel C, we assume a firm's exposure to climate risk is likely to be persistent across time, hence, we replace zeros with the first available non-zero observation before time $t$ of each firm. We also try to exclude each of the

48 Kenneth French industries to test whether the results are driven by any particular industry. The results are visualised in Figure 3.8. Overall, these results validate the outperformance of the high-climate-risk portfolio to the low-climate-risk portfolio.

## 3.6 Conclusion

We provide a unified framework to process topic-specific risk nested in text data. Focusing on the variation of climate-related terms in SEC documents, we extract a sufficient projection as a proxy for climate risk exposure. The benchmark in use is the supervised learning algorithm, MNIR by Taddy (2013). The contribution of our method in risk exposure estimation relies on its sufficiency in variational projection and computational scalable efficiency. We find that this text-implied latent factor is testable within the classical Fama-MacBeth regression following the Gentzkow et al. (2019). On the empirical side, we find that our proposed climate risk exposure ($\beta^*$) triggered pricing inefficiency. The premium is both statistically significant and of great economic magnitude.

In addition, we show that the interpretation of our projected measure is sensible at firm-level. We provide lots of evidence internally and externally to verify our conjectures. First, we display the relative importance of terms in the whole documents and emphasize their sensibility for climate topics. Second, we validate its relationship with important climate indicators including carbon intensity in Bolton and Kacperczyk (2021b) and WSJ in Engle et al. (2020b). Finally, we provide event comparison and show that the method is sensible for key event studies. Overall, the MNIR extracted climate exposure should be not just technically important but empirically matter in the climate finance literature, especially climate asset pricing.

From the perspective of text mining application in economics and finance, our paper is generally targeting at providing a unified framework to process specific text-implied risk in documents. Even though we limit our contributions within the field of climate finance, it's technically accessible to extend it to other controversial text-implied risk factors, like labor idiosyncratic risk or cybersecurity risk, and test their relative importance within the classical asset pricing framework. Moreover, the proposed framework is also applicable to other fields like the nowcasting in macroeconomics or key macroeconomic forecasting. We will investigate these directions in future work.

Table A1: List of the IPCC Reports

This table lists the reports we used for keywords discovery and their descriptions. We considered all available IPCC synthesis reports (i.e. first - fifth) and some special reports prepared by their working groups.

| Title | Description |
|---|---|
| The 1990 and 1992 IPCC Assessments | IPCC First Synthesis Report (IPCC, 1992) |
| IPCC Second Assessment Climate Change 1995 | IPCC Second Synthesis Report (IPCC, 1995) |
| Climate Change 2001: Synthesis Report | IPCC Third Synthesis Report (IPCC, 2001a) |
| Climate Change 2007 Synthesis Report | IPCC Fourth Synthesis Report (IPCC, 2007) |
| Climate Change 2014 Synthesis Report | IPCC Fifth Synthesis Report (IPCC, 2014) |
| Climate Change and Land | An IPCC special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems (Shukla et al., 2019) |
| The Ocean and Cryosphere in a Changing Climate | This special report was prepared following an IPCC Panel decision in 2016 to prepare three Special Reports during the Sixth Assessment Cycle (the other two are Global Warming of 1.5°C and Climate Change and Land) (Pörtner et al., 2019) |
| Global Warming of 1.5 $^{o}$C | An IPCC special report on the impacts of global warming of 1.5 °C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty (Zhai et al., 2018) |
| Renewable Energy Sources and Climate Change Mitigation | This report impartially assesses the scientific literature on the potential role of renewable energy in the mitigation of climate change for policymakers, the private sector, academic researchers and civil society (Edenhofer et al., 2011) |

Table A2: Top-5 Similar Words of Initial Unigrams

This table shows our initial set of unigrams and the 5 most similar words based on the word embeddings from word2vec. We use 200 neurons in the hidden layer of the Skip-gram's neural network model, which means we have a dense vector with size 200 to represent each unigrams in the corpus. Cosine similarity is used to measure the similarities between words. The 5 most similar words of each of the unigrams in the initial set are reported in the table. Note that since we removed all digital numbers during the data cleaning, "co" represents "CO2" (carbon dioxide), "ch" represents "CH4" (methane), "oh" stands for Hydroxide ion and "n" represents "N2O" (nitrous oxide).

| Initial Unigrams | 1st | 2nd | 3rd | 4th | 5th |
|---|---|---|---|---|---|
| methane | ch | oxide | dioxide | emission | co |
| weather | extreme | event | meteorological | drought | storm |
| climate | change | climatic | adaptation | impact | human |
| emission | greenhouse | co | gas | methane | emit |
| agriculture | agricultural | forestry | food | smallholder | management |
| rain | precipitation | drought | runoff | rainfall | tropical |
| extinction | specie | invasion | endemic | richness | unique |
| rps | fit | quota | enact | incentive | enforcement |
| ozone | stratospheric | depletion | aerosol | stratosphere | precursor |
| habitat | specie | wetland | biodiversity | ecosystem | mammal |
| emit | fossil | emission | release | absorb | zero |
| co | dioxide | emission | methane | concentration | carbon |
| oil | palm | crude | coal | fuel | ethanol |
| renewable | energy | electricity | technology | deployment | nuclear |
| temperature | warm | warming | surface | mean | precipitation |
| ecosystem | ecosystems | biodiversity | wetland | habitat | terrestrial |
| drought | flood | extreme | event | rainfall | stress |
| flood | drought | flooding | storm | surge | landslide |
| adaptation | adaptive | adapt | mitigation | climate | resilience |
| reserve | conventional | oil | corridor | spot | fossil |
| green | urban | street | roof | brown | clean |
| mitigation | policy | option | adaptation | emission | strategy |
| hurricane | cyclone | storm | typhoon | maria | island |
| rainwater | irrigation | harvesting | dam | catchment | irrigate |
| pollutant | pollution | air | particulate | pm | ozone |
| atmosphere | atmospheric | earth | dioxide | sink | release |
| ch | methane | oxide | oh | oxidation | n |
| precipitation | rainfall | runoff | frequency | surface | temperature |
| rainfall | precipitation | runoff | extreme | event | monsoon |

Table A3: Initial Bigrams for Searching Climate Change Bigrams

This table shows our initial set of bigrams for the bigram discovery algorithm. These initial bigrams are defined by Sautner et al. (2020) since they are unambiguously related to climate change.

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

Table A4: Top-100 Bigrams Captured by Climate Change Exposure

This table reports the top-100 bigrams associated with the frequency in target set. The target set is identified via the keyword discovery algorithm, which contains potential climate related bigrams other than the initial bigrams.

| Bigram | Freq | Bigram | Freq | Bigram | Freq |
|---|---|---|---|---|---|
| change climate | 263 | renewable resource | 30 | reduce emission | 19 |
| land use | 81 | climate sensitivity | 28 | induce change | 19 |
| sea ice | 66 | project change | 28 | climate response | 19 |
| climate system | 60 | food security | 27 | energy source | 19 |
| climate impact | 55 | wind solar | 27 | cost energy | 19 |
| climate variability | 52 | cover change | 26 | energy service | 19 |
| sustainable development | 49 | impact change | 26 | water quality | 18 |
| use change | 48 | offshore wind | 26 | carbon capture | 18 |
| climate relate | 47 | climatic change | 26 | time scale | 18 |
| energy supply | 47 | high level | 26 | sea surface | 18 |
| global climate | 44 | surface temperature | 25 | change regional | 18 |
| impact climate | 42 | change change | 25 | greenhouse warm | 18 |
| climate policy | 42 | variability change | 24 | warm climate | 17 |
| environmental change | 40 | solar power | 24 | climate resilient | 17 |
| climate extreme | 38 | electricity generation | 23 | greenhouse effect | 17 |
| natural gas | 38 | energy technology | 23 | air pollutant | 17 |
| fossil fuel | 37 | system change | 23 | industrial level | 17 |
| temperature change | 36 | climate induce | 23 | technical potential | 17 |
| land cover | 35 | future change | 22 | wind farm | 17 |
| large scale | 35 | adaptation strategy | 22 | soil carbon | 17 |
| global temperature | 35 | power generation | 22 | emission scenario | 17 |
| climate adaptation | 35 | cross box | 21 | energy use | 17 |
| regional climate | 34 | climate risk | 21 | global regional | 16 |
| global change | 34 | natural system | 21 | energy consumption | 16 |
| low carbon | 33 | global energy | 21 | climate land | 16 |
| renewable source | 32 | adaptation mitigation | 21 | limit warm | 16 |
| future climate | 32 | change ocean | 21 | change temperature | 16 |
| wind turbine | 32 | energy demand | 21 | carbon cycle | 15 |
| change global | 32 | climate projection | 21 | carbon stock | 15 |
| energy system | 31 | climate warming | 21 | climate climate | 15 |
| change land | 30 | relate change | 21 | coral reef | 15 |
| mitigation adaptation | 30 | atmospheric concentration | 21 | emission reduction | 15 |
| power plant | 30 | global environmental | 20 | | |
| climate mitigation | 30 | ocean energy | 20 | | |

Table A5: Summary Statistics

This table reports the summary statistics of the climate projection (i.e. CRS) in Panel A, the summary statistics of firm characteristics we used as controls and the carbon intensity from MSCI ESG (see e.g. Bolton and Kacperczyk, 2021b) in Panel B, and the summary statistics of risk factors in Panel C. The definitions of firm characteristics are in 3.9.1. *MKT* is the market aggregate return and *SMB*, *HML*, *RMW* and *CMA* are the mimicking portfolios for the size risk, book-to-market risk, profitability risk and investment risk, respectively, in the Fama and French (2015). *UMD* is the momentum factor introduced by Carhart (1997). The columns in the tables represent the minimum, median, maximum, mean, standard deviation and the range (i.e. the difference between maximum and minimum) of the projections respectively. The sample period is from 2005 to 2017.

Panel A: Summary Statistics of Climate Projections

| Year | Min | Medium | Max | Mean | Std | MaxMinDiff |
|---|---|---|---|---|---|---|
| 2005 | -0.244 | -0.048 | 0.214 | -0.033 | 0.056 | 0.458 |
| 2006 | -0.290 | -0.040 | 0.204 | -0.024 | 0.053 | 0.494 |
| 2007 | -0.492 | -0.138 | 0.412 | -0.080 | 0.141 | 0.903 |
| 2008 | -0.256 | 0.091 | 0.230 | 0.065 | 0.081 | 0.486 |
| 2009 | -0.382 | -0.029 | 0.106 | -0.022 | 0.034 | 0.488 |
| 2010 | -0.226 | -0.051 | 0.130 | -0.038 | 0.052 | 0.355 |
| 2011 | -0.270 | 0.027 | 0.216 | 0.024 | 0.037 | 0.486 |
| 2012 | -0.157 | 0.034 | 0.111 | 0.025 | 0.034 | 0.268 |
| 2013 | -0.119 | -0.050 | 0.117 | -0.037 | 0.043 | 0.235 |
| 2014 | -0.519 | 0.166 | 0.320 | 0.102 | 0.147 | 0.839 |
| 2015 | -0.182 | 0.035 | 0.202 | 0.028 | 0.046 | 0.385 |
| 2016 | -0.205 | 0.080 | 0.225 | 0.055 | 0.075 | 0.430 |
| 2017 | -0.154 | -0.003 | 0.058 | -0.002 | 0.016 | 0.212 |

Panel B: Summary Statistics of Firm Characteristics and Carbon Intensity

| | Min | Medium | Max | Mean | Std |
|---|---|---|---|---|---|
| Book-to-market | 0 | 0.58 | 720.83 | 0.77 | 2.80 |
| Sales | -25623 | 98.30 | 136267 | 831.41 | 3698.34 |
| Leverage | 0 | 0.16 | 39.59 | 0.23 | 0.30 |
| Carbon Intensity | 0 | 6.60 | 10 | 6.41 | 2.55 |

Panel C: Summary Statistics of Risk Factors

| Factor | Min | Medium | Max | Mean | Std |
|---|---|---|---|---|---|
| MKT | -0.172 | 0.011 | 0.114 | 0.007 | 0.041 |
| SMB | -0.047 | 0.002 | 0.065 | 0.001 | 0.024 |
| HML | -0.114 | -0.002 | 0.081 | -0.001 | 0.026 |
| RMW | -0.041 | 0.001 | 0.053 | 0.002 | 0.016 |
| CMA | -0.032 | -0.0002 | 0.035 | 0.0001 | 0.014 |
| UMD | -0.343 | 0.003 | 0.128 | 0.001 | 0.045 |

Table A6: Paris Agreement and Stock Returns: Difference-in-differences Estimation

Panel A of this table reports the results of difference-in-difference analysis on Paris Agreement. The dependent variable is the stock return. We use firms in the largest quintile of the distribution of firms sorted by CRS in 2014. Matching is conducted to find a control group of firms with similar characteristics. In particular, we use the nearest neighbour method. The matching characteristics are the same as those in our previous regressions. Similar to Bolton and Kacperczyk (2021b), TREAT is a generic indicator variable taking the value one for firms in the treatment sample and zero for firms in the control sample, and AFTER is an indicator variable equal to zero for the period 2015/05–2015/11 and equal to one for the period 2015/12–2016/05. Firm and year-month fixed effects are also included in the regression. For our difference-in-difference specifications we report coefficients from the following regression $\text{Return}_{i,t} = \alpha + \beta_1 \text{TREAT} * \text{AFTER}_{i,t} + \beta_2 \text{Controls}_{i,t} + \gamma_i + \delta_t + \varepsilon_{i,t}$. $\gamma_i$ and $\delta_t$ represent stock $i$ and year-month $t$ fixed effects, respectively. Standard errors are clustered at the firm level. We include industry fixed effect in column 2. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively. Panel B reports the balance tests results of the matching along firm-level dimensions in the pre-period.

Panel A: Difference-in-differences Estimation

| | *Dependent variable:* | |
|---|---|---|
| | Returns | |
| | (1) | (2) |
| TREAT*AFTER | 0.02078*** | 0.02080*** |
| | (0.00421) | (0.00421) |
| Controls | Yes | Yes |
| Firm F.E. | Yes | Yes |
| Year/month F.E. | Yes | Yes |
| Industry F.E. | No | Yes |
| Observations | 13,588 | 13,588 |
| Adjusted $R^2$ | 0.14829 | 0.14522 |

Panel B: Balance Test

| Characteristic | Control Group Mean | Treatment Group Mean | t-statistic |
|---|---|---|---|
| Book-to-market | 0.744 | 0.789 | -0.973 |
| Sales | 1,262.424 | 1,501.675 | -0.793 |
| Sales Growth | 0.216 | 0.251 | -0.182 |
| Leverage | 0.279 | 0.272 | 0.639 |

Table A7: Climate Risk and Carbon Intensity

This table reports the panel regressions of climate projections on a climate indicator (i.e. carbon intensity in Bolton and Kacperczyk (2021b)), controlling for several firm characteristics (book-to-market, sales, leverage, etc.) and firm and year fixed effects. Standard errors are clustered at the firm level. Analytical variable definitions are provided in 3.9.1. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| | *Dependent variable:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Carbon Intensity | | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Climate Risk Index | 0.240*** | 0.217*** | 0.111*** | 0.053*** | 0.054*** | 0.051*** | 0.051*** | 0.051*** |
| | (0.012) | (0.013) | (0.010) | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Book-to-market | | | | | −0.009 | −0.009 | −0.009 | −0.015 |
| | | | | | (0.018) | (0.018) | (0.018) | (0.018) |
| Sales | | | | | | 0.075** | 0.075** | 0.081*** |
| | | | | | | (0.032) | (0.032) | (0.031) |
| Sales Growth | | | | | | | −0.00002 | 0.00001 |
| | | | | | | | (0.0003) | (0.0003) |
| Leverage | | | | | | | | 0.011 |
| | | | | | | | | (0.027) |
| Clustered SE | Firm | Firm | Firm | Firm | Firm | Firm | Firm | Firm |
| Year fixed effects | No | Yes | No | Yes | Yes | Yes | Yes | Yes |
| Firm fixed effects | No | No | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 55,923 | 55,923 | 55,923 | 55,923 | 54,219 | 54,193 | 53,737 | 49,080 |
| Adjusted $R^2$ | 0.057 | 0.105 | 0.592 | 0.699 | 0.701 | 0.701 | 0.702 | 0.706 |

Table A8: Climate Risk Portfolios

This table reports average (monthly) excess returns, CAPM alphas, alphas regressing with Fama-French three-factor model (Fama and French, 1993), alphas regressing with Carhart four-factor model (Carhart, 1997) and alphas regressing with Fama-French five-factor model (Fama and French, 2015) for portfolios constructed on the basis of our Climate Risk Index. Starting from January 2005, we sort stocks at the end of each year in ascending order on the basis of their Climate Risk and allocate them into five groups (Based on quintiles). We form the spread strategy H-L that is long the portfolio with the highest climate risk stocks (H) and short the portfolio with the lowest climate risk stocks (L). Table reports returns for equally-weighted (ew) and value-weighted (vw) portfolios over the period January 2005- December 2017. The standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| Alphas | Equal/Value-Weighted | L | 2 | 3 | 4 | H | H-L |
|---|---|---|---|---|---|---|---|
| Excess return | ew | -0.006 | 0.002 | 0.005 | 0.007 | 0.014*** | 0.020*** |
| | | (0.006) | (0.005) | (0.006) | (0.006) | (0.005) | (0.003) |
| | vw | -0.004 | 0.002 | 0.004 | 0.007* | 0.009** | 0.012*** |
| | | (0.004) | (0.004) | (0.004) | (0.004) | (0.003) | (0.002) |
| CAPM alpha | ew | -0.009 | 0.000 | 0.002 | 0.005 | 0.012** | 0.020*** |
| | | (0.005) | (0.005) | (0.005) | (0.005) | (0.005) | (0.003) |
| | vw | -0.005 | 0.001 | 0.003 | 0.005 | 0.008** | 0.013*** |
| | | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.002) |
| FF 3-factor alpha | ew | -0.008 | 0.001 | 0.002 | 0.005 | 0.012** | 0.020*** |
| | | (0.006) | (0.005) | (0.006) | (0.006) | (0.006) | (0.003) |
| | vw | -0.005 | 0.001 | 0.002 | 0.005 | 0.008** | 0.013*** |
| | | (0.004) | (0.004) | (0.005) | (0.004) | (0.004) | (0.002) |
| Carhart 4-factor alpha | ew | -0.008 | 0.001 | 0.003 | 0.005 | 0.012** | 0.020*** |
| | | (0.006) | (0.005) | (0.005) | (0.006) | (0.005) | (0.003) |
| | vw | -0.004 | 0.001 | 0.003 | 0.005 | 0.008** | 0.013*** |
| | | (0.004) | (0.004) | (0.005) | (0.004) | (0.004) | (0.002) |
| FF 5-factor alpha | ew | -0.007 | 0.002 | 0.004 | 0.006 | 0.014*** | 0.021*** |
| | | (0.006) | (0.005) | (0.006) | (0.006) | (0.005) | (0.003) |
| | vw | -0.004 | 0.002 | 0.003 | 0.006 | 0.009*** | 0.013*** |
| | | (0.004) | (0.004) | (0.005) | (0.004) | (0.003) | (0.002) |

Table A9: Correlations and Auto-correlations of Different Factors

A correlation matrix of seven factors' monthly returns is summarised in Panel A. *MKT* is the market risk factor. *SMB*, *HML*, *RMW* and *CMA* are the mimicking portfolios for the size risk, book-to-market risk, profitability risk and investment risk respectively in Fama and French (2015). *UMD* is a momentum factor introduced by Carhart (1997). *CRF* is the long-short portfolio based on CRS. The table covers the period 2005-2017. Panel B summarises the mean, standard deviation and auto correlations of different factor's monthly returns. Note that we show the means as percentage.

Panel A: Correlations between Different Factors

| | CRS | MKT | SMB | HML | RMW | CMA | UMD |
|---|---|---|---|---|---|---|---|
| CRS | 1 | -0.119 | -0.079 | -0.108 | 0.009 | -0.040 | 0.009 |
| MKT | -0.119 | 1 | 0.390 | 0.284 | -0.473 | -0.031 | -0.341 |
| SMB | -0.079 | 0.390 | 1 | 0.289 | -0.415 | 0.107 | -0.197 |
| HML | -0.108 | 0.284 | 0.289 | 1 | -0.239 | 0.505 | -0.441 |
| RMW | 0.009 | -0.473 | -0.415 | -0.239 | 1 | 0.037 | 0.240 |
| CMA | -0.040 | -0.031 | 0.107 | 0.505 | 0.037 | 1 | -0.119 |
| UMD | 0.009 | -0.341 | -0.197 | -0.441 | 0.240 | -0.119 | 1 |

Panel B: Auto-correlations of Factors

| | Mean | S.d. | Lag-1 | Lag-2 | Lag-3 |
|---|---|---|---|---|---|
| CRS | 1.954 | 0.032 | 0.130 | -0.078 | -0.135 |
| MKT | 0.727 | 0.041 | 0.140 | -0.035 | 0.095 |
| SMB | 0.136 | 0.024 | -0.154 | 0.021 | 0.039 |
| HML | -0.065 | 0.026 | 0.210 | -0.063 | -0.051 |
| RMW | 0.218 | 0.016 | 0.217 | 0.011 | -0.085 |
| CMA | 0.006 | 0.014 | 0.122 | 0.079 | 0.011 |
| UMD | 0.064 | 0.045 | 0.221 | -0.077 | 0.052 |

Table A10: Double-Sorted Portfolios

This table reports average excess returns, 4-factor alphas from Carhart four-factor model (Carhart, 1997) and 5-factor alphas from the FF5 model (Fama and French, 2015) for double-sorted portfolios on the basis of the climate risk index and firm characteristics (book-to-market, sales, leverage, etc.). Analytical variable definitions are provided in 3.9.1. Starting from January 2005, we sort stocks at the end of each year in ascending order on the basis of their Climate Risk and allocate them into five groups, and we also independently sort stocks into ascending order according to the value of each characteristic mentioned above and allocate them into two portfolios (LOW and HIGH) based on median values for each month. The intersection of these two classifications yields the double-sorted portfolios. We track the performance of the intersection portfolios over the following month until these are rebalanced. We report both equal-weighted and value-weighted average returns and five-factor alphas for the spread strategy High-Low Climate Risk Stocks within each HIGH and LOW classification. The standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| Firm Characteristics | High/Low | Avg.Ret(ew) | Carhart4 Alp(ew) | FF5 Alp(ew) | Avg.Ret(vw) | Carhart4 Alp(vw) | FF5 Alp (vw) |
|---|---|---|---|---|---|---|---|
| Market Value | LOW | 0.023*** | 0.023*** | 0.025*** | 0.021*** | 0.022*** | 0.023*** |
| | | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.003) |
| | HIGH | 0.015*** | 0.016*** | 0.016*** | 0.012*** | 0.013*** | 0.013*** |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Book-to-market | LOW | 0.018*** | 0.019*** | 0.019*** | 0.012*** | 0.012*** | 0.013*** |
| | | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| | HIGH | 0.018*** | 0.018*** | 0.019*** | 0.012*** | 0.013*** | 0.013*** |
| | | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Sales | LOW | 0.022*** | 0.023*** | 0.024*** | 0.020*** | 0.021*** | 0.022*** |
| | | (0.003) | (0.003) | (0.003) | (0.004) | (0.004) | (0.004) |
| | HIGH | 0.016*** | 0.017*** | 0.017*** | 0.012*** | 0.013*** | 0.013*** |
| | | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Sales Growth | LOW | 0.019*** | 0.020*** | 0.021*** | 0.012*** | 0.013*** | 0.013*** |
| | | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| | HIGH | 0.018*** | 0.019*** | 0.019*** | 0.013*** | 0.014*** | 0.014*** |
| | | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Leverage | LOW | 0.017*** | 0.018*** | 0.018*** | 0.012*** | 0.012*** | 0.012*** |
| | | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) |
| | HIGH | 0.022*** | 0.022*** | 0.023*** | 0.013*** | 0.014*** | 0.015*** |
| | | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |

Table A11: Individual-stock-level Fama-MacBeth Cross-sectional Regressions

This table reports the results from Fama-MacBeth regressions on the relation between our Climate Risk Index and subsequent monthly stock returns (1-month to 12-month). For each month of our sample we run cross sectional regressions of excess stock returns on lagged climate risk and lagged firm characteristics (book-to-market, sales, leverage, etc.). All variables are defined in 3.9.1. The continuous variables are standardised to have a mean of 0 and standard deviation of 1. The coefficients are reported as time-series averages of the estimates from the cross sectional regressions. The standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

| | Returns(t+1) | Returns(t+1) | Returns(t+2) | Returns(t+3) | Returns(t+6) | Returns(t+9) | Returns(t+12) |
|---|---|---|---|---|---|---|---|
| Climate Risk Index | 0.1254*** | 0.1214*** | 0.117*** | 0.0867*** | 0.0534** | 0.0808 | 0.1374 |
| | (0.0182) | (0.0175) | (0.0185) | (0.021) | (0.0254) | (0.0514) | (0.1113) |
| Book-to-market | | 0.01*** | 0.0075* | 0.0038 | -0.0043 | -0.0055 | 0.0003 |
| | | (0.0038) | (0.0041) | (0.0035) | (0.0042) | (0.0035) | (0.0042) |
| Sales | | 0.0002 | 0.0002 | 0.0001 | 0.0001 | -0.0013 | -0.0011 |
| | | (0.0003) | (0.0003) | (0.0003) | (0.0004) | (0.0013) | (0.0013) |
| Sales Growth | | 0.015 | 0.0361** | 0.0385** | 0.001 | 0.1621 | -0.5453 |
| | | (0.011) | (0.0171) | (0.0152) | (0.0159) | (0.5876) | (0.8334) |
| Leverage | | -0.0011 | -0.0009 | -0.0008 | -0.0006 | 0.0016 | 0.004 |
| | | (0.001) | (0.001) | (0.001) | (0.0011) | (0.002) | (0.0061) |

Table A12: Climate Risk Portfolios - Robustness Tests

This table reports average excess returns, 4-factor alphas from Carhart four-factor model (Carhart, 1997) and 5-factor alphas from the FF5 model (Fama and French, 2015) for the spread strategy that is long the portfolio with the highest climate risk stocks and short the portfolio with the lowest climate risk stocks. Results are reported both for equally-weighted and value-weighted portfolios. In Panel A we form our portfolios based on different cut points. In Panel B, we form our portfolios based on another revised climate risk measure, which replaces all zeros with the industry median value in any given year. We reorganise firms into the Kenneth French 48 industry groupings based on their 4-digit SIC codes. In Panel C, we form our portfolios based on a revised climate risk measure, which replaces all zeros with the previous non-zero observation for each firm. The standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. ***, ** and * denote statistical significance at 1%, 5% and 10% levels, respectively.

Panel A: Different Portfolio Splits

| Portfolio Split | Avg.Ret(ew) | Carhart4 Alp(ew) | FF5 Alp(ew) | Avg.Ret(vw) | Carhart4 Alp(vw) | FF5 Alp(vw) |
|---|---|---|---|---|---|---|
| Tercile Portfolios (P3-P1) | 0.014*** | 0.014*** | 0.014*** | 0.009*** | 0.010*** | 0.010*** |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Quartile Portfolios (P4-P1) | 0.017*** | 0.017*** | 0.018*** | 0.011*** | 0.012*** | 0.012*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Quintile Portfolios (P5-P1) | 0.020*** | 0.020*** | 0.021*** | 0.012*** | 0.013*** | 0.013*** |
| | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Decile Portfolios (P10-P1) | 0.027*** | 0.028*** | 0.029*** | 0.017*** | 0.018*** | 0.019*** |
| | (0.004) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) |

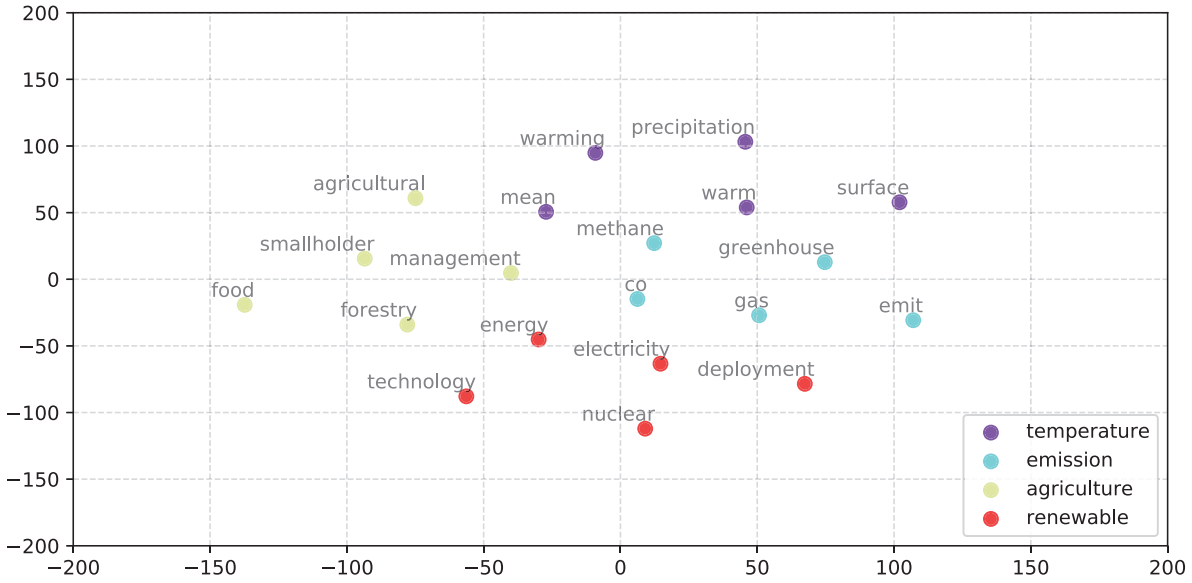
Panel B: Replacing Zeros with Industry Medians

| Portfolio Split | Avg.Ret(ew) | Carhart4 Alp(ew) | FF5 Alp(ew) | Avg.Ret(vw) | Carhart4 Alp(vw) | FF5 Alp(vw) |
|---|---|---|---|---|---|---|
| Tercile Portfolios (P3-P1) | 0.014*** | 0.014*** | 0.015*** | 0.009*** | 0.010*** | 0.010*** |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Quartile Portfolios (P4-P1) | 0.017*** | 0.018*** | 0.018*** | 0.011*** | 0.011*** | 0.012*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Quintile Portfolios (P5-P1) | 0.020*** | 0.020*** | 0.021*** | 0.012*** | 0.013*** | 0.013*** |
| | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Decile Portfolios (P10-P1) | 0.028*** | 0.028*** | 0.029*** | 0.017*** | 0.018*** | 0.019*** |
| | (0.004) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) |

Panel C: Replacing Zeros with Previous Non-zero Observation

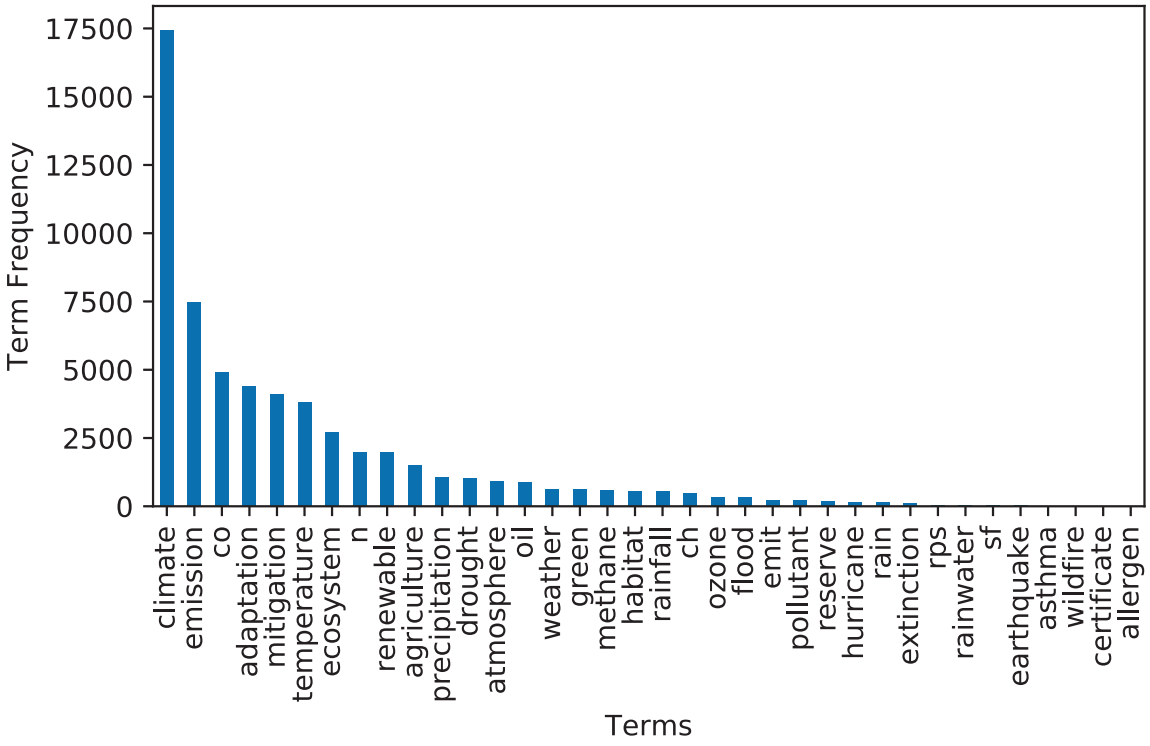
| Portfolio Split | Avg.Ret(ew) | Carhart4 Alp(ew) | FF5 Alp(ew) | Avg.Ret(vw) | Carhart4 Alp(vw) | FF5 Alp(vw) |
|---|---|---|---|---|---|---|
| Tercile Portfolios (P3-P1) | 0.014*** | 0.014*** | 0.014*** | 0.009*** | 0.010*** | 0.010*** |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.001) | (0.001) |
| Quartile Portfolios (P4-P1) | 0.017*** | 0.018*** | 0.018*** | 0.011*** | 0.011*** | 0.012*** |
| | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) | (0.002) |
| Quintile Portfolios (P5-P1) | 0.020*** | 0.020*** | 0.020*** | 0.012*** | 0.013*** | 0.013*** |
| | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| Decile Portfolios (P10-P1) | 0.027*** | 0.028*** | 0.029*** | 0.017*** | 0.018*** | 0.018*** |
| | (0.004) | (0.004) | (0.004) | (0.003) | (0.003) | (0.003) |

Figure 3.1: Clusters of Similar Words from IPCC Reports (preplexity=20)

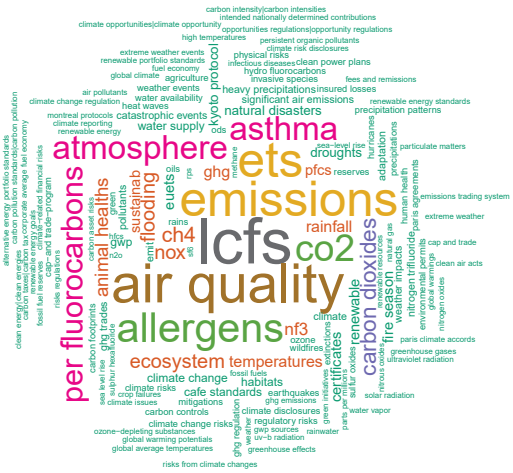



**Notes:** This figure is a demonstration of visualising the word embeddings of similar words in a 2-dimensional space. The word embeddings are generated by word2vec (Mikolov et al., 2013a). We use 200 neurons in the hidden layer of the Skip-gram's neural network model, which means we have a dense vector with size 200 to represent each unigrams in the corpus. We then use t-SNE (Van der Maaten and Hinton, 2008) to get the low-dimensional projections. The dots shown in the figure are the 5 most similar words of words "temperature", "emission", "agriculture" and "renewable" based on the Skip-gram model. We can visualise the clustering of similar words of the same unigram.

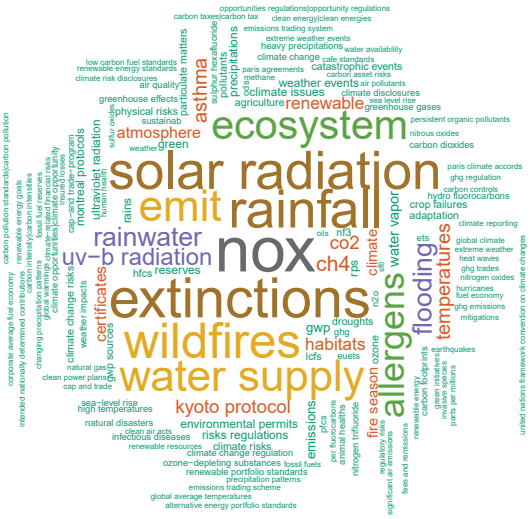
Figure 3.2: Term Frequency of Initial Unigrams in IPCC Reports



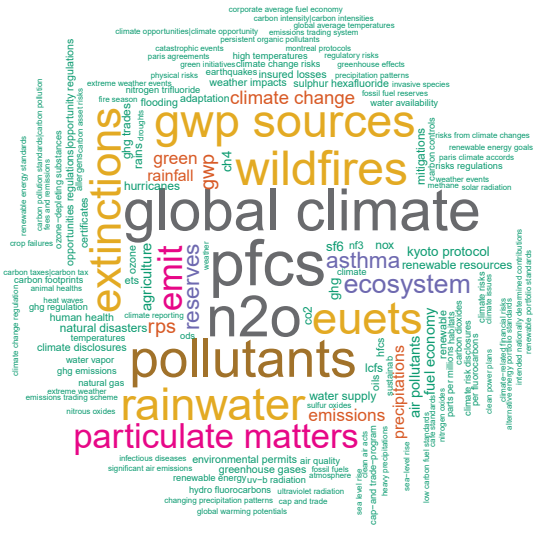

**Notes:** This figure visualises the term frequency of initial unigrams in several IPCC reports (listed in Table A1). Note that since we removed all digital numbers during the data cleaning, "co" represents "CO2" (carbon dioxide), "ch" represents "CH4" (methane), "n" represents "N2O" (nitrous oxide) and "sf" stands for "SF6" (sulfur hexafluoride).



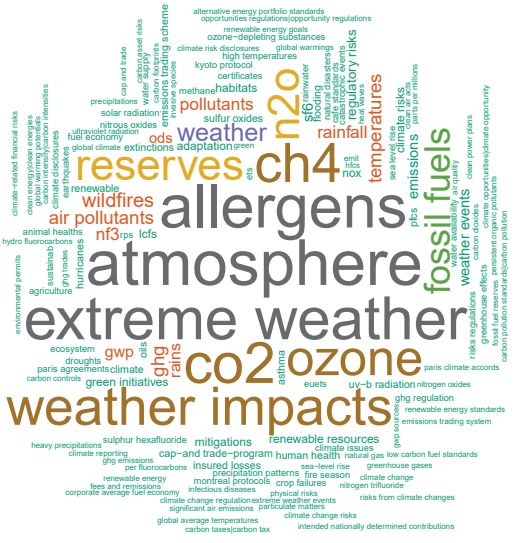

(a) Word Cloud of Term Frequency in 2009




(b) Word Cloud of Term Frequency in 2010



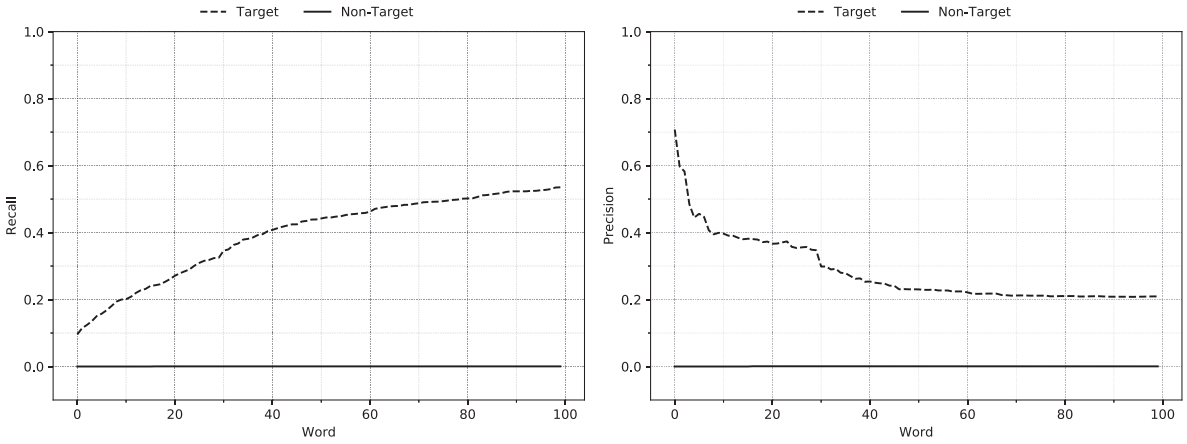

(c) Word Cloud of Term Frequency in 2011




(d) Word Cloud of Term Frequency in 2012

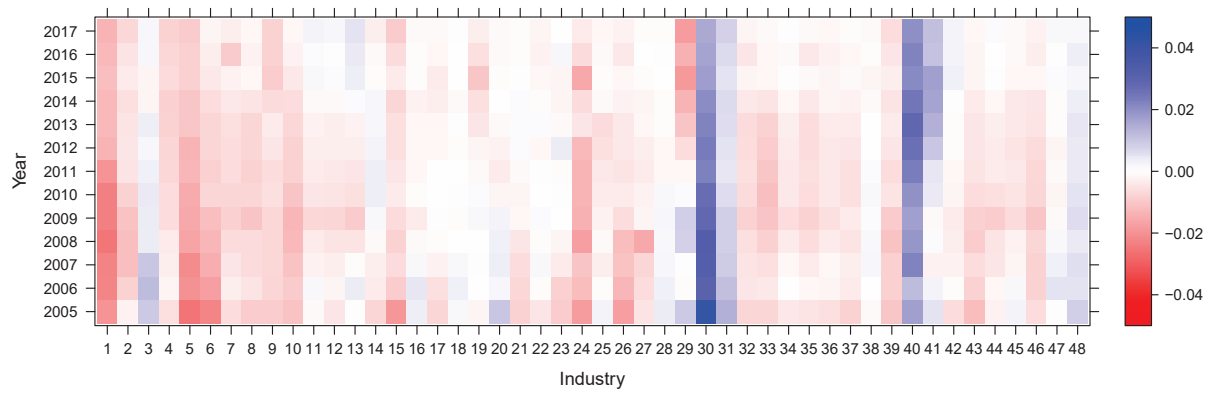
Figure 3.3: Word Cloud of the Aggregate Document-Term Matrix from 2009 to 2012

Figure 3.4: Cumulative Recall and Precision



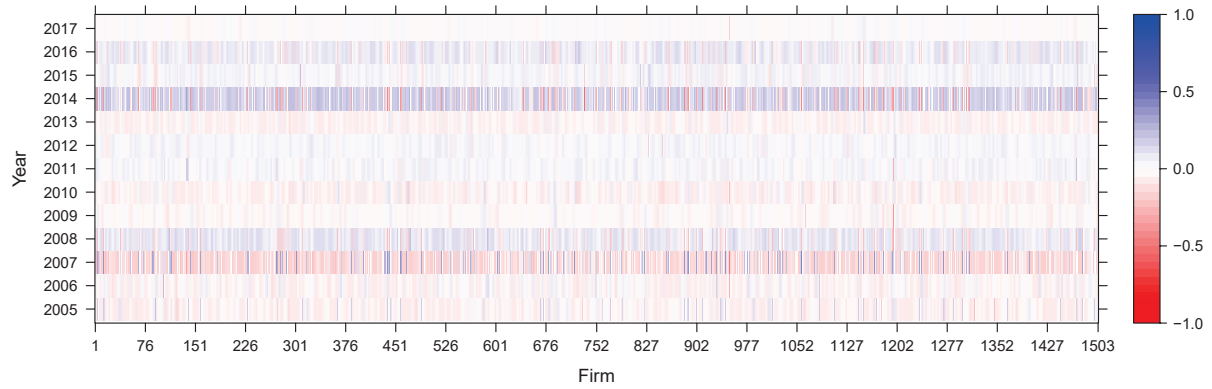

**Notes:** This figure reports precision, the proportion of retrieved documents from each keyword that contains documents of interest, and recall, the proportion of all documents of interest that are retrieved by the keyword. Note that the non-target set lines are not zero, but are close to zero.

Figure 3.5: Projection Heatmap at Industry-Level




**Notes:** We visualise the projections at industry-level from 2005 to 2017. We reorganise firms into the Kenneth French 48 industry groupings based on their 4-digit SIC codes.

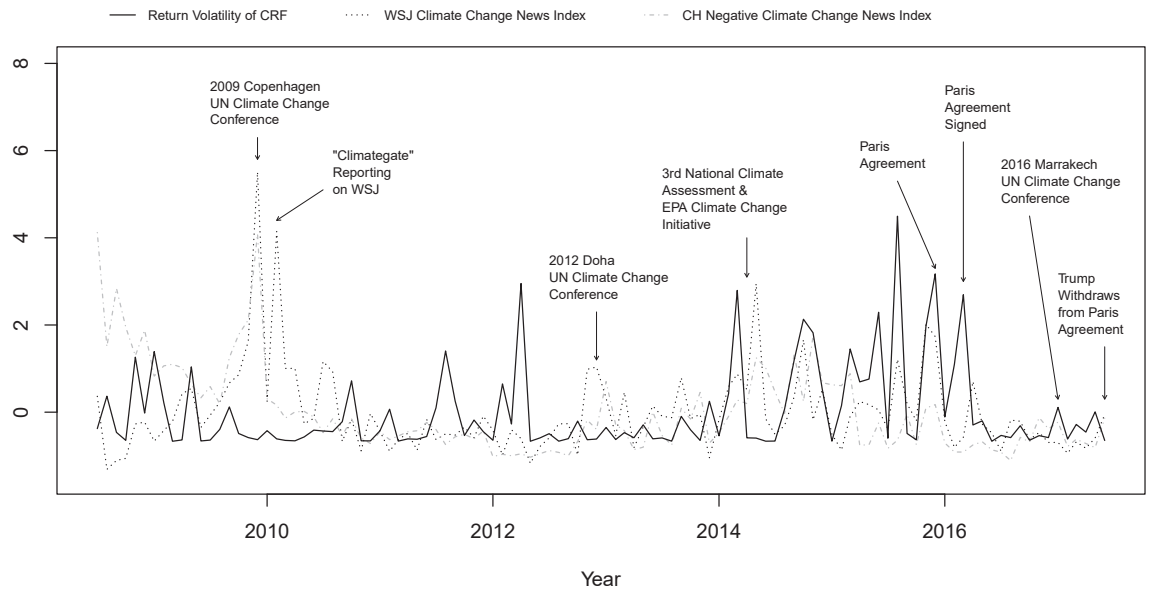
Figure 3.6: Projection Heatmap of Firms



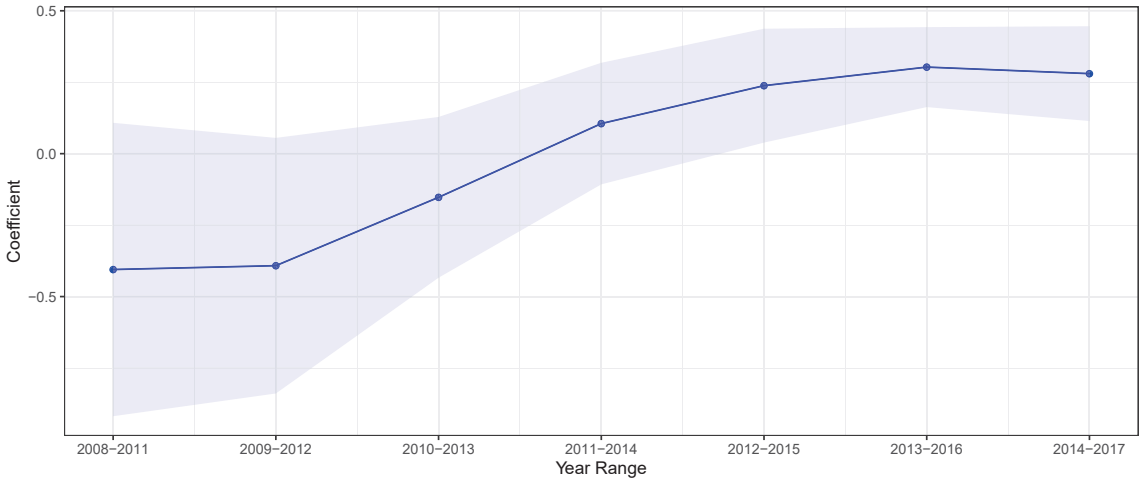

**Notes:** We visualise the projections of firms over different years. Note that these are fiscal years ending at the end of June. We only show projections of those firms having returns in all 13 years (i.e. stocks not having full record in all 13 years are omitted).

Figure 3.7: Return Volatility of the Factor Mimicking Portfolio

(a) Return Volatility of CRF

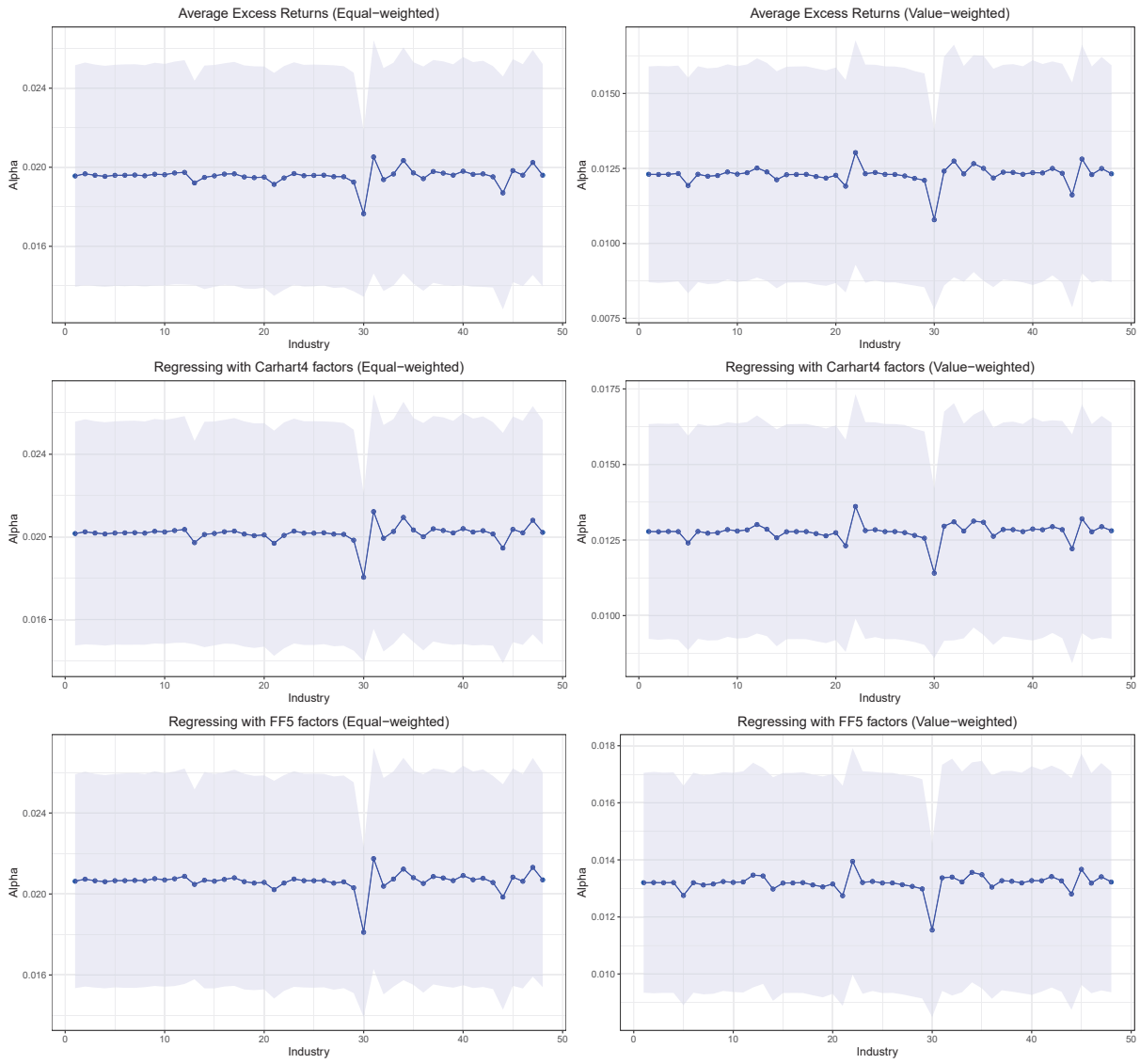



(b) Regressing WSJ Climate Change News Index on Portfolio Returns based on CRS




**Notes:** Panel (a) visualises the return volatility of the factor mimicking portfolio CRF, the WSJ Climate Change News Index and the CH Negative Climate Change News Index from 2008 to 2017 from (Engle et al., 2020b), annotated with climate-relevant news announcements. Panel (b) visualises the results of regressing WSJ Climate Change News Index on CRF. We ran rolling regressions on the 4-year window. The shaded area represents the 90% confidence interval.

Figure 3.8: Climate Risk Portfolios - Robustness Tests (Industry Excl.)




**Notes:** This figure visualises average excess returns, 4-factor alphas from Carhart four-factor model (Carhart, 1997) and 5-factor alphas from the FF5 model (Fama and French, 2015) for the spread strategy that is long the portfolio with the highest climate risk stocks and short the portfolio with the lowest climate risk stocks. We repeat the analysis 48 times after excluding each of the Kenneth French 48 industries in turn to flush out abnormal impact of any particular industry group. We reorganise firms into the Kenneth French 48 industry groupings based on their 4-digit SIC codes. The standard errors are corrected for the Newey-West heteroskedasticity and autocorrelation. The shaded area represents the 95% confidence interval.

## 3.7 Derivation

### 3.7.1 Proof of Proposition 1

**Proof of Proposition 1.** The core idea follows the (Taddy, 2013) where the observed words counts' likelihood that appears in the documents is as follows,

$$\begin{aligned} L_y^i &= \binom{m_y^i}{\mathbf{x}_y^i} \exp\left[\left(\mathbf{x}_y^i\right)' \eta_y^i - A\left(\eta_y^i\right)\right] \\ &= \binom{m_y^i}{\mathbf{x}_y^i} \exp\left(\left(\mathbf{x}_y^i\right)' \phi\right) \exp\left[\left(\left(\mathbf{x}_y^i\right)' \mathbf{\Phi}\right) r_{y+1}^i - A\left(\eta_y^i\right)\right], \end{aligned}$$

where,

$$A\left(\eta_y^i\right) = m_y^i \log\left[\sum_{l=1}^{D} \exp\left(\eta_{yl}^i\right)\right], \quad \mathbf{\Phi}' = [\gamma_1, \ldots, \gamma_D].$$

Take the logarithm of the likelihood,

$$\begin{aligned} l_y^i &= \log\left(m_y^i!\right) + \sum_{j=1}^{D} \log\left(\frac{\left(q_{yj}^i\right)^{x_{yj}^i}}{x_{yj}^i!}\right) \\ &= \log\left(m_y^i!\right) - \sum_{j=1}^{D} \log\left(x_{yj}^i!\right) + \sum_{j=1}^{D} x_{yj}^i \phi_j + \sum_{j=1}^{D} x_{yj}^i \gamma_j r_{y+1}^i - m_y^i \log \sum_{l=1}^{D} \exp\left(\eta_{yl}^i\right). \end{aligned}$$

We define the terms:

$$\begin{aligned} h\left(\mathbf{x}_y^i, m_y^i\right) &\equiv \log\left(m_y^i!\right) - \sum_{j=1}^{D} \log\left(x_{yj}^i!\right) + \sum_{j=1}^{D} x_{yj}^i \phi_j \\ g\left(\mathbf{\Phi}' \mathbf{x}_y^i, m_y^i, r_{y+1}^i\right) &\equiv \sum_{j=1}^{D} x_{yj}^i \gamma_j r_{y+1}^i - m_y^i \log \sum_{l=1}^{D} \exp\left(\eta_{yl}^i\right). \end{aligned}$$

Therefore, the log-likelihood can be represented as,

$$l_y^i = h\left(\mathbf{x}_y^i, m_y^i\right) + g\left(\mathbf{\Phi}' \mathbf{x}_y^i, m_y^i, r_{y+1}^i\right).$$

Hence, according to the usual sufficiency factorisation and follow (Taddy, 2013), (Schervish, 1995, 2.21),

$$P\left(\mathbf{x}_y^i \,|\, \mathbf{\Phi}' \mathbf{x}_y^i, m_y^i, r_{y+1}^i\right) = P\left(\mathbf{x}_y^i \,|\, \mathbf{\Phi}' \mathbf{x}_y^i, m_y^i\right),$$

and $r_{y+1}^i$ is independent of $\mathbf{x}_y^i$ given $\mathbf{\Phi}'\mathbf{x}_y^i$, meaning that,

$$P\left(r_{y+1}^i \,|\, \mathbf{x}_y^i, \mathbf{\Phi}'\mathbf{x}_y^i, m_y^i\right) = P\left(r_{y+1}^i \,|\, \mathbf{\Phi}'\mathbf{x}_y^i, m_y^i\right).$$

Under the conditions of Theorem 6.3 in Lehmann and Scheffé (1950), there exists a minimal sufficient statistic,

$$G\left(\mathbf{f}_y^i\right) = T\left(\mathbf{x}_y^i\right) = \tilde{G}\left(\mathbf{\Phi}'\mathbf{f}_y^i, m_y^i\right).$$

Since $m_y^i$ cannot be recovered from $\mathbf{f}_y^i$, it is not possible for $\tilde{G}\left(\mathbf{\Phi}'\mathbf{f}_y^i, m_y^i\right)$ to vary with $m_y^i$, while $G\left(\mathbf{f}_y^i\right)$ does not. Therefore, it must be the case that $\tilde{G}\left(\mathbf{\Phi}'\mathbf{f}_y^i, m_y^i\right) = \tilde{G}\left(\mathbf{\Phi}'\mathbf{f}_y^i\right)$. Hence, $\mathbf{\Phi}'\mathbf{f}_y^i$ is a sufficient statistic for $r_{y+1}^i$, meaning,

$$P\left(r_{y+1}^i \,|\, \mathbf{x}_y^i, \mathbf{\Phi}'\mathbf{f}_y^i\right) = P\left(r_{y+1}^i \,|\, \mathbf{\Phi}'\mathbf{f}_y^i\right),$$

which equivalents to,

$$r_{y+1}^i \perp \mathbf{x}_y^i \,|\, \mathbf{\Phi}'\mathbf{f}_y^i.$$

This completes the proof that $\mathbf{\Phi}'\mathbf{f}_y^i$ is a sufficient statistic for $r_{y+1}^i$. ■

## 3.8 Dictionary Construction Details

We provide details about the procedures to construct the extended unigrams and bigrams, respectively.

### 3.8.1 Discovery of Climate Change Unigrams

We create a set of unigrams $\mathbb{C}^U$ from the union of two separate unigrams: i) an initial set of climate-related unigrams $\mathbb{C}^{U_0}$ (reported in Table A2), which contains the unigrams of the vocabulary used by Chou and Kimbrough (2019b) [4], and ii) a set $\mathbb{C}^{U_S}$ created by finding similar words from the word embeddings of word2vec algorithms (Mikolov et al., 2013a,b).

A dense vector containing the estimated parameters of the word2vec models can be used to represent a word. Therefore, each unigram in our corpus $\mathbb{C}$ and the initial unigram set $\mathbb{C}^{U_0}$ can be represented as a dense vector. We then use cosine similarity to measure the similarities between words. For each unigram in $\mathbb{C}^{U_0}$, we put the 5 most similar words to the set $\mathbb{C}^{U_S}$. For example, we have the word "rainwater" in the initial set. We then put "irrigation", "harvesting", "dam", "catchment" and "irrigate", which are the 5 most similar words to "rainwater" based on the cosine similarity, into $\mathbb{C}^{U_S}$. The most similar unigrams are reported in Table A2. Note that on the word2vec training stage, we used the following hyperparameters: i) Dimensionality of the feature vector is 200 (i.e. setting the neuron amount of the hidden layer as 200), and ii) Ignores all words with the total frequency lower than 30 per corpus. We also put a demonstration to visualise the clusters of similar unigrams in Figure 3.1.

**Creating a final climate change unigrams library** We define the final vocabulary of unigrams $\mathbb{C}^U$ as $\mathbb{C}^U = \mathbb{C}^{U_0} \bigcup \mathbb{C}^{U_S}$. In this case, we can find more words that are related to the initial unigrams and climate change. For example, "CO2" is a similar word to "carbon" since "CO2" (i.e. carbon dioxide) contains the element "carbon". Or "coal" and "fuel" are similar words of "oil" given both "coal" and "oil" are "fossil fuels".

### 3.8.2 Discovery of Climate Change Bigrams

Bigrams are also considered to be included in our vocabulary. We create a set of bigrams $\mathbb{C}^B$ from the union of two separate bigrams: i) an initial set of climate-related bigrams $\mathbb{C}^{B_0}$, and ii)

[4] We visualise the term frequency of this initial set of unigrams in IPCC reports in Figure 3.2. We might have skipped some terms in the initial set due to the low frequency. During the training process of the word2vec model, we ignore all words with the total frequency lower than 30 per corpus.

a set $\mathbb{C}^{Bs}$ created by the machine learning keyword discovery algorithm proposed by King et al. (2017), which constructs potential climate-related bigrams from the IPCC reports. We believe the synthesis reports and some special reports from the IPCC contain enough information and conversation about global climate change. The algorithm requires a small set of initial bigrams $\mathbb{C}^{B_0}$, which are listed in Table A3. These initial bigrams are defined by Sautner et al. (2020) since they are unambiguously related to climate change.

**Defining the reference set and the search set** Sentences in the corpus $\mathbb{C}$ including the initial set of bigrams $\mathbb{C}^{B_0}$ are identified as the reference set, which unambiguously contains information about climate change. The remaining sentences that do not include the initial bigrams are identified as the search set. The search set may mention climate-related topics not captured by the initial bigrams, which means we may find potential climate-related bigrams other than the initial bigrams (i.e. discovering new keywords).

**Partitioning the search set** The algorithm depends on several supervised learning models and can extract features that predict climate change from the identified sentences of interest. 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 $\mathbb{S}$ (with the size around $1/3$ of $\mathbb{S}$). We fit several machine learning classifiers with the training set, which are Multinomial Naive Bayes, Support Vector Machine, logistic regression and neural network. Training the models can predict whether or not a sentence is related to climate change. We then apply the trained models to the search set and group sentences into a target set $\mathbb{T}$ if at least $3/4$ models predict the sentences are climate-related. Therefore, we have a search set $\mathbb{S}$ not containing bigrams in $\mathbb{C}^{B_0}$. And $\mathbb{S}$ can be partitioned by a target set $\mathbb{T}$ (not containing any bigrams in $\mathbb{C}^{B_0}$, but are likely to mention climate-related contents or topics not captured by $\mathbb{C}^{B_0}$) and a non-target set $\mathbb{S} \setminus \mathbb{T}$.
We report the cumulative recall and precision for the first 100 keywords (ranked by likelihood) in each of the target and non-target lists in Figure 3.4. We introduce one keyword at a time from left to right. We can visualise that the target set line is well above the non-target set line in both figures, showing good performances of separating the search set into target set and non-target set. We can also visualise a downward trend of the cumulative precision, which validates the ordering of the keywords (with more precise words near the top of the list).

**Finding climate change bigrams** To discover "new" bigrams, we reverse-engineer the machine learning process and trace back the bigrams that best discriminate climate-change-related from the non-initial sentences. For each of the bigrams in $\mathbb{S}$, we compute the document frequencies of this bigram in $\mathbb{T}$ and $\mathbb{S} \setminus \mathbb{T}$ respectively. If the bigram appears more frequently in the target set than in the non-target set, this bigram will be identified as a keyword in $\mathbb{T}$. For example, we have a bigram $b$, 10 sentences in total in the target set $\mathbb{T}$ and 100 sentences

in total in the non-target set $\mathbb{S} \setminus \mathbb{T}$. If $b$ appears in 3 out of 10 sentences in $\mathbb{T}$, the frequency of $b$ in $\mathbb{T}$ is 0.3. Similarly, if $b$ appears in 10 out of 100 sentences in $\mathbb{S} \setminus \mathbb{T}$, the frequency of $b$ in $\mathbb{S} \setminus \mathbb{T}$ is 0.1. In this case, the bigram appears more frequently in $\mathbb{T}$. We then follow King et al. (2017) to compute the likelihood metric of each bigram. We rank these bigrams via the likelihood and take the first 100 into the set $\mathbb{C}^{B_S}$. We report the Top-100 bigrams in Table A4. Note that we use a log-gamma function instead of a gamma function to deal with the large size of the search set.

**Creating a final climate change bigrams library** We define the final vocabulary of bigrams $\mathbb{C}^B$ as $\mathbb{C}^B = \mathbb{C}^{B_0} \bigcup \mathbb{C}^{B_S}$. In this case, we can capture more climate-related contents or topics that have not been included in $\mathbb{C}^{B_0}$ and expand the initial set of climate-related bigrams. We summarise the algorithm in Algorithm 2.

### 3.8.3 Integrated Dictionary

**Creating a final climate change library** We can create a new vocabulary by taking the union of the defined vocabulary of unigrams $\mathbb{C}^U$ and the vocabulary of bigrams $\mathbb{C}^B$, which is the expanded version (created via several machine learning techniques) of the initial sets of unigrams and bigrams. Our final vocabulary also contains the terms raised by Chou and Kimbrough (2019b) (i.e. $\mathbb{C}^U \bigcup \mathbb{C}^B \bigcup \mathbb{C}^{\text{Chou}}$).

## 3.9 Extra Experimental Results

### 3.9.1 Variable Definitions

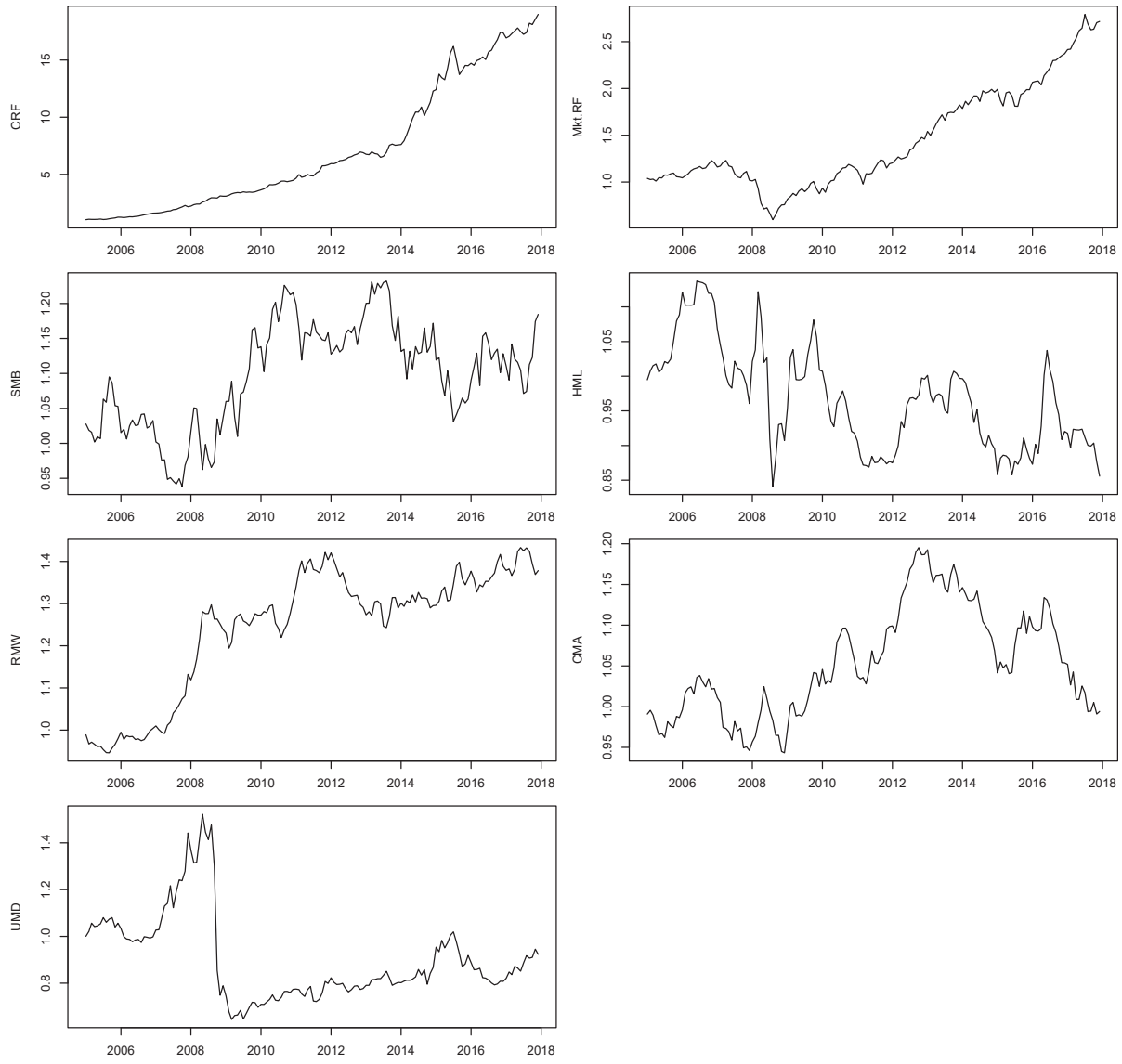
Table A13: Variable Definitions

This table provides definitions of key variables used in our analysis.

| Variable | Description | Source |
|---|---|---|
| Book-to-market | Ratio of book value of equity to market equity | Compustat/CRSP |
| Climate Risk Index | The sufficient projection of the document-term matrix by MNIR | 10-K |
| Leverage | The book value of debt (short-term + long-term) divided by the book value of assets | Compustat |
| Market Value | Natural logarithm of market value | CRSP |
| Sales | Sales divided by the net turnover | Compustat |
| Sales Growth | Sales growth rate ($\text{Sales}_t$ / $\text{Sales}_{t-1} - 1$) | Compustat |

### 3.9.2 Accumulated Returns of Different Factor-Mimicking Portfolios

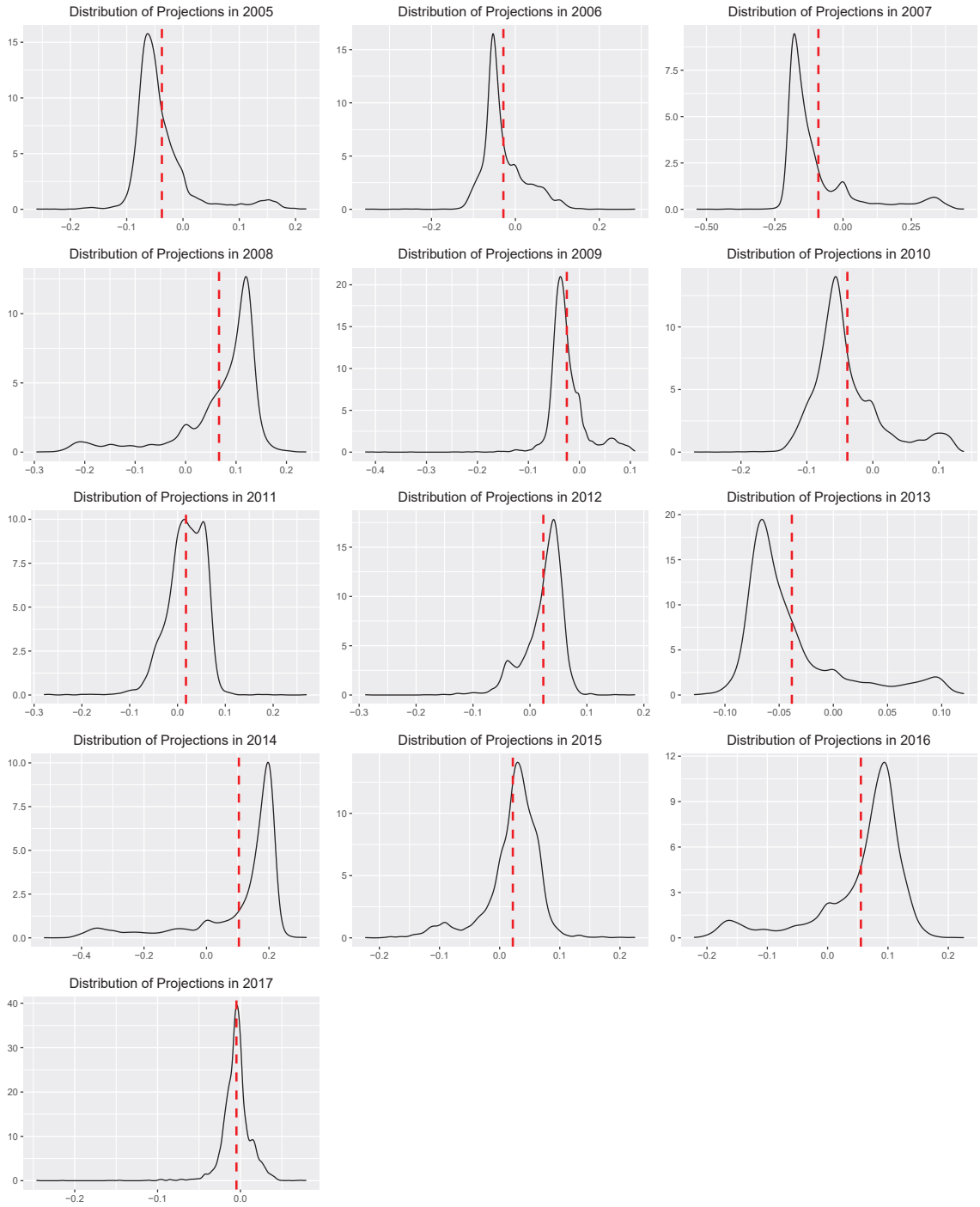
Figure 3.9: Accumulated Returns of Different Factor-Mimicking Portfolios




**Notes:** These figures show the accumulated monthly excess returns of different factor-mimicking portfolios, covering the period 2005 to 2017. *MKT* is the market risk factor. *SMB*, *HML*, *RMW* and *CMA* are the mimicking portfolios for the size risk, book-to-market risk, profitability risk and investment risk respectively in Fama and French (2015). *UMD* is a momentum factor introduced by Carhart (1997). *CRF* is the long-short portfolio based on CRS.

### 3.9.3 Cross-sectional Distribution of Projections

Figure 3.10: Cross-sectional Distribution of Projections




**Notes:** This figure visualises the distribution density of each year's projections across firms. The vertical dashed red line shows the mean of the projections.

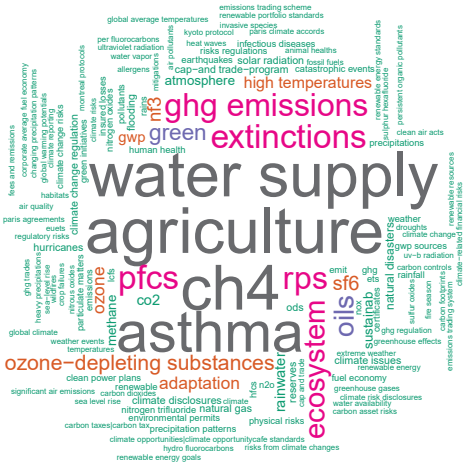
### 3.9.4 Climate Change Vocabulary [5]

quota, wave energy, natural disasters, climate induce, renewable portfolio standards, option, ghg emissions, nitrogen oxides, climatic change, carbon dioxide, typhoon, air pollutant, allergens, climate sensitivity, sulphur hexafluoride, renewable, air quality, air pollutants, ghg trades, sea water, solar thermal, gwp, human health, carbon intensity|carbon intensities, climate opportunities|climate opportunity, carbon emission, atmospheric, relate change, technical potential, absorb, energy efficient, energy conversion, euets, variability change, change land, precipitation patterns, greenhouse gases, coral reef, environmental change, impact, climate risk, ecosystems, aerosol, emissions trading system, monsoon, cafe standards, invasion, carbon tax, carbon, impact climate, renewable energy, pre-industrial levels of carbon dioxides, weather impacts, green initiatives, carbon asset risks, per fluorocarbons, use change, ozone-depleting substances, future climate, cost energy, paris agreements, high level, greenhouse gas, climate system, flue gas, habitat, greenhouse warm, nitrogen trifluoride, energy, extreme weather events, warming, electricity generation, concentration, fuel, emission scenario, time scale, regulation or disclosure of gh emissions, emissions, mitigations, clean water, air, fossil fuels, snow ice, renewable energy goals, stress, animal healths, carbon cycle, adaptation mitigation, lcfs, catastrophic events, uv-b radiation, kyoto protocol, crude, climate impact, rainfall, land use, climate resilient, air pollution, mitigation adaptation, pollutants, warm, physical risks, drought, emissions trading scheme, global regional, event, change regional, clean energy|clean energies, extinctions, hurricanes, solar energy, climate adaptation, global average temperatures, costal region, weather climate, energy use, new energy, wind power, sustainab, energy demand, climate change, carbon dioxides, rps, clean energy, natural system, risks from climate changes, sink, ozone, zero, electric vehicle, carbon stock, rains, carbon taxes|carbon tax, certificates, landslide, climate climate, land cover, wetland, spot, hydro fluorocarbons, energy supply, particulate, alternative energy portfolio standards, large scale, climate extreme, wind farm, flood, global decarbonization, montreal protocols, global environmental, ets, terrestrial, regional climate, irrigation, climate risks, emit, industrial level, adaptation strategy, atmosphere, reduce emission, nuclear, coal, intended nationally determined contributions, climate issues, habitats, low carbon fuel standards, enact, rainwater, renewable energy standards, clean air acts, global change, cyclone, frequency, clean air, impact change, conventional, water vapor, climate disclosures, catchment, carbon controls, harvesting, crop failures, storm, urban, carbon energy, green, renewable resource, wind energy, sf6, runoff, intergovernmental panel on climate changes, food, mitigation, united nations framework convention on climate changes, cross box, wind turbine, roof, sea-level rise, fossil fuel reserves, energy technology, climate mitigation, greenhouse effects, atmospheric concentration, corporate average fuel economy, weather events, earthquakes, changing precip-

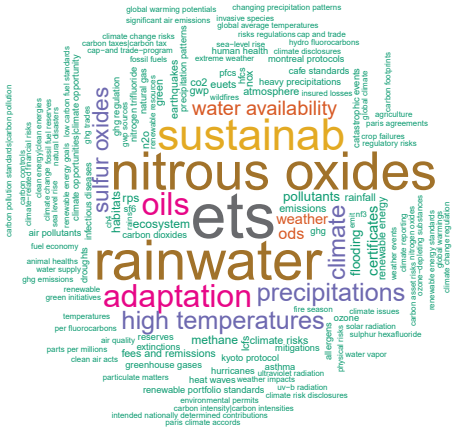
[5] Terms in the list were converted to a regular expression in order to capture encompassing patterns and be recognised when extending over more than one line, the multi-term expressions were coded. Also, the capitalisation is ignored when we conduct the searching.

itation patterns, wildfires, environmental sustainability, food security, precipitation, dioxide, cap and trade, adapt, earth, greenhouse, nf3, global warm, meteorological, invasive species, induce change, smallholder, energy climate, reserves, change temperature, sulfur oxides, temperatures, regulatory risks, solar power, biodiversity, dam, climate change regulation, ch4, ocean energy, island, global energy, ods, oxide, power generation, ecosystem, cover change, insured losses, human, pfcs, mean, maria, soil carbon, surface, high temperatures, heat waves, change climate, ultraviolet radiation, agriculture, weather, biomass energy, change ocean, fire season, natural hazard, renewable resources, strategy, agricultural, limit warm, ozone layer, low carbon, water availability, carbon pollution standards|carbon pollution, water supply, surface temperature, wind resource, gas emission, precursor, policy, climate land, future change, emission reduction, nitrous oxides, forest land, oh, sea level rise, environmental permits, energy environment, droughts, surge, climate reporting, richness, brown, street, sea ice, temperature change, energy consumption, global warmings, climate, endemic, ghg emission, climate projection, change, change change, climate warming, sustainable development, irrigate, persistent organic pollutants, global temperature, sea level, paris climate accords, extreme, sustainable energy, climate-related financial risks, ethanol, hfcs, climate change risks, climate policy, opportunities regulations|opportunity regulations, technology, clean, heat power, release, incentive, temperature, extreme weather, carbon footprints, renewable source, offshore wind, emission, unique, climatic, particulate matters, palm, carbon capture, methane, global climate, fossil fuel, warm climate, natural gas, conference of the parties|conference of the party, enforcement, specie, mammal, parts per millions, water resource, electricity, infectious diseases, precipitations, fossil, oil, energy service, pollution, fit, n2o, wind solar, tropical, ghg, management, climate variability, natural gas, power plant, depletion, carbon neutral, carbon price, project change, solar radiation, water quality, cap-and trade-program, air temperature, clean power plans, oxidation, asthma, climate relate, climate risk disclosures, energy system, stratosphere, global warming potentials, flooding, fees and remissions, gas, sea surface, heavy precipitations, adaptive, corridor, forestry, adaptation, gwp sources, greenhouse effect, deployment, system change, nox, risks regulations, co2, resilience, significant air emissions, carbon sink, pm, change global, ghg regulation, climate response, stratospheric, oils, energy source, fuel economy

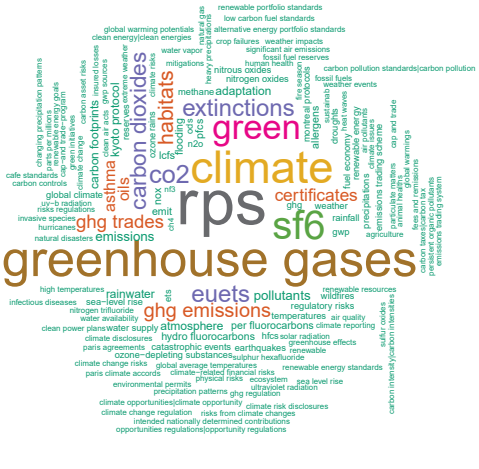
### 3.9.5 Visualisation of the Aggregate DTM



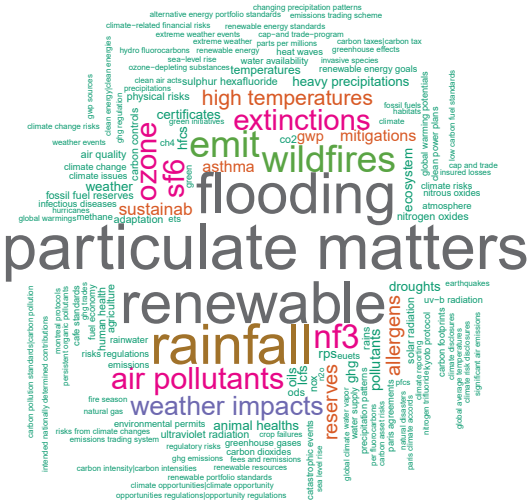

(a) Word Cloud of Term Frequency in 2005




(b) Word Cloud of Term Frequency in 2006

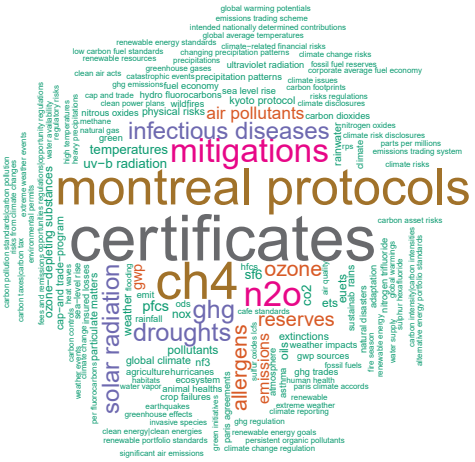



(c) Word Cloud of Term Frequency in 2007




(d) Word Cloud of Term Frequency in 2008

Figure 3.11: Word Cloud of the Aggregate Document-Term Matrix from 2005 to 2008




(a) Word Cloud of Term Frequency in 2013

(b) Word Cloud of Term Frequency in 2014

(c) Word Cloud of Term Frequency in 2015

(d) Word Cloud of Term Frequency in 2016

Figure 3.12: Word Cloud of the Aggregate Document-Term Matrix from 2013 to 2016

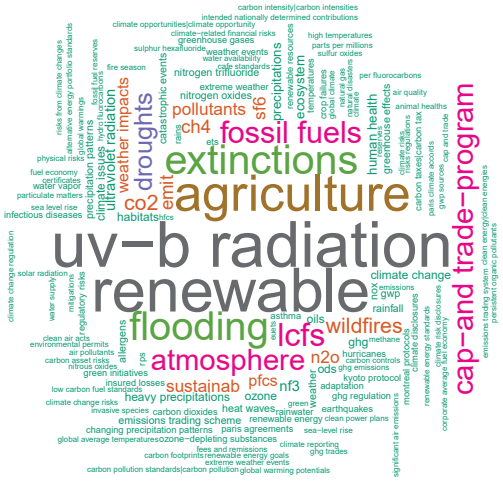



Figure 3.13: Word Cloud of the Aggregate Document-Term Matrix in 2017

# Bibliography


Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012a). The environment and directed technical change. *American economic review*, 102(1):131–66.

Acemoglu, D., Aghion, P., Bursztyn, L., and Hemous, D. (2012b). The environment and directed technical change. *American Economic Review*, 102(1):131–66.

Addoum, J. M., Ng, D. T., and Ortiz-Bobea, A. (2020). Temperature shocks and establishment sales. *The Review of Financial Studies*, 33(3):1331–1366.

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. (2020). Climate change concerns and the performance of green versus brown stocks. *National Bank of Belgium, Working Paper Research*, (395).

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K. (2022). Climate change concerns and the performance of green vs. brown stocks. *Management Science*.

Baker, S. R., Bloom, N., and Davis, S. J. (2016a). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baker, S. R., Bloom, N., and Davis, S. J. (2016b). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Baldauf, M., Garlappi, L., and Yannelis, C. (2020). Does climate change affect real estate prices? only if you believe in it. *The Review of Financial Studies*, 33(3):1256–1295.

Bansal, R., Kiku, D., and Ochoa, M. (2021). Climate change risk. *Federal Reserve Bank of San Francisco Working Paper*.

Bansal, R., Ochoa, M., and Kiku, D. (2016). Climate change and growth risks. Working Paper 23009, National Bureau of Economic Research.

Bansal, R. and Shaliastovich, I. (2013). A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets. *The Review of Financial Studies*, 26(1):1–33.

Bansal, R. and Yaron, A. (2004a). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance*, 59(4):1481–1509.

Bansal, R. and Yaron, A. (2004b). Risks for the long run: A potential resolution of asset pricing puzzles. *The journal of Finance*, 59(4):1481–1509.

Barnett, M. (2017). Climate change and uncertainty: An asset pricing perspective. Technical report, Working Paper.

Barnett, M. (2020). A run on oil? The implications of climate policy and stranded assets risk. *ASU Working paper*.

Baz, S., Cathcart, L., Michaelides, A., and Zhang, Y. (2022a). Firm-level climate regulatory exposure. *Available at SSRN 3873886*.

Baz, S., Cathcart, L., Michaelides, A., and Zhang, Y. (2022b). Firm-level climate regulatory exposure. *Available at SSRN 3873886*.

Ben-Rephael, A., Da, Z., and Israelsen, R. D. (2017). It Depends on Where You Search: Institutional Investor Attention and Underreaction to News. *The Review of Financial Studies*, 30(9):3009–3047.

Berkman, H., Jona, J., and Soderstrom, N. S. (2019). Firm-Specific Climate Risk and Market Valuation. Available at https://ssrn.com/abstract=2775552.

Bernard, R., Tzamourani, P., and Weber, M. (2022). Climate change and individual behavior.

Bhattacharya, U. and Daouk, H. (2002). The World Price of Insider Trading. *The Journal of Finance*, 57(1):75–108.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3(Jan):993–1022.

Bolton, P., Després, M., da Silva, L. A. P., Samama, F., and Svartzman, R. (2020). *The Green Swan: Central banking and financial stability in the age of climate change*. Bank for International Settlements – Other Publications. Available at `https://www.bis.org/publ/othp31.pdf`.

Bolton, P. and Kacperczyk, M. (2020a). Do Investors Care about Carbon Risk? *Journal of Financial Economics*. Forthcoming.

Bolton, P. and Kacperczyk, M. (2020b). Global Pricing of Carbon-Transition Risk. Available at `https://ssrn.com/abstract=3550233orhttp://dx.doi.org/10.2139/ssrn.3550233`.

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

Bolton, P. and Kacperczyk, M. (2021b). Do investors care about carbon risk? *Journal of Financial Economics*.

Bolton, P. and Kacperczyk, M. (2021c). Global pricing of carbon-transition risk. Technical report, National Bureau of Economic Research.

Bornstein, G., Krusell, P., and Rebelo, S. (2022). A World Equilibrium Model of the Oil Market. *The Review of Economic Studies*. rdac019.

Bretscher, L., Hsu, A., and Tamoni, A. (2020a). The real response to uncertainty shocks: The risk premium channel. *Georgia Tech Scheller College of Business Research Paper*, (17-13).

Bretscher, L., Hsu, A., and Tamoni, A. (2020b). The real response to uncertainty shocks: The risk premium channel. *Georgia Tech Scheller College of Business Research Paper*, (17-13).

Bybee, L., Kelly, B. T., and Su, Y. (2022). Narrative asset pricing: Interpretable systematic risk factors from news text. *Johns Hopkins Carey Business School Research Paper*, (21-09).

Cai, Y. and Lontzek, T. S. (2019). The social cost of carbon with economic and climate risks. *Journal of Political Economy*, 127(6):2684–2734.

Campbell, J. Y. and Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of political Economy*, 107(2):205–251.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of Finance*, 52(1):57–82.

Child, T., Massoud, N., Schabus, M., and Zhou, Y. (2020). Surprise Election for Trump Connections. *Journal of Financial Economics*. Forthcoming.

Choi, D., Gao, Z., and Jiang, W. (2020). Attention to global warming. *The Review of Financial Studies*, 33(3):1112–1145.

Chou, C. and Kimbrough, S. O. (2019a). Talking about Climate Change: What Are Enterprises Saying in Their SEC filings? Available at https://ssrn.com/abstract=3509765.

Chou, C. and Kimbrough, S. O. (2019b). Talking about climate change: What are enterprises saying in their sec filings? *Available at SSRN 3509765*.

Colacito, R., Croce, M., Ho, S., and Howard, P. (2018). Bkk the ez way: International long-run growth news and capital flows. *American Economic Review*, 108(11):3416–49.

Cong, L., Liang, T., and Zhang, X. (2019). Textual factors: A scalable, interpretable, and data-driven approach to analyzing unstructured information. *Available at https://ssrn.com/abstract=3307057*.

Croce, M. M., Kung, H., Nguyen, T. T., and Schmid, L. (2012). Fiscal Policies and Asset Prices. *The Review of Financial Studies*, 25(9):2635–2672.

Drake, M. S., Roulstone, D. T., and Thornock, J. R. (2015). The Determinants and Consequences of Information Acquisition via EDGAR. *Contemporary Accounting Research*, 32(3):1128–1161.

Edenhofer, O., Pichs-Madruga, R., Sokona, Y., Seyboth, K., Kadner, S., Zwickel, T., Eickemeier, P., Hansen, G., Schlömer, S., von Stechow, C., et al. (2011). *Renewable energy sources and climate change mitigation: Special report of the intergovernmental panel on climate change*. Cambridge University Press.

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

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

Fama, E. F. and French, K. R. (1992). The Cross-Section of Expected Stock Returns. *Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1):1–22.

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, 81(3):607–636.

Feng, G., Giglio, S., and Xiu, D. (2020). Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370.

Florackis, C., Louca, C., Michaely, R., and Weber, M. (2022). Cybersecurity Risk. *The Review of Financial Studies*. hhac024.

Florackis, C., Louca, C., Michaely, R., and Weber, M. (2023). Cybersecurity risk. *The Review of Financial Studies*, 36(1):351–407.

Fowlie, M., Goulder, L., Kotchen, M., Borenstein, S., Bushnell, J., Davis, L., Greenstone, M., Kolstad, C., Knittel, C., Stavins, R., et al. (2014). An Economic Perspective on the EPA's Clean Power Plan. *Science*, 346(6211):815–816.

Fried, S. (2018a). Climate Policy and Innovation: A Quantitative Macroeconomic Analysis. *American Economic Journal: Macroeconomics*, 10(1):90–118.

Fried, S. (2018b). Climate policy and innovation: A quantitative macroeconomic analysis. *American Economic Journal: Macroeconomics*, 10(1):90–118.

Fried, S. (2018c). Climate Policy and Innovation: A Quantitative Macroeconomic Analysis. *American Economic Journal: Macroeconomics*, 10(1):90–118.

Fried, S., Novan, K., and Peterman, W. (2020). The Macro Effects of Climate Policy Uncertainty. Mimeo.

Gao, L., Hitzemann, S., Shaliastovich, I., and Xu, L. (2021). Oil volatility risk. *Journal of Financial Economics*.

Garcia, D., Hu, X., and Rohrer, M. (2023). The colour of finance words. *Journal of Financial Economics*, 147(3):525–549.

Gentzkow, M. and Shapiro, J. M. (2010). What drives media slant? evidence from us daily newspapers. *Econometrica*, 78(1):35–71.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019). Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica*, 87(4):1307–1340.

Gerarden, T. D., Reeder, W. S., and Stock, J. H. (2020). Federal coal program reform, the clean power plan, and the interaction of upstream and downstream climate policies. *American Economic Journal: Economic Policy*, 12(1):167–99.

Ghoneem, M. Y. M. (2016). Planning for climate change, why does it matter?(from phenomenon to integrative action plan). *Procedia-Social and Behavioral Sciences*, 216:675–688.

Giglio, S., Maggiori, M., Stroebel, J., and Weber, A. (2020). Climate Change and Long-Run Discount Rates: Evidence from Real Estate. *Review of Financial Studies*. Forthcoming.

Gillan, S. L., Koch, A., and Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, 66:1–16.

Goldberg, Y. and Levy, O. (2014). word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method. *arXiv preprint arXiv:1402.3722*.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

Görgen, M., Jacob, A., Nerlinger, M., Riordan, R., Rohleder, M., and Wilkens, M. (2019). Carbon Risk. Available at `https://ssrn.com/abstract=2930897`.

Hambel, C., Kraft, H., and der Ploeg, R. v. (2021). Asset diversification versus climate action. *Available at SSRN 3528239*.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Hassan, T. A., Hollander, S., van Lent, L., and Tahoun, A. (2019a). Firm-Level Political Risk: Measurement and Effects. *The Quarterly Journal of Economics*, 134(4):2135–2202.

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

Hassler, J., Krusell, P., and Olovsson, C. (2018). The consequences of uncertainty: Climate sensitivity and economic sensitivity to the climate. *Annual Review of Economics*, 10(1):189–205.

Hassler, J., Krusell, P., and Olovsson, C. (2021). Directed technical change as a response to natural resource scarcity. *Journal of Political Economy*, 129(11):3039–3072.

Hershbein, B. and Kahn, L. B. (2018). Do recessions accelerate routine-biased technological change? evidence from vacancy postings. *American Economic Review*, 108(7):1737–72.

Hirose, Y., Yamashita, K., and Hijiya, S. (1991). Back-propagation algorithm which varies the number of hidden units. *Neural Networks*, 4(1):61–66.

Hoberg, G. and Phillips, G. (2016). Text-based network industries and endogenous product differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Hong, H., Li, F. W., and Xu, J. (2019). Climate risks and market efficiency. *Journal of Econometrics*, 208(1):265–281.

Hsu, P.-H., Li, K., and Tsou, C.-Y. (2020). The Pollution Premium. *Available at https://ssrn.com/abstract=3578215*.

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

Iaria, A., Schwarz, C., and Waldinger, F. (2018). Frontier knowledge and scientific production: Evidence from the collapse of international science. *The Quarterly Journal of Economics*, 133(2):927–991.

Ilhan, E., Sautner, Z., and Vilkov, G. (2020). Carbon Tail Risk. *The Review of Financial Studies*. forthcoming.

IPCC (1992). *Climate Change 1992: Synthesis Report. Contribution of Working Groups I, II, and III to the First Assessment Report of the Inter-governmental Panel on Climate Change*. IPCC Geneva.

IPCC (1995). *Climate Change 1995: Synthesis Report. Contribution of Working Groups I, II, and III to the Second Assessment Report of the Inter-governmental Panel on Climate Change*. IPCC Geneva.

IPCC (2001a). *Climate Change 2001: Synthesis Report. A Contribution of Working Groups I, II, and III to the Third Assessment Report of the Integovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom, and New York, NY, USA.

IPCC (2001b). United nations framework convention on climate change (unfccc) consolidated version with amendments to annexes. The Panel Geneva.

IPCC (2007). *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC Geneva.

IPCC (2014). *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. IPCC Geneva.

Jaccard, I., Benmir, G., and Vermandel, G. (2020). Green asset pricing.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics*, 41(2):257–275.

Ke, Z. T., Kelly, B. T., and Xiu, D. (2019). Predicting returns with text data. Technical report, National Bureau of Economic Research.

King, G., Lam, P., and Roberts, M. E. (2017). Computer-assisted keyword and document set discovery from unstructured text. *American Journal of Political Science*, 61(4):971–988.

Koijen, R. S. J., Philipson, T. J., and Uhlig, H. (2016). Financial Health Economics. *Econometrica*, 84(1):195–242.

Krueger, P., Sautner, Z., and Starks, L. T. (2020). The Importance of Climate Risks for Institutional Investors. *The Review of Financial Studies*, 33(3):1067–1111.

Kölbel, J., Leippold, M., Rillaerts, J., and Wang (2020). Does the CDS Market Reflect Regulatory Climate Risk Disclosures? Available at `https://ssrn.com/abstract=3616324`.

Larsen, V. H. and Thorsrud, L. A. (2019). The value of news for economic developments. *Journal of Econometrics*, 210(1):203–218.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436.

Lehmann, E. L. and Scheffé, H. (1950). Completeness, similar regions, and unbiased estimation: Part i. *Sankhyā: The Indian Journal of Statistics (1933-1960)*, 10(4):305–340.

Li, Q., Shan, H., Tang, Y., and Yao, V. (2020). Corporate climate risk: Measurements and responses. *Available at SSRN 3508497*.

Liu, Y. and Matthies, B. (2022). Long-run risk: Is it there? *The Journal of Finance*, 77(3):1587–1633.

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

Loughran, T. and McDonald, B. (2017). The Use of EDGAR Filings by Investors. *Journal of Behavioral Finance*, 18(2):231–248.

Malmendier, U., Nagel, S., and Yan, Z. (2020). The Making of Hawks and Doves: Inflation Experiences on the FOMC. *Journal of Monetary Economics*. Forthcoming.

Massimiliano Croce, M. (2014). Long-run productivity risk: A new hope for production-based asset pricing? *Journal of Monetary Economics*, 66:13–31.

Mcauliffe, J. D. and Blei, D. M. (2008). Supervised topic models. In *Advances in neural information processing systems*, pages 121–128.

Meeuwis, M., Parker, J. A., Schoar, A., and Simester, D. I. (2018). Belief Disagreement and Portfolio Choice. Mimeo.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013a). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

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

Mueller, H. and Rauh, C. (2018). Reading between the lines: Prediction of political violence using newspaper text. *American Political Science Review*, 112(2):358–375.

Nordhaus, W. D. (1992). An optimal transition path for controlling greenhouse gases. *Science*, 258(5086):1315–1319.

Nordhaus, W. D. and Moffat, A. (2017). A survey of global impacts of climate change: Replication, survey methods, and a statistical analysis. Working Paper 23646, National Bureau of Economic Research.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2021a). Dissecting green returns. Technical report, National Bureau of Economic Research.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2021b). Dissecting green returns. Working Paper 28940, National Bureau of Economic Research.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. (2021c). Sustainable investing in equilibrium. *Journal of Financial Economics*, 142(2):550–571.

Pörtner, H.-O., Roberts, D. C., Masson-Delmotte, V., and Zhai, P. (2019). *The Ocean and Cryosphere in a Changing Climate: A Special Report by the Intergovernmental Panel on Climate Change*. Intergovernmental Panel on Climate Change.

Ramadorai, T. and Zeni, F. (2019). Climate regulation and emissions abatement: Theory and evidence from firms' disclosures. *Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3469787*.

Ramelli, S., Wagner, A. F., Zeckhauser, R. J., and Ziegler, A. (2018). Investor rewards to climate responsibility: Evidence from the 2016 climate policy shock. Working Paper 25310, National Bureau of Economic Research.

Ready, R. C. (2018). Oil consumption, economic growth, and oil futures: The impact of long-run oil supply uncertainty on asset prices. *Journal of Monetary Economics*, 94:1–26.

Rong, X. (2014). word2vec parameter learning explained. *arXiv preprint arXiv:1411.2738*.

Ryans, J. (2017). Using the EDGAR Log File Data Set. *Available at https://ssrn.com/abstract=2913612*.

Sautner, Z., van Lent, L., Vilkov, G., and Zhang, R. (2020). Firm-level climate change exposure.

Schervish, M. J. (1995). *Theory of Statistics*. Springer series in statistics. Springer-Verlag, New York.

Seltzer, L., Starks, L. T., and Zhu, Q. (2020). Climate regulatory risks and corporate bonds. *Nanyang Business School Research Paper*, (20-05).

Shanken, J. (2015). On the Estimation of Beta-Pricing Models. *The Review of Financial Studies*, 5(1):1–33.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442.

Shukla, P., Skea, J., Calvo Buendia, E., Masson-Delmotte, V., Pörtner, H., Roberts, D., Zhai, P., Slade, R., Connors, S., Van Diemen, R., et al. (2019). Ipcc, 2019: Climate change and land: an ipcc special report on climate change, desertification, land degradation, sustainable land management, food security, and greenhouse gas fluxes in terrestrial ecosystems.

Skjærseth, J. B. and Skodvin, T. (2018). *Climate change and the oil industry: Common problem, varying strategies*. Manchester University Press.

Swanson, E. T. (2018). Risk aversion, risk premia, and the labor margin with generalized recursive preferences. *Review of Economic Dynamics*, 28:290–321.

Taddy, M. (2013). Multinomial inverse regression for text analysis. *Journal of the American Statistical Association*, 108(503):755–770.

Tomar, S. (2022). Greenhouse gas disclosure and emissions benchmarking. *SMU Cox School of Business Research Paper*, (19-17).

UNFCCC (2005). Summary for policymakers: A report of working group i of the intergovernmental panel on climate change. UNFCCC Secretariat with support from UNOG. FCCC-Informal-103/ FCCC-Informal-140.

Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11).

Van der Ploeg, F. and Rezai, A. (2020). Stranded assets in the transition to a carbon-free economy. *Annual Review of Resource Economics*, 12:281–298.

Van Ooyen, A. and Nienhuis, B. (1992). Improving the convergence of the back-propagation algorithm. *Neural Networks*, 5(3):465–471.

Wagner, A. F., Zeckhauser, R. J., and Ziegler, A. (2018). Company Stock Price Reactions to the 2016 Election Shock: Trump, Taxes, and Trade. *Journal of Financial Economics*, 130(2):428 – 451.

Zhai, P., Pörtner, H., Roberts, D., Skea, J., Shukla, P., Pirani, A., Moufouma-Okia, W., Péan, C., Pidcock, R., Connors, S., et al. (2018). Global warming of 1.5○ c. *An IPCC Special Report on the impacts of global warming*.