The Carbon Premium and Policy Risk Exposure: A Text-Based Approach

Sarah Duffy*

June 2025

Shifts in climate policy stringency have heterogeneous effects on firms' profitability. Does the market price this risk? This paper provides new evidence on this question, utilising a supervised machine learning algorithm to construct a firm-level measure of climate policy risk exposure. Firms exposed to climate policy risk have negative abnormal returns on climate policy announcement days. I build a set of such dates and characterize abnormal return responses using *Risk Factors* discussions in 10-K filings. The algorithm uncovers predictors of policy risk exposure in the text which are used to construct an exposure score for each firm. This exposure score is correlated with emissions, environmental lobbying behaviour, and is predictive out of sample. Higher exposure is not associated with a premium. Green preference shifts are considered as a mechanism to rationalize this result. I find that empirically identified preference shocks can partly explain the lack of a climate policy risk premium.

*Department of Economics, University of Oxford. I thank Andrea Ferrero and Rick van der Ploeg for their guidance and support, and Stephen Hansen for kindly sharing the text data employed in this paper. I also thank participants at the University of Oxford internal seminar, HEC Paris PhD conference, EAERE annual conference, MMF annual conference, and the LSE Environment Day for their comments.

1. Introduction

Substantial reductions of carbon emissions are required in order to stay below the temperature limits laid out in international climate agreements.[1]. It is widely agreed that such emissions reductions require policy intervention, such as carbon taxes or emissions caps. Climate policy interventions have heterogeneous effects on the profitability of firms depending on their emissions, their ease of substituting to greener technologies, and their positioning in the network of firms who themselves have differential exposure. These factors constitute a risk for firms who will be exposed to these policies: climate policy risk.

Theory predicts that firms with higher exposure to climate policy risk pay higher returns to investors to compensate them for exposure to this risk. This is the climate policy risk premium. In this paper I test for the existence of such a premium. The (non)existence of a climate policy risk premium has important implications. If governments delay in implementing climate policy and this risk is not priced in by markets, a sudden policy shift could in the worst instance lead to fire sales with financial stability implications (Carney 2015).

Motivated by the disagreement in this empirical literature, most of which uses emissions data or ESG scores[2] to proxy for risk exposure (Bolton and Kacperczyk 2021, 2023; Bauer et al. 2022), I build a novel firm-level measure of exposure to climate policy risk utilising equity price movements on climate policy announcement dates and text data from firms' form 10-K. Firm-level emissions data has many limitations: release lags, cross-sectional incompleteness, and instability across vendors and time. Moreover, emissions are an incomplete measure of exposure to climate policy risk. A given firm may have low emissions but may be very vulnerable to climate policy risks due to reliance on inputs from high emissions firms or high emissions associated with the consumption of their products. Exposure to climate policy risk will also be related to how easily a firm can abate their emissions. 'Scope 3' emissions data present a potential solution to the first of these problems by including indirect emissions up and down the

[1]The UNEP finds that the world is heading for a 2.5-2.9°C temperature rise unless countries deliver more than promised under the Paris Agreement (UNEP 2023). The IPCC states that global warming of 2°C will be exceeded during the 21st century unless deep reductions in CO2 emissions occur in the coming decades (IPCC 2021).

[2]Environmental, social, and governance (ESG) investing is an investing approach that prioritizes environmental issues, social issues, and corporate governance. To aid investors in this approach, vendors of financial data have begun to construct and disseminate ESG scores for listed firms. These scores are designed to summarize the performance of firms along environmental, social, and governance dimensions which are not as easily captured with traditional financial indicators.

value chain as well as direct emissions from production. However, scope 3 emissions are reported only for a small subset of firms and are likely measured with a great deal of error as there is ongoing disagreement on what exactly should (not) be included (Bauer et al. 2022). Therefore, rather than relying on emissions data I build a market implied measure of risk using text data from corporate reporting.

I proceed in five steps. First, I outline a simple asset pricing model to motivate the empirical approach. In the model, higher exposure to climate policy risk is associated with i) a larger fall in one period returns in response to the announcement of an increase in climate policy stringency, and ii) a climate policy risk premium in equilibrium. Therefore, I collate a list of salient and unexpected climate policy announcements which are informative for exposure of US-listed firms. I list climate policy announcements by US Presidents, Congressional bill passages related to climate, Supreme Court decisions on climate related legislation, agreements reached at UN Climate Conferences, and major IPCC report releases. Out of these dates I select those that i) are covered in either the New York Times or Wall Street Journal, ii) are associated with detectable movements in stock prices, and iii) aren't confounded by other major non-climate news. Examples of announcement dates included are the signing of the Paris agreement and the passage of the Inflation Reduction Act through the senate. Examples of announcements which are excluded are Executive Orders on climate which are not covered in the news, and Inflation Reduction Act related events which occurred as COVID lockdown measures were also announced.

Second, I use text data from *Risk Factor* discussions in firms' form 10-K to understand in natural language the drivers of firms' exposure to climate policy risk, and to build a firm -level measure of exposure to that risk. In these Risk Factor discussions, firms describe the risks that they are exposed to. The SEC advises that firms should include any item that could impact future earnings. Firms can be sued for omitting material information on risks. Following Davis, Hansen, and Seminario-Amez (2021), the idea is that the language firms use to describe their risks is predictive of their financial performance when these risks materialize. I use this text to predict abnormal returns on the climate policy announcement dates and find that terms associated with fossil fuels are most predictive of negative abnormal return responses, and terms associated with cleaner activities are most predictive of positive abnormal returns.

To implement this procedure I adopt the multinomial inverse regression (MNIR) method introduced by Taddy (2013) and applied by Gentzkow, Shapiro, and Taddy (2019) and Davis, Hansen, and Seminario-Amez (2021). MNIR represents a methodological improvement in this literature as it considers all terms in the *Risk Factor* discussions to be candidate predictors of abnormal returns. Therefore I do not pre-specify a set of terms I believe will be predictive. Instead, terms are selected into the model based on the strength of their association with abnormal returns on climate policy announcement dates. Once predictive terms are found, a sufficient reduction of the text for each firm can be defined as the sum of term counts multiplied by their coefficient in the inverse regression. This is my climate policy risk exposure score. Therefore, MNIR uncovers both which terms are predictive and how they should be weighted. An additional advantage of this method is its relative simplicity and similarity to econometric models of discrete choice.

Third, the exposure score is validated by its ability to pick out climate policy related terms and its ranking of fossil intensive firms as highly exposed, and green technology firms as negatively exposed. I conduct three further exercises to validate the risk measure. First, I show that it is correlated with firm-level emissions. Second, a higher exposure score is associated with a greater likelihood of lobbying on environmental issues.[3] Third, I re-estimate the exposure score using a subset of the climate policy announcement dates and show it is predictive on the left-out 'test' dates. In this case it also out-performs alternative measures of risk exposure from the literature.

Fourth, with this measure of risk exposure in hand I investigate whether climate policy risk is priced. Both cross sectional regressions following Bolton and Kacperczyk (2021), and portfolio sorts indicate that there is no premium over the sample considered. In fact, the high minus low exposure portfolio earns a negative premium over the period 2012-2020 which turns weakly positive only over the last two years 2021-2022. Markets react to climate policy announcements, however I do not observe a premium associated with this risk.

Finally, I turn to rationalize this empirical result within the framework of the asset pricing model. Green preference shocks which reduce cash-flows for browner firms offer a potential explanation. When a green preference shock hits, abnormal returns fall for brown firms. Therefore, although this also implies a brown premium in equilibrium, persistent preference shocks over the estimation period imply lower realized returns for brown firms. This is plausible given the rise of sustainable investing over the sample period. In order to separate these preference shocks from policy shocks in the data, I draw on a distinction between equities with low policy risk exposure, and equities

[3] I obtain data on firm-level lobbying from Open Secrets: https://www.opensecrets.org/federal-lobbying/

perceived to be ‘sustainable’ by ESG investors.

I utilise a sign-restriction SVAR to separate these shocks. van der Beck (2021) constructs a representative ESG portfolio. I compare this portfolio to the low-high risk portfolio constructed using the text-based exposure score and construct an over-under portfolio which goes long in equities over-weight in the ESG portfolio relative to their policy risk exposure score (e.g. a relatively green oil company), and short in equities under-weight in the ESG portfolio relative to their exposure score (e.g. a tobacco company). The idea of the sign restriction is that the over-under portfolio should earn positive returns when a preference shock materializes but not when a policy shock does, whereas the low-high portfolio should increase in both instances. I show that the empirically identified preference shocks explain about as much of the return dynamics of the low-high portfolio as the policy shocks do.

Relation to the literature. This paper contributes to an extensive empirical literature that investigates the relationship between environmental and financial performance. The first strand of this literature asks whether stocks of firms with higher carbon emissions earn higher returns. Influential papers by Bolton and Kacperczyk (2021, 2023) provide evidence of a carbon premium: stocks of firms with higher emissions earn higher returns which can’t be explained by differences in other known risk factors. These results are interpreted as evidence of investors seeking compensation for holding stocks which are exposed to ‘carbon risks’ e.g. from regulation to limit emissions or technology risk from renewable energy. Hsu, Li, and Tsou (2023) similarly find a premium associated with non-carbon industrial pollution. Ilhan, Sautner, and Vilkov (2022) find that climate policy uncertainty is priced in the option market.

Subsequent papers have questioned the existence of such a carbon premium. Aswani, Aneesh, and Shiva (2023) find that the carbon premium result is not robust to restricting the sample to firm-disclosed rather than vendor-imputed emissions data. They also stress the point that returns are not correlated with emissions intensities - emissions scaled by firm size - which may be more informative for climate risk exposure. Bauer et al. (2022) use disclosed emissions data for G7 countries and find a negative premium. Enders et al. (2023) observe a negative carbon premium in Europe. Zhang (2023) argues that after accounting for lags in the release of carbon emissions data, the premium also turns negative. Hong, Weikai Li, and Xu (2019) find that food stock prices under-react to climate change risk. I contribute to this literature by asking a related question: is there a *climate policy* risk premium? In order to answer this question and avoid the problems with emissions data outlined above, I construct a market implied measure of risk using text data from corporate reporting.

Closely related to this paper is a strand of literature which constructs firm-level measures of risk exposure. Hassan et al. (2019) use text data from earnings calls to build firm-level measures of political risk exposure. The share of quarterly earnings conference calls devoted to discussing political risks is used as a measure of exposure. Sautner et al. (2023) extend this approach to the case of climate risks. They provide a list of initial keywords associated with climate change and use a keyword discovery algorithm to construct sets of related terms whose frequency in earnings calls is defined as climate risk exposure. In a later paper Sautner et al. (2024) find that their climate change exposure measure carries a positive and increasing premium since the financial crisis. Li, Tang, and Yao (2199) also use earnings calls to identify climate related risks, in this case using a pre-specified dictionary. Baz, Michaelides, and Zhang (2022) construct another dictionary based measure of exposure to risk using firms form 10-K. They document that more exposed firms experience positive stock return effects after the 2016 Trump election. Engle et al. (2020) construct a climate news series. Huynh and Xia (2021) find that bonds with a higher climate change news beta earn lower future returns. Faccini, Martin, and Skiadopoulos (2022) construct aggregate measures of climate risk from textual analysis of news coverage. Blasberg, Kiesel, and Taschini (2021) use Credit Default Swap spreads to construct a market-implied carbon risk factor.

I contribute to this strand of the literature using Risk Factor discussions from firms 10-K filings and a supervised machine learning algorithm utilising information from climate policy announcement dates. Davis, Hansen, and Seminario-Amez (2021)construct firm-level exposures to COVID-19 news using this data, showing that the ML approach yields richer characterizations of risk than the dictionary approach and outperforms it in terms of goodness-of-fit. My contribution is to apply this more sophisticated ML method to the climate risk case. Instead of taking pre-specified terms, consistent with the asset pricing model, this algorithm selects terms into the model which are predictive of stock price movements on climate policy announcement dates. Therefore, this paper is also related to empirical literature employing climate policy event studies.

Seltzer, Starks, and Zhu (2022) use the Paris Agreement as a shock to expected climate regulations and show that these risks affect corporate bond risk assessment and pricing. Barnett (2023) uses transition related events to show that shifts in transition risk lead to patterns consistent with a 'run on fossil fuel'. Cassidy (2023) uses tools from natural language processing to construct a set of climate policy announcements from Executive

Branch communication. Hengge, Panizza, and Varghese (2023) use EU ETS surprises following Känzig (2023) to show that events resulting in higher carbon prices lead to negative abnormal returns which increase with a firm's carbon intensity. I contribute to this literature by constructing a set of climate policy announcements which are informative for the risk exposure of listed firms.

A final strand of the literature considers 'ESG' investing more broadly. Pástor, Stambaugh, and Taylor (2021) develop an asset pricing model which includes both policy and preference shocks. Green firms outperform brown firms when policy shocks hit, but brown firms earn higher expected returns in equilibrium. Pástor, Stambaugh, and Taylor (2022) provide some evidence that in fact stocks of green firms earn higher realized returns potentially due to increased concerns about climate change. van der Beck (2021) argues that the high realized returns of ESG portfolios can be explained by price pressure from flows towards sustainable funds. Therefore these high realized returns need not reflect high expected returns. In this paper I utilise the theory of Pástor, Stambaugh, and Taylor (2021). By considering stock price movement on climate policy announcement days I can build a measure of risk exposure. I use the ESG portfolios build by van der Beck (2021) to separate policy and preference shocks in a sign restriction SVAR.

Outline. Section 2 presents an Organizing Framework to motivate the empirical approach. Section 3 outlines the Data and Empirical Approach. Section 4 discusses Results, section 5 Interprets these results in the context of the Organizing Framework, and section 6 concludes.

2. Organizing Framework

This section presents a simple theoretical framework useful for understanding climate policy risk and its relation with stock returns. The framework follows closely the model of Giglio et al. (2021) and is similar in structure to models of long run risk a la Bansal and Yaron (2004).

Assume that aggregate consumption follows

$$\Delta c_{t+1} = \mu + x_t - J_{t+1}, \tag{1}$$

where

$$x_{t+1} = \mu_x + \rho x_t + \phi J_{t+1}. \tag{2}$$

c_t is the log of aggregate consumption which corresponds to aggregate output in equilibrium.[4] The process x_t captures persistent changes in the growth rate of consumption. J_t is a jump process that takes a value of ξ with probability λ in each period and 0 otherwise. I interpret this shock as a climate policy shock i.e. an unexpected shift in climate policy stringency.[5] In line with the empirical evidence, an increase in climate policy stringency depresses consumption (Känzig 2023). A calibration of $\phi > 0, \rho > 0$ implies partial mean reversion in consumption after a policy shock.

I specify a separate cash flow process for risky assets which in this setting correspond to equity. There is a cross section of equities in the model indexed by i. The process for equity cash flows is similar to that for consumption:

$$\Delta d_{i,t+1} = \mu_d + y_t - \eta_i J_{t+1} \tag{3}$$

$$y_{t+1} = \mu_y + \omega y_t + \psi J_{t+1} \tag{4}$$

Equities are differentially exposed to the climate policy shock. This is captured by the η_i coefficients which measure exposure. Equities with larger positive η_i are more negatively exposed to climate policy, equities with larger negative η_i are more positively exposed to climate policy.

I assume the existence of a representative agent who maximizes lifetime utility and faces a complete set of financial instruments. The per period utility function features constant relative-risk aversion (γ):

$$U(C_t) = \delta^t \frac{C_t^{1-\gamma}}{1-\gamma} \tag{5}$$

where δ is the rate of time preference. I am interested in one period returns and risk premia. One period returns of equity i are given by:

[4]The economy is closed and does not feature investment.
[5]Giglio et al. (2021) interpret J_t as a physical climate disaster in their setting.

(6)
$$R_{i,t,t+1} = \frac{P_{i,t+1}}{P_{i,t}}$$

Therefore log one period returns are:

(7)
$$r_{i,t,t+1} = \ln P_{i,t+1} - \ln P_{i,t} = p_{i,t+1} - p_{i,t} + \Delta d_{t+1}$$

Assuming that log returns and stochastic discount factor are jointly normally distributed, the risk premium is then given by:

(8)
$$E_t\left[R_{i,t,t+1}\right] - R^f_{t,t+1} \simeq \gamma \,\mathrm{Cov}_t\left[r_{i,t,t+1}, \Delta c_{t+1}\right]$$

where $E_t\left[R_{i,t,t+1}\right]$ is the one period expected return of equity i and $R^f_{t,t+1}$ is the risk-free rate.

Solving the model recursively, it can be shown that

(9)
$$r_{i,t,t+1} = -[a_i + b_i x_{t+1} + e y_{t+1}] + [\mu_d + y_t] - \eta_i J_{t+1}$$

(10)
$$E_t\left[R_{i,t,t+1}\right] - R^f_{t,t+1} \simeq \gamma \eta_i \xi^2 \lambda(1 - \lambda)$$

where a_i, b_i, and e solve recursive equations derived analytically by Giglio et al. (2021). The exact expressions are reported in Appendix A.

Discussion. A climate policy shock (e.g. an announcement of an increase in a carbon tax or a fall in an emissions cap) leads to a fall in consumption. When the climate policy shock hits, one period returns fall for those firms (equities) which are negatively exposed. One period returns increase for those firms which are positively exposed. Therefore, climate policy exposure constitutes a risk and commands a risk premium: investors demand compensation for holding equity which is negatively exposed to climate policy. The model predicts a risk premium which is increasing in exposure (η_i).

This insight will guide the empirical approach of this paper: climate policy jump days will be used to identify equities which are exposed to climate policy risk.

3. Data and Empirical Method

3.1. Data

Firm-level returns. I consider a sample from 2012-2022 with attention restricted to U.S-incorporated firms with share prices quoted in U.S. Dollars. Financial data is obtained from Compustat. Daily and monthly returns are obtained for 3,894 equity securities for which there is return data on at least three of the jump days listed below and a non-empty part 1a) in form 10-K. See Appendix B for more information about the sample. Abnormal returns are constructed on each climate policy announcement date in the standard way per the Capital Asset Pricing Model (CAPM):

$$\text{Abn}_{i,t} = \log\left(\frac{p_{i,t}}{p_{i,t-1}}\right) - R_{f,t} - \text{ beta }_i \times \left(R_{M,t} - R_{f,t}\right) \tag{11}$$

where $p_{i,t}$ denotes the share price for stock i on day t, R_f denotes the four-week treasury bill rate, $beta_i$ is the stock's CAPM beta, and R_M is the value-weighted average market return. Each stock's beta is estimated using an OLS regression of daily excess return on the market excess return in the sample of all trading days for the year prior.

In cross-sectional regressions I use a number of variables following Bolton and Kacperczyk (2021) which are constructed using Compustat. These variables are outlined in Appendix B.

Risk Factors Text. The Securities and Exchange Commission (SEC) requires the vast majority of publicly held firms to include a discussion of *Risk Factors* in their annual corporate reporting - their form 10-K. Firms are advised to include any factor which could impact future earnings in these reports. If the firm omits material information or risks, investors can sue for compensation. Machine-readable versions of these reports are available on EDGAR. The raw text is processed into a corpus of words and phrases as outlined in Appendix B. After pre-processing there are 21,876 terms. Following Davis, Hansen, and Seminario-Amez (2021), this large number of terms necessitates some form of dimension reduction. Therefore I employ the MNIR model which operates on all terms (Taddy 2013).

3.2. Empirical Approach

Return movements on climate policy announcement dates contain information on firms' relative exposure to climate policy risks. I use a supervised machine learning algorithm to uncover predictors of this risk exposure and build exposure scores for each firm. As the US does not have a national carbon tax, I use a narrative approach to identify dates on which other climate policy announcements have been made.

Jump Days. In order for a candidate date to be useful in this setting, the announcement must lead to movements in abnormal returns which I can be confident are due to the climate policy announcement. Therefore, the announcements must be i) unexpected, ii) reasonably large, and iii) not confounded by other non-climate announcements at the same time. Therefore, I select jump days according to the following steps:

a. Collate a list of climate policy announcements from 2012-2022. These candidate announcements come from a number of sources:
    - Important dates relating to the passage of the Inflation Reduction Act (IRA) from Bistline, Mehrotra, and Wolfram (2023).
    - Important dates relating to the passage and challenge of the Clean Power Plan (CPP).
    - Tabling of other relevant climate legislation e.g. the Green New Deal.
    - Signing of International Agreements and Announcements from UN Conferences.
    - Executive orders relating to climate.
    - Supreme Court decisions on climate legislation.
b. Remove those candidate announcements of which there was no coverage in either the *Wall Street Journal* or the *New York Times*.
c. Remove those candidate announcements which were made at the same time as more salient market news materialized.
d. Remove those candidate announcements which were not associated with negative abnormal returns for especially brown industries: Oil, Coal and Mining.

Step a) results in a list of 53 candidate dates. After step b) 45 dates remain. After step c) 33 dates remain, and step d) leaves us with the final 15 dates. Details of this process are confined to Appendix B. The final set of 15 dates are listed in Table 1.

TABLE 1. Climate Policy Announcement Dates

Announcement	Date
Loss and damage agreement at COP27	21/11/22
Senate Passage of IRA	08/08/22
Congressional Budget Office Pricing of IRA	03/08/22
Senator Manchin U-turn on IRA	27/07/22
Supreme Court Strikes down CPP	30/06/22
Build Back Better bill passes the house	19/11/21
Glasgow Pact at COP26	13/11/21
IPCC Sixth Assessment Report	09/08/21
Biden Executive Order to rejoin Paris Agreement	20/01/21
Formal notice of intention to withdraw from Paris Agreement	04/11/19
Green New Deal Resolution presented to Congress	07/02/19
Trump's Executive order to begin review of CPP	28/03/17
US and China announce they will sign the Paris Agreement	01/04/16
Paris Agreement Reached	14/12/15
Obama unveils CPP	03/08/15

Note: List of Climate Policy Announcement Dates used. Dates in red are instances of climate policy becoming less stringent.

Multinomial inverse regression. I am interested in what features of the *Risk Factors* text are predictive of abnormal return responses to the climate policy announcements, and in building an exposure score for each firm. I utilise the MNIR method developed by Taddy (2013) and applied by Gentzkow, Shapiro, and Taddy (2019) and Davis, Hansen, and Seminario-Amez (2021). The text for each firm is represented by a bag-of-words i.e. it is a V-dimensional vector $\mathbf{x_i}$ of terms. V is the number of unique terms in the corpus, and $x_{i,v}$ gives the number of instances of term v in firm $i's$ filings. MNIR models the relationship between the term counts and the object of interest - here abnormal returns on climate policy announcement dates - directly. The structure of the model is very similar to econometric models of discrete choice. MNIR posits

$$\mathbf{x}_i \sim \text{MN}\left(\mathbf{q}_i, N_i\right) \tag{12}$$

where $\mathbf{q}_i$ is a multinomial V-dimensional probability vector and N_i is the total number of terms in firm i's filings. The probabilities of each term for each firm are therefore given by $\mathbf{q}$. The probability of term v for firm i is given by:

$$q_{i,v} = \frac{\exp\left(a_v + \mathbf{y}_i^T \mathbf{b}_v\right)}{\sum_v \exp\left(a_v + \mathbf{y}_i^T \mathbf{b}_v\right)} \tag{13}$$

where $\mathbf{y}_i = (\text{Abn}_{i,t}, \mathbf{c}_i)$ contains the firm's abnormal returns on the announcement dates as well as a vector of controls.[6] The baseline frequency of term v is set by the parameter a_v, and b_v is a vector of coefficients describing how firm observables map to the probability of term v appearing in the firm's filings.

Therefore, we have a multinomial logistic regression over V terms where the outcome variable is the probability of a given term appearing in a random draw from the firm's filings. This is a standard econometric model of discrete choice but for the high dimensionality of V. In order to reduce the dimensions of the problem, the model is estimated using Bayesian regularization methods with a Gamma-Laplace prior structure on the regression coefficients. This is a generalization of the standard LASSO penalty with the possibility for term-specific penalties. As shown in Taddy (2013), this prior trades off goodness-of-fit and model complexity via the maximization of an information criterion.

I now have a model of the inverse relationship of term probabilities given returns. To move forward to predictors of abnormal returns I follow Taddy (2013) and Davis, Hansen, and Seminario-Amez (2021) and define a sufficient reduction projection

$$z_i = \sum_v x_{i,v} b_{1,v} \tag{14}$$

with the property that

$$\text{Abn}_i \perp \mathbf{x}_i | z_i, N_i, \mathbf{c}_i \tag{15}$$

This implies that conditional on the scalar projection z_i - my exposure score - the high-dimensional text data contains no extra predictive information for returns. Therefore, we can model Abn_i as a function of z_i, N_i, $\mathbf{c}_i$ while disregarding $\mathbf{x}_i$. z_i summarizes the predictive information in the text for how abnormal returns respond to climate policy

[6] I control for the number of terms in the part 1a), leverage, NAICS3, and market capitalization.

announcements.

Portfolio sorts and cross-sectional regressions. I use the constructed exposure scores to test for a climate policy risk premium. I proceed in two ways. First, I construct quintile portfolios sorted on exposure score. The low portfolio contains firms with the lowest exposure scores and the high portfolio contains firms with the highest exposure scores. After forming the portfolios I calculate value-weighted monthly returns on these portfolios. I also form a high-minus-low portfolio that takes a long position in the most exposed firms and a short position in the least exposed firms. I calculate the value-weighted return on this long-short portfolio and test whether it earns a positive return over the sample. Such a result would be indicative of a climate policy risk premium.

Following Bolton and Kacperczyk (2021) I also run cross-sectional return regressions to test for a premium. Here, the dependent variable is firm level monthly returns, the independent variable is the exposure score, and I include the same vector of controls as in Bolton and Kacperczyk (2021). In this case a significant positive coefficient on the exposure score would be evidence for a climate policy risk premium.

4. Results

4.1. Inverse Regression

The MNIR model is fit to the climate policy announcement dates. I control for industry (SIC), market capitalization, and leverage. The model selects 18,054 terms. Of these, it puts positive weight on 9,573 and negative weight on 8,481. 5,946 terms are not selected into the model. Table 2 shows the terms with the highest positive and negative coefficients in the inverse regression - the first element of $\mathbf{b_v}$. As is standard I weight the coefficients by their tf-idf. The table indicates the terms which are most predictive of abnormal return responses on the climate policy announcement dates. Therefore it is useful in uncovering factors which contribute to exposure to climate policy risk.

The table shows that terms associated with negative abnormal returns are often related to fossil fuels ("coal mining", "gas price", "natural gas cost"). The model therefore captures that as well as fossil fuel producers, downstream firms who are dependent on fossil energy may be exposed to climate policy. The model also picks up terms directly related to climate policy ("fuel surcharge", "environmental permit", "permit requirement"). Terms associated with positive abnormal returns are related to activities

TABLE 2. Most Predictive Terms

rank	term	rank	term
1	mineable	1	battery
2	regulatory_clearance	2	solar_panel
3	refined	3	prc_foreign_exchange
4	politically	4	electric_transmission
5	mineralize	5	wind
6	fuel_surcharge	6	clinician
7	capital_requirement	7	secondary_trading
8	coal_mining	8	solar
9	mineral_resource	9	financially
10	environmental_permit	10	financial_statement
11	power_shortage	11	fda_approval
12	energy_use	12	variable_interest_rate
13	power_system	13	available_therapy
14	refractory	14	innovator
15	boeing	15	robot
16	gas_exploration	16	department_of_defense
17	gas_price	17	obamas
18	production_from_oil	18	willingness_of_patient
19	permit_requirement	19	solar_energy
20	natural_gas_cost	20	clinical_research

Note: The table shows the terms with the largest coefficients weighted by their tf-idf score in the MNIR model fit to climate policy announcement days. The left hand side of the table shows the 20 terms which have the largest negative tf-idf weighted coefficients. The right hand side of the table shows the 20 terms with the largest positive weighted coefficients. The tf-idf is computed as the count of uses of the term in the whole corpus multiplied by $log(\frac{24{,}000}{df_v})$ where df_v is the number of firms who use term v in their filings. Terms that appear in the filings of more than two-thirds of firms are excluded.

which are less likely to be affected by climate policy. Renewable technologies, healthcare, and financial services all show up in the list of predictive terms.

Table 3 shows the coefficients on each of the terms in the Sautner et al. (2023) list of initial keywords which are included in my corpus. They specify a list of initial climate-related keywords and use a keyword-discovery algorithm to build lists of similar terms. I restrict attention to their ‘opportunity‘ and ‘regulatory’ bigram lists. In cases where the exact bigram is not included in my corpus, I instead take the closest related term. Table 3 shows that the MNIR model correctly sorts the terms. Bigrams specified as initial ‘regulatory’ terms are selected into the model with negative coefficients i.e. the model picks up that they are related to negative climate policy risk exposure. Bigrams specified as related to 'opportunity' are selected into the model with positive coefficients. Therefore, by using abnormal return responses on climate policy announcement dates as labels in a supervised machine learning algorithm, I circumvent the need for specifying sets of initial keywords. Instead, I allow the data to tell me which terms are most informative for climate policy risk exposure. The model succeeds in selecting terms which we would imagine a priori would be associated with positive or negative climate-policy risk exposure.

There are two additional advantages of this approach. Firstly, the inverse regression makes clear not only whether a term is positively or negatively related to climate policy risk exposure but also provides an indicator of the strength of this association. Therefore, very predictive terms will be highly weighted, and less predictive terms less highly weighted when constructing firm-level exposure scores. Secondly, the Risk Factor text by definition only discusses downside risk. Therefore, I do not have to be concerned with performing corrections for sentiment.

TABLE 3. Sautner et al initial keywords

coefficient	term	coefficient	term
-25.72	greenhouse_gas	5.46	solar
-19.52	greenhouse_gas_emissions	5.96	solar_energy
-20.41	emission_reduction	5.32	solar_panel
-17.39	pollution	4.61	renewable_energy
-17.01	carbon_dioxide	3.61	renewable
-31.97	emission_of_carbon	6.22	wind
-13.76	environmental_standard	3.81	wave
-12.58	epa	2.69	electric_vehicle
-5.02	dioxide	4.51	geothermal
-22.09	energy		
-5.32	nox		

Note: The table shows the coefficients on terms included in Sautner et al's initial keyword list. Where the exact bigram does not appear in my corpus I instead take the closest term available. The left column shows coefficients on terms in the 'regulatory' category, the right column shows coefficients on terms in the 'opportunity' category. The mean coefficient value in the corpus is –2.39, with a standard deviation of 7.09. The maximum value is 9.1 and the minimum is –31.

4.2. Exposure Scores

Having fit the MNIR model to abnormal returns on climate policy announcement dates, I can now construct the sufficient reduction projection:

$$z_i = \sum_v x_{i,v} b_{1,v}$$

This scalar summarizes the information in the text which is relevant for explaining abnormal returns on climate policy announcement dates. This z_i will be my text-based exposure score for each firm. Table 4 shows the ten firms with the most negative exposure scores.

TABLE 4. Most exposed firms

	name	industry
1	Peabody Energy Corp	Mining
2	Fieldpoint Petroleum Corp	Oil and Gas Extraction
3	Whiting Petroleum Corp	Oil and Gas Extraction
4	Continental Resources Inc	Oil and Gas Extraction
5	Epizyme Inc	Chemical Manufacturing
6	Ring Energy Inc	Mining
7	Chevron Corp	Petroleum and Coal Manufacturing
8	Jones Energy Inc	Oil and Gas Extraction
9	Consol Energy Inc	Oil and Gas Extraction
10	Antrero Resources Corp	Support Activities for Mining

Note: The table shows the 10 firms with the most negative values of z - the sufficient reduction projection from the MNIR model.

These firms discuss risks in their filings which the inverse regression of abnormal returns on term counts has discovered are predictive of negative responses to climate policy. For example, the most predictive paragraph from the filings of the most exposed firm, Peabody Energy Corp, discusses environmental regulation extensively:

"Concerns about the environmental impacts of coal combustion, including perceived impacts on global climate issues, are resulting in increased regulation of coal combustion in many jurisdictions and unfavorable lending policies by government-backed lending institutions and development banks toward the financing of new overseas coal-fueled power plants, and interest in further such regulation and policies, which could significantly affect demand for our products. Global climate issues continue to attract public and scientific attention. Numerous

reports, such as the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, have also engendered concern about the impacts of human activity, especially fossil fuel combustion, on global climate issues. In turn, increasing government attention is being paid to global climate issues and to emissions of what are commonly referred to as greenhouse gases, including emissions of carbon dioxide from coal combustion by power plants. Enactment of laws or passage of regulations regarding emissions from the combustion of coal by the U.S., some of its states or other countries, or other actions to limit such emissions, could result in electricity generators switching from coal to other fuel sources. Further, policies limiting available financing for the development of new coal-fueled power plants could adversely impact the global demand for coal in the future."

As well as producers of fossil fuels, the model also identifies firms that rely on fossil fuels, for example airlines, as exposed to climate policy risk. Similarly, the model identifies firms that rely on renewables as positively exposed to climate policy risk. Table 5 shows two such examples. Tesla, the electric vehicle company is shown as having one of the lowest exposure scores due to the inclusion of terms related to renewable energy in its filings. Delta airlines on the other hand is quite exposed to the climate policy risk.

4.3. Exposure Score Validation

The MNIR model is validated by its ability to select climate risk related terms. The subsequent exposure score is validated by its high ranking of firms with known intensity in 'brown' activities. I perform three additional checks on the validity of the exposure score. First, I show that the firm-level exposure score is correlated with emissions intensity.[7] The table in Appendix C shows that a higher exposure score is strongly correlated with higher reported emissions, higher estimated emissions, and higher emissions intensities (emissions scaled by revenues). A one standard deviation increase in exposure score is associated with more than 1 million tonnes of additional CO2 emissions.

Secondly, I show that firms with higher exposure scores are more likely to lobby on environmental issues. Lobbying firms are required to provide a good-faith estimate rounded to the nearest $10, 000 of all lobbying-related income from their clients in each quarter. They are also required to report the purpose of the lobbying and the organiza-

[7] I obtain firm-level emissions data for 2022 from Refinitiv. Summary statistics are included in Appendix C.

TABLE 5. Three Example Firms

Firm	Terms	tf-idf x MNIR coeff.
S M ENERGY CO #11 ↗	crude oil gathering	$475 \cdot 2$
	drill rig	$297 \cdot 5$
	ngl	$136 \cdot 9$
	greenhouse gas emissions	$99 \cdot 4$
	hydraulic fracturing fluid	$96 \cdot 1$
DELTA AIRLINES INC #76 ↗	airbus	$943 \cdot 2$
	airline	$445 \cdot 6$
	crude oil gathering	$243 \cdot 5$
	passage of legislation	$216 \cdot 6$
	refined	$81 \cdot 9$
TESLA INC #30 ↘	electric utility	$185 \cdot 7$
	alternative energy source	$135 \cdot 3$
	solar	$112 \cdot 0$
	batch	$95 \cdot 5$
	modulator	$86 \cdot 8$

Note: The table shows the 5 terms with the largest tf-idf weighted coefficients for three example firms.

tion being targeted. I obtain data on this lobbying behaviour from OpenSecrets.[8] This dataset includes a list of 80 possible topics that firms can lobby on. I designate five of these topics as climate related: "Clean Air and Water", "Environment and Superfund", "Fuel, Gas, and Oil", "Natural Resources", and "Hazardous and Solid Waste". Following Hassan et al. (2019), I generate a dummy variable that equals 1 if a firm lobbies on a climate topic between 2016 and 2019, and 0 otherwise. I run the following regression:

$$\mathbb{1}\left[\text{Lobbying}_{\,i} > 0\right] = \delta_j + \theta z_i + \gamma X_i + \epsilon_i, \tag{16}$$

where δ_j is an industry fixed effect at the NAICS 3 digit level. X_i is a vector of controls, and θ measures the association of a firm's climate risk exposure score and its propensity to lobby on climate issues.

Table 6 indicates the results from this regression. The coefficient estimate (0.039, s.e.=0.006) implies that a one-standard-deviation increase exposure to climate policy

[8]https://www.opensecrets.org/federal-lobbying/

TABLE 6. Climate related Lobbying

	$1[\text{Lobbying}_i > 0]$		
	(1)	(2)	(3)
z_i	0.052***	0.041***	0.039***
	(0.003)	(0.006)	(0.006)
Market Cap			0.000***
			(0.000)
Constant	0.022***	−0.010	−0.015
	(0.004)	(0.043)	(0.041)
Industry Fixed Effects		*Yes*	*Yes*
Observations	3,887	3,865	3,982
R^2	0.062	0.161	0.197
Adjusted R^2	0.062	0.142	0.178
F Statistic	256.856***	8.239***	10.711***

Note: This table shows the results from regressions of a dummy variable that equals one if firm i lobbies on climate issues on the firm's climate-risk exposure. z_i is standardized by its standard deviation. ***, **, and * denote statistical significance at the 1, 5, and 10 % level, respectively.

related risk is associated with a 3.9-percentage-point increase in the probability that a given firm lobbies on climate related issues. This exercise suggests that the climate policy risk exposure measure is meaningful economically: firms manage their exposure to this risk by lobbying on climate related issues. As discussed in Hassan et al. (2019), this is consistent with theoretical models that predict that exposure to political risks is associated with an increase in active management through lobbying as well as donations (Peltzman 1976).

Thirdly, I test the predictive power of my exposure measure and compare it to that of alternative measures. Figure A.1 in Appendix C shows that on the set of climate policy announcement dates, the R squared of the regression using my exposure score as a predictor is higher than that using the Sautner et al. (2023) exposure score. This is not surprising given that my exposure score is constructed using abnormal return responses on those dates. Therefore, I randomly select half of the climate policy announcement dates as a training set and leave the other half as a test set. I construct my exposure score using only the training set and compare the predictive performance of my exposure score and the Sautner et al. (2023) exposure score on the remaining dates. Again, as shown in figure A.2, the MNIR model provides an exposure score with better predictive performance.

4.4. Portfolio Sorts

Armed with this validated exposure score, I test whether higher exposure to climate policy risk is associated with higher returns - whether there exists a climate policy risk premium. Table 7 presents the results of the size weighted quintile portfolio sorts on exposure.

For the period 2012-2022, the high-minus-low portfolio generates a significant excess return of –0.39% per month. The pattern is similar for the period 2012-2019. Between 2020 and 2022, the high-minus-low portfolio earns insignificant positive excess returns. Therefore, there is no evidence of a climate policy risk premium. In fact, *less risky* firms have been paying higher returns over this period. Figure 1 presents the cumulative return of the return spread between the high and low exposure portfolios. The high-minus-low portfolio loses value between 2012 and 2020, regaining somewhat from 2020 to 2022.

TABLE 7. Portfolio Sorts

	L	2	3	4	H	H-L
2012-2022	1.33***	1.65***	1.29***	1.18***	0.94**	–0.39**
	(0.35)	(0.37)	(0.34)	(0.32)	(0.52)	(0.29)
2012-2019	1.56**	1.42***	1.38***	1.20**	1.00**	–0.564*
	(0.34)	(0.35)	(0.33)	(0.37)	(0.41)	(0.32)
2020-2022	0.37	1.78**	1.31	1.74*	1.1*	0.72
	(0.6)	(0.58)	(0.65)	(0.61)	(0.53)	(0.58)

Note: This table presents the monthly value-weighted returns of the exposure score sorted portfolios. The standard errors are reported in the parenthesis below the coefficients. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

4.5. Cross-sectional Regressions

As an additional check I run the following regression:

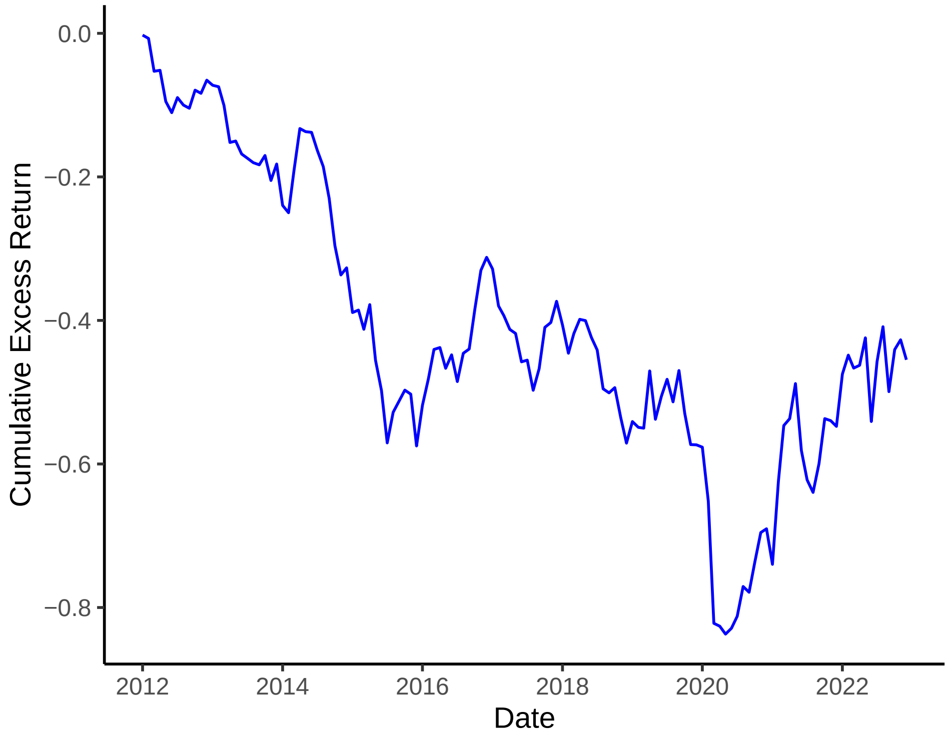

FIGURE 1

This figure plots the cumulative return of the return spread between the high and low climate policy risk exposure portfolio.

$$R_{i,t} = \alpha_0 + \beta z_i + \gamma \mathbf{x}_{i,t-1} + \mu_t + \mu_j + \epsilon_{i,t}, \tag{17}$$

where $R_{i,t}$ is the return on equity of firm i in month t, z_i is the exposure score and $\mathbf{x}$ is a vector of controls known to predict returns.[9] The coefficient of interest is β. Standard errors are clustered at the firm and year levels. This specification is identical to that in Bolton and Kacperczyk (2021) but for the use of exposure score rather than emissions as an explanatory variable. Contrary to Bolton and Kacperczyk (2021), however, I find no evidence for a carbon policy risk premium. Table 8 presents the results of this regression.

A one standard deviation increase in exposure score is associated with a 58 bps decrease in returns. This result is robust to the inclusion of narrow industry fixed effects. The negative premium also remains if the sample is restricted to post 2015, or

[9]The list of controls along with summary statistics are included in Appendix A.

TABLE 8. Cross-sectional Regression

Variable	Model 1	Model 2
z	-0.5818**	-0.5786**
	(0.1744)	(0.2259)
Market Cap	0.0000	0.0000
	(0.0000)	(0.0000)
Book to Market	1.5296***	1.7406***
	(0.4304)	(0.4914)
Leverage	0.4643	0.5703
	(0.4716)	(0.4914)
Investa/a	0.2991	0.9567
	(1.6082)	(1.7729)
HHI	-2.3468	-2.7809
	(2.3227)	(3.7085)
PPE	0.0834	-0.1145*
	(0.0599)	(0.0526)
SALESGR	-0.2819	-2.7809
	(0.4599)	(0.4356)
EPSGR	1.8133***	1.8177***
	(0.1296)	(0.1237)
beta	-0.1908*	-0.2712***
	(0.0939)	(0.0551)
ROE	0.3826*	0.3289
	(0.1776)	(0.2399)
MOM	11.3085*	11.1848*
	(5.8725)	(5.7765)
VOL	8.4160**	8.4967**
	(3.2632)	(3.182)
Observations:	177,625	177,625
Year-Month Fixed Effects:	YES	YES
Industry Fixed Effects:	NO	YES
RMSE:	17.5	17.5
Adj. R^2:	0.1388	0.1392

Note: This table presents the results of the cross sectional regression with the dependent variable being firms' monthly returns. The columns present coefficient estimates and standard errors.The sample period is 2012-2022. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively. Model 1 includes year-month fixed effects. Model 2 additionally includes industry (NAICS4) fixed effects.

even post 2019. Therefore, in neither the cross-sectional regressions nor the portfolio sorts do we find evidence for a climate policy risk premium. Rather, it seems that less risky firms from a climate policy perspective are trading at a premium. In section 5 I attempt to explain this puzzle.

5. Preference Shocks: Rationalizing the Results

The organizing framework presented above predicts that firms with higher exposure to climate policy risk should earn higher returns on their equity. However, the empirical evidence indicates the opposite. I show that empirically identified preference shocks can rationalize this.

In the theory developed by Pástor, Stambaugh, and Taylor (2021), though there is a brown premium in equilibrium, there could be a negative gap between returns on brown and green assets if green preferences strengthen unexpectedly. In this case, a preference shock leads to higher realized returns for green than brown firms. Therefore, persistent strengthening of aggregate demand for green equities over the estimation period could explain the empirical results presented above. A similar mechanism is explored by van der Beck (2021) to explain the high realized returns from ESG investing. Given the rise of ESG investing and increasing salience of climate-related issues over the sample period considered, it is plausible that green aggregate demand could have increased in a persistent manner (Ardia et al. 2023).

The full treatment of consumer and investor preferences for green equity in general equilibrium is considered in Pástor, Stambaugh, and Taylor (2021). In my partial equilibrium framework, green consumer preference shocks can be modelled as entering the dividends process: consumer preferences becoming greener reduces the cashflow for browner firms. Therefore, equations 3 and 4 from above become:

$$\Delta d_{i,t+1} = \mu_d + y_t - \eta_i J_{t+1} - \upsilon_i V_{t+1}$$

$$y_{t+1} = \mu_y + \omega y_t + \psi J_{t+1} + \zeta V_{t+1}$$

where V is the consumer preference shock and υ_i is the exposure of firm i to the shock. Browner firms will therefore have larger υ_i. As above, the incidence of a preference shock will reduce one period returns of firms with positive υ_i, and those firms will earn higher expected returns in equilibrium. It is clear that such preference shocks could in theory explain the high realized returns of low policy risk equities in my sample as firms with higher υ_i are also likely to have higher η_i. Firms which are more exposed to preference shocks are likely to also be exposed to policy shocks e.g. an oil company will suffer from an increase in climate policy stringency and from consumer/investor preferences becoming greener.

This similarity in how policy and preference shocks affect returns makes it difficult to separate them empirically to understand their relative effects. I identify preference shocks by considering the different responses of the climate policy risk exposure portfolios and the representative ESG portfolio identified by van der Beck (2021). The idea is that while in general a higher η_i implies a higher υ_i, the rankings of firms by these two parameters need not be the same - some firms which are highly exposed to policy shocks may not be so highly exposed to preference shocks and vice versa.

5.1. Data

van der Beck (2021) builds a representative ESG portfolio by identifying a set of 551 ESG mutual funds using a list of sustainability keywords. Using data on these funds' portfolio holdings, he then constructs a representative ESG portfolio that pools their holdings. The deviations of a stock's holdings in the ESG portfolio from the aggregate mutual fund portfolio is therefore a revealed preference measure of how sustainable investors perceive that stock to be. I will consider this as a measure of υ_i: how exposed a firm is to preference shocks.

These deviations are compared to the weightings in my climate policy risk exposure portfolio: the zs or η_i's. I designate a stock to be over-weight in the ESG portfolio if it has a positive deviation in van der Beck's data and has a positive exposure score (z). An overweight stock is one which the supervised machine learning algorithm has picked out as negatively exposed to climate policy risk and is also held in a higher proportion in the ESG portfolio than in the aggregate portfolio. Such a stock is risky from a climate policy perspective however is perceived to be green. Examples of such stocks are fossil fuel companies which are exposed to climate policy risk but are less polluting than their competitors and are therefore included in the ESG portfolio as the 'best in class'. I designate a stock to be under-weight in the ESG portfolio if it has a negative deviation in van der Beck's data and a negative exposure score (z). Such a stock is not risky from a climate policy risk perspective but is perceived not to be sustainable. This category leverages the fact that ESG is a broader category than climate performance. Some stocks which are not negatively exposed to climate policy risk are nonetheless perceived as anti-ESG, for example tobacco companies or firms with poor management practices.

I construct a portfolio that goes long in these overweight stocks and short in those that are underweight (over-under). The performance of this portfolio will be compared to the performance of the portfolio that goes long in the lowest quintile of exposure score stocks and short in the highest quintile. As outlined above, this second portfolio earns a positive return over the sample. As a shorthand I will call this the green-brown portfolio, though it should be clear that I am discussing exposure to climate policy risk.

5.2. Sign-Restriction SVAR

These two portfolios are included in a sign-restriction SVAR in order to separate preference and policy shocks. The intuition is that the over-under portfolio should earn positive returns when there is a preference shock but not a policy shock. The green-brown portfolio should earn positive returns in both cases.

The VAR of the form:

$$Y_t = B(L)Y_{t-1} + U_t \tag{18}$$

where the vector Y_t includes monthly returns for the over-under portfolio, monthly returns for the green-brown portfolio, the oil price, and GDP is estimated following the standard sign-restriction methodology (Uhlig 2005; Fry and Pagan 2011), using the efficient algorithm from Arias, Rubio-Ramírez, and Waggoner (2018). The Bayesian estimation procedure uses a Normal-Wishart prior and a constant and 6 lags are included. In order to identify the structural preference and policy shocks I impose a sign restriction on the over-under and green-brown portfolios. The sign restrictions hold for the impact period only.

$$\begin{bmatrix} \\ \text{over-under} \\ \text{green-brown} \\ \text{GDP} \\ \text{Oil price} \end{bmatrix} = \begin{bmatrix} \epsilon_{pref} & \epsilon_{pol} & \epsilon_y & \epsilon_o \\ + & - & ? & ? \\ + & + & ? & ? \\ ? & ? & ? & ? \\ ? & ? & ? & ? \end{bmatrix}$$

5.3. Results

The structural VAR identifies the policy and preference shocks with the sign restrictions. This allows for a historical decomposition of the contribution of the shocks as well as a forecast error variance decomposition. Figure 2 shows that over the sample period, preference shocks have contributed to the returns process of the green-brown portfolio. The forecast error variance decomposition in Figure 3 shows that, in fact the contribution of preference shocks has been as important to explaining the dynamics of the green-brown portfolio returns as policy shocks.[10] This is indicative evidence that substantial policy shocks over the estimation period could explain the positive returns of the less risky portfolio in the data. In next steps I plan to calibrate my asset pricing model and input the empirically identified shocks to quantify the contribution of preference shocks to the surprising 'negative premium' result.

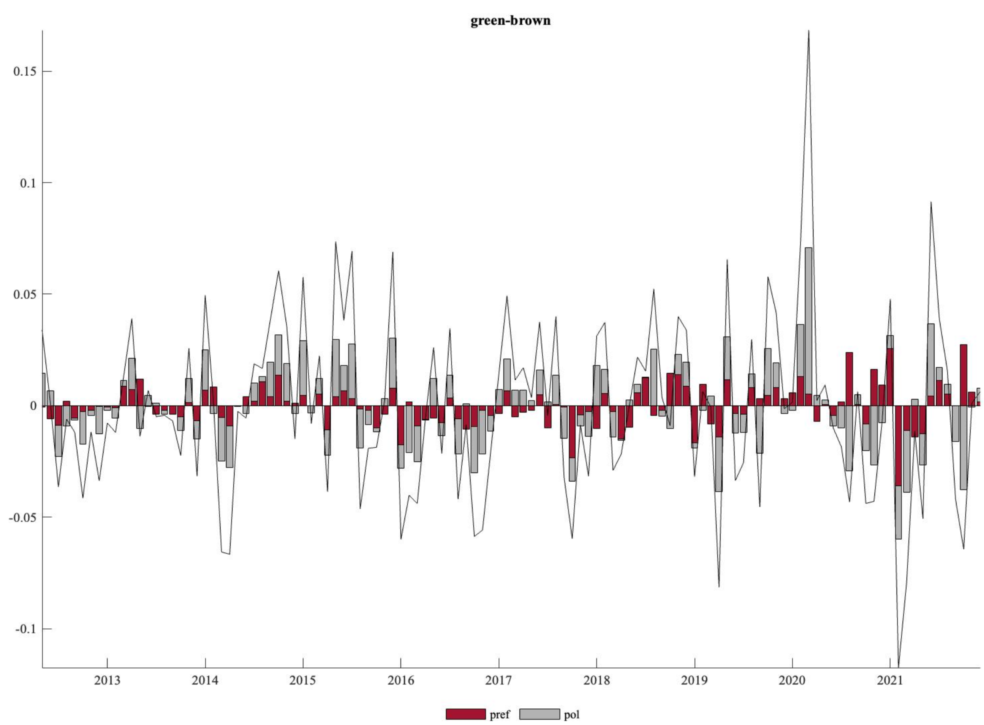

FIGURE 2. Historical decomposition of green-brown portfolio

6. Conclusion

This paper exploits text from firms' annual filings and return volatility on climate policy announcement dates to build a firm-level measure of exposure to climate policy risk.

[10]The Structural shocks and IRFS are presented in Appendix 3.

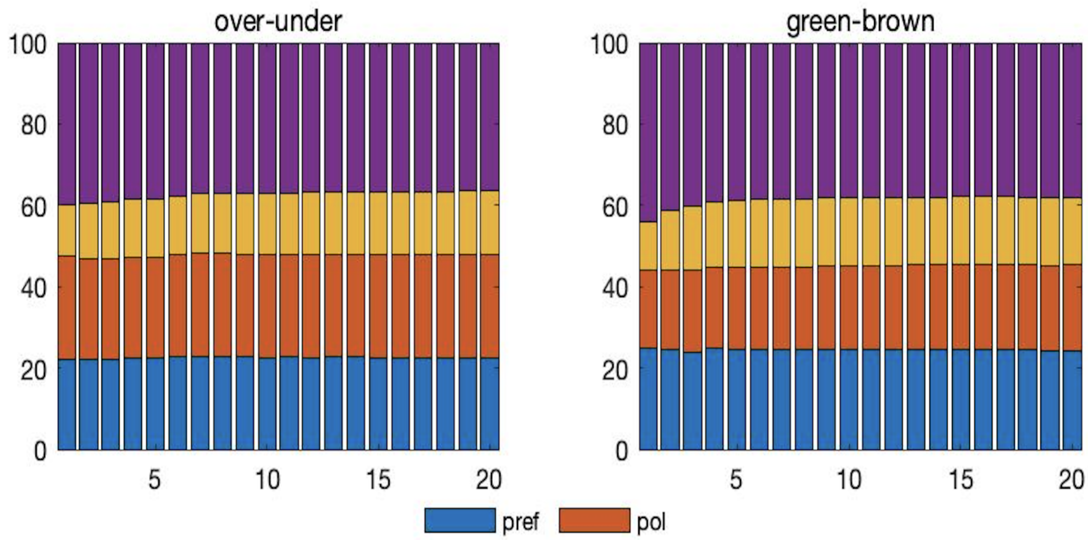

FIGURE 3. Forecast Error Variance Decomposition

The supervised machine learning algorithm employed, MNIR, considers all terms in the text and selects into the model those with predictive power for returns on the announcement dates. The subsequent exposure score is correlated with firm-level emissions, environmental lobbying behaviour, and is predictive out of sample. Exposure to climate policy risk is not associated with higher returns over the sample. In fact there is evidence for a negative premium on climate policy risk exposure. Preference shocks are considered as a candidate explanation for this discrepancy with theory. Empirically identified preference shocks are shown to have explanatory power for the spread between low and high risk exposure returns, explaining a similar amount of the variation as policy shocks. Further investigation of the consequences of changing aggregate demand for green products and assets is a promising avenue for future research.

References

Ardia, David, Keven Bluteau, Kris Boudt, and Koen Inghelbrecht. 2023. "Climate Change Concerns and the Performance of Green Versus Brown Stocks." *Management Science* 69 (12).

Arias, Jonas E, Juan F Rubio-Ramírez, and Daniel F Waggoner. 2018. "Inference Based on SVARs Identified with Sign and Zero Restrictions: Theory and Applications." *Econometrica* 86 (2).

Aswani, Jitendra, Raghunandan Aneesh, and Rajgopal Shiva. 2023. "Are Carbon Emissions Associated with Stock Returns?" *Review of Finance*.

Bansal, Ravi, and Amir Yaron. 2004. "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles." *The Journal of Finance* 59 (4).

Barnett, Michael. 2023. "A Run on Fossil Fuel? Climate Change and Transition Risk." *Working Paper*.

Bauer, Michael D, Daniel Huber, Glenn D Rudebusch, and Ole Wilms. 2022. "Where is the Carbon Premium? Global Performance of Green and Brown Stocks." *Journal of Climate Finance*.

Baz, Salim, Alex Michaelides, and Yi Zhang. 2022. "Firm-level climate regulatory exposure." *Working paper*.

van der Beck, Philippe. 2021. "Flow-Driven ESG Returns." *SSRN Electronic Journal*.

Bistline, John, Neil Mehrotra, and Catherine Wolfram. 2023. "Economic Implications of the Climate Provisions of the Inflation Reduction Act." *NBER working paper* (31267).

Blasberg, Alexander, Ruediger Kiesel, and Luca Taschini. 2021. "Carbon Default Swap - Disentangling the Exposure to Carbon Risk Through CDS."

Bolton, Patrick, and Marcin Kacperczyk. 2021. "Do Investors Care about Carbon Risk?" *Journal of Financial Economics* 142 (2): 517–549.

Bolton, Patrick, and Marcin Kacperczyk. 2023. "Global Pricing of Carbon-Transition Risk." *Journal of Finance* 78 (6).

Carney, Mark. 2015. "Breaking the tragedy of the horizon - climate change and financial stability."

Cassidy, William. 2023. "Elections Have Consequences: The Impact of Political Agency on Climate Policy and Asset Prices." *Working Paper*.

Davis, Steven J, Stephen Hansen, and Cristhian Seminario-Amez. 2021. "Firm-level Risk Exposures and Stock Returns in the Wake of COVID-19." *Working Paper*.

Enders, Arthur, Thomas Lontzek, Karl Schmedders, and Marco Thalhammer. 2023. "Carbon Risk and Equity Prices." *Available at SSRN 4476587*.

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel. 2020. "Hedging Climate Change News." *The Review of Financial Studies* 33 (3).

Faccini, Renato, Rastin Martin, and George Skiadopoulos. 2022. "Dissecting Climate Risks: Are They Priced in the US Stock Market?" *Journal of Banking and Finance*.

Fry, Renee, and Adrian Pagan. 2011. "Sign Restrictions in Structural Vector Autoregressions: A Critical Review." *Journal of Economic Literature* 49 (4).

Gentzkow, Matthew, Jesse Shapiro, and Matt Taddy. 2019. "Measuring polarization in high-dimensional data: Method and application to congressional speech." *Econometrica* 87 (4).

Giglio, Stefano, Matteo Maggiori, Krishna Rao, Johannes Stroebel, and Andreas Weber. 2021. "Climate Change and Long-Run Discount Rates: Evidence from Real Estate." *The Review of Financial Studies* 34 (8): 3527–3571.

Hassan, Tarek A, Stephan Hollander, Ahmed Tahoun, and Laurence van Lent. 2019. "Firm-Level Political Risk: Measurement and Effects." *The Quarterly Journal of Economics* 134 (4).

Hengge, Martina, Ugo Panizza, and Richard Varghese. 2023. "Carbon Policy Surprises and Stock Returns: Signals from Financial Markets." *IMF working paper*.

Hong, Harrison G, Frank Weikai Li, and Jiangmin Xu. 2019. "Climate Risks and Market Efficiency." *Journal of Econometrics* 208 (1).

Hsu, Po-Hsuan, Kai Li, and Chi-Yang Tsou. 2023. "The Pollution Premium." *The Journal of Finance*

78 (3).
Huynh, Thanh D, and Ying Xia. 2021. "Climate Change News Risk and Corporate Bond Returns." *Journal of Financial and Quantitative Analysis* 56 (6).
Ilhan, Emirhan, Zacharias Sautner, and Grigory Vilkov. 2022. "Carbon Tail Risk." *The Review of Financial Studies* 34 (3).
IPCC. 2021. "Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.."
Känzig, Diego. 2023. "The Unequal Consequences of Carbon Pricing." *Working Paper*.
Li, Qing, Yuehua Tang, and Vincent Yao. "Corporate Climate Risk: Measurements and Responses." *The Review of Financial Studies*.
Peltzman, Sam. 1976. "Toward a more general theory of regulation.." *Journal of Law and Economics* (19).
Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor. 2021. "Sustainable investing in equilibrium." *Journal of Financial Economics* 142 (2): 550–571.
Pástor, Ľuboš, Robert F. Stambaugh, and Lucian A. Taylor. 2022. "Dissecting green returns." *Journal of Financial Economics* 146 (2).
Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. 2023. "Firm-Level Climate Change Exposure." *The Journal of Finance* 78 (3): 1449–1498.
Sautner, Zacharias, Laurence Van Lent, Grigory Vilkov, and Ruishen Zhang. 2024. "Pricing Climate Change Exposure." *Management Science*.
Seltzer, Lee H., Laura Starks, and Qifei Zhu. 2022. "Climate Regulatory Risk and Corporate Bonds." *NBER working paper* 29994.
Taddy, Matt. 2013. "Multinomial Inverse Regression for Text Analysis." *Journal of the American Statistical Association* 108 (503). Number: arXiv:1012.2098 arXiv:1012.2098 [stat].
Uhlig, Harald. 2005. "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure.." *Journal of Monetary Economics* 52 (2).
UNEP. 2023. "Emissions Gap Report 2023."
Zhang, Shaojun. 2023. "Carbon Returns Across the Globe." *Working Paper*.

Appendix A. Model Solutions

The return on an equity is given by by:

$$R_{i,t,t+1} = \frac{P_{i,t+1}}{P_{i,t}}.$$

Consequently, the log return is simply:

$$r_{i,t,t+1} = \ln P_{i,t+1} - \ln P_{i,t} = p_{i,t+1} - p_{i,t} + \Delta d_{t+1}$$

The adjustment for Δd_{t+1} is necessary as p denotes the log price-dividend ratio, not just the log price. The model is solved to find the log price-dividend ratio:

$$p_{i,t} = a_i + b_i x_t + e_i y_t$$

Where the expressions of the coefficients in terms of primitives are as follows:

$$\begin{aligned}
a_i &= \ln \delta - \gamma\mu + \mu_d + \ln\left[1 + \bar{\lambda}\left(\exp\left\{\left(\gamma - \eta_i\right)\xi\right\} - 1\right)\right] \\
b_i &= -\gamma \\
e_i &= 1
\end{aligned}$$

Therefore, substituting for the log price-dividend ratio and for dividend growth in returns:

$$r_{i,t,t+1} = -\left[a_i + b_i x_t + e_i y_t\right] + \left[\mu_d + y_t - \eta_i J_{t+1}\right].$$

For the expected return of the stock, we have:

$$\begin{aligned}
E_t\left[R_{i,t,t+1}\right] &= E_t\left[\exp\left\{r_{i,t,t+1}\right\}\right] \\
&= \exp\{-\left[a_i + b_i x_t + e_i y_t + \left[\mu_d + y_t\right]\right\} \times \left[1 + \lambda\left(\exp\left\{-\eta_i\right)\xi\right\} - 1\right)],
\end{aligned}$$

where the last line follows from J_{t+1} only taking value $\xi \in (0, 1)$ with probability λ and zero otherwise.

$$E_t\left[R_{i,t,t+1}\right] - R^f_{t,t+1} \simeq \gamma \operatorname{Cov}_t\left[r_{i,t,t+1}, \Delta c_{t+1}\right].$$

Substituting for consumption growth and the log return and dropping constant terms:

$$\gamma \operatorname{Cov}_t\left[r_{i,t,t+1}, \Delta c_{t+1}\right] = \gamma \operatorname{Cov}_t[-\eta_i J_{t+1}, -J_{t+1}].$$

Since $\operatorname{Var}_t\left[J_{t+1}\right] = \xi^2 \lambda (1 - \lambda)$, we obtain:

$$\gamma \operatorname{Cov}_t[-\eta_i J_{t+1}, -J_{t+1}] = \gamma\left[\eta_i\right]\xi^2\lambda(1 - \lambda),$$

and therefore:

$$E_t\left[R_{i,t,t+1}\right] - R^f_{t,t+1} \simeq \gamma\eta_i\xi^2\lambda(1 - \lambda)$$

Appendix B. Data

B.1. Sample

The following are the details on how we construct our analysis sample:

- 5,239 firms are identified with return data on at least three of the climate policy announcement dates
- 23 firms with no leverage information are removed
- In order to compute abnormal returns, I first need to get estimates of stock- level betas.I keep stocks for which there are at least 125 daily return observations for the year prior to at least three of the jump days. I lose 18 firms in this step.
- I remove 7 firms with no available NAICS2 code in the dataset.
- Finally, I link these firms with the text data from form 10-K. I end up with a sample of 3,894 stocks with a non-empty 10-K.

B.2. Cross-sectional Controls

Financial data for cross-sectional regressions are obtained from Compustat.

The control variables are as follows:

- Market Cap is the natural logarithm of market capitalization (price times shares outstanding) at the end of year.
- Book to Market is the book value divided by market cap and the end of the year.
- Leverage is the book leverage of the company.
- Invest/a represents the firm's capital expenditures divided by the book value of its assets.
- HHI Is the Herfindahl concentration index of firms with respect to industry.
- PPE is the natural logarithm of the firm's property, plan and equipment
- SALESGR is the dollar change in annual firm revenues normalized by last month's market capitalization.
- EPSGR is the dollar change in annual earnings per share normalized by the firm's equity price.
- beta is the market beta of the firm calculated over the one year period using daily data.
- ROE is the firm's earnings performance, given by the ratio of its net yearly income divided by the value of its equity.
- MOM is the average of the most recent 12 months' returns on the stock.

- VOL is the standard deviation of returns based on past 12 months of monthly returns.

Following Bolton and Kacperczyk (2021), to eliminate the impact of outliers, I winsorize Book to Market, Leverage and Invest/a at the 2.5% level and MOM, VOL, SALESGR and EPSGR at the 0.5% level. Summary statistics for the variables are in Table A1.

	Mean	Median	sd
Market Cap	6.7	6.65	10.8
Book to Market	0.633	0.489	0.541
Leverage	0.371	0.342	0.316
Invest/a	0.034	0.018	0.052
HHI	0.0	0.342	0.316
PPE	6.612	5.598	8.236
SALESGR	-0.021	0.0176	0.448
EPSGR	-0.016	0.004	0.583
beta	1.048	1.026	0.821
ROE	-0.061	0.073	0.585
MOM	0.014	0.011	0.065
VOL	0.137	0.102	0.133

TABLE A1. Summary Statistics

B.3. Text preprocessing

I follow Davis, Hansen, and Seminario-Amez (2021) in processing the text from the 10Ks. First, 10Ks from 2011-2022 are scraped from EDGAR. Terms are created as follows:

- 4,237 phrases correspond to named entities that appear more than 25 times in the corpus. These are identified with the named entity recognizer (NER) from the Stanford NLP group.
- 26,729 additional multi-word expressions (MWE) are identified. In order to identify these, all words in the corpus are tagged using a part-of-speech tagger from the Stanford NLP group. Tag patterns likely to correspond to meaningful sentences are then tabulated. The final set of MWE is the set of resulting trigrams that appear more than 150 times in the corpus, and bigrams that appear more than 500 times.

The following standard steps complete the preprocessing:

- Lowercase all text (case-folding)

- Tokenize text by breaking it into individual terms
- Drop common words from a standard stop word list
- Drop any terms that appear in the Risk Factors text over fewer than 25 firms over the sample

  After preprocessing there are 21,876 unique terms.

B.4. Climate Policy Announcement Dates

This table shows the 38 candidate climate policy dates from the initial 53 which are not included in the final set. The third column specifies the reason for which the candidate date was removed. The options are: no news coverage, confounded by other announcements, insignificant return response from brown industries.

TABLE A2. Candidate Climate Policy Announcement Dates Removed

Announcement	Date	Reason Removed
Biden Executive Order to reduce ghg from power plants	11/05/23	No Coverage
Green New Deal reintroduced	20/04/23	Insignificant
Biden Executive Order to reduce air pollution	12/04/23	No Coverage
IPCC sixth assessment report	20/03/23	Insignificant
Biden Executive order to require federal contractors to disclose climate impact	14/11/22	No Coverage
House passage of IRA	12/8/22	Insignificant
Announcement of end of Manchin-Schumer negotiation	14/7/22	Cofounded
Manchin announces decision to vote against BBB	19/12/21	Confounded
Biden revokes Keystone Pipeline permit	09/06/21	Insignificant
Moratorium on new oil leases on federal land	27/01/21	Insignificant
Actual withdrawal from Paris agreement	04/11/20	Cofounded
Wheeler confirmed as EPA head	12/04/18	Insignificant
Trump plans to cut Obama's coal emissions standards	21/08/18	Confounded
Trump weakens fuel economy standards	24/08/18	Confounded
IPCC special report on 1.5 degrees	08/10/18	Insignificant
Trump announces US will leave Paris	01/06/17	Insignificant
Trump executive order reviving pipeline plans	24/01/17	Insignificant
EPA budget cut	01/03/17	Insignificant
Trump proposes Wheeler to EPA	26/07/17	Insignificant
Pruitt begins CPP appeal	09/10/17	Cofounded
US and China agree to ratify Paris	03/09/16	No Coverage
Trump nomination of Pruitt to EPA	07/12/16	Insignificant
Supreme Court blocks CPP	09/02/16	Insignificant
House rejects Obama's climate regulations	01/12/15	Insignificant
Senate blocks Obama's climate regulations	17/11/15	Cofounded
Obama vetoes push to kill climate regulations	24/02/15	Insignificant
Obama vetoes Keystone XL pipeline	24/02/15	Insignificant
Obama rejects Keystone XL pipeline	06/11/15	Cofounded
Bilateral climate agreement with Brazil	30/06/15	Cofounded
EPA announces rules on power plants	02/06/14	No Coverage
Obama announces CPP	18/06/14	Insignificant
US-China deal on carbon cuts	11/11/24	Cofounded
Third national climate assessment	06/05/14	Insignificant
Obama announcement of Climate Action Plan	25/06/13	Insignificant
Warsaw COP walkouts	20/11/13	Cofounded
IPCC report	27/09/13	Insignificant
Obama executive order for stricter fuel economy standards	15/10/12	Insignificant
Doha COP	08/12/12	Insignificant

Appendix C. Results

C.1. Emissions

TABLE A3. Emissions Regression Results

	Dependent Variable: Emissions		
	Emissions in Tonnes	Emissions Intensities	Reported Emissions
	(1)	(2)	(3)
z	$1,163,635.000^{***}$	124.047^{***}	$1,533,491.000^{***}$
	(144, 015.300)	(32.851)	(336, 516.500)
Constant	$1,061,955.000^{***}$	174.999^{***}	$2,323,302.000^{***}$
	(132, 421.900)	(31.582)	(367, 936.300)
Observations	2,149	1,878	797
R^2	0.030	0.008	0.025
Adjusted R^2	0.029	0.007	0.024
Residual Std. Error	6,121,662.000 (df = 2147)	1,359.725 (df = 1876)	9,756,542.000 (df = 795)
F Statistic	65.285^{***} (df = 1; 2147)	14.259^{***} (df = 1; 1876)	20.766^{***} (df = 1; 795)

Note: This table presents the results of the cross sectional regression with the dependent variable being emissions and the independent variable emissions. Column 1 presents the case where the dependent variable is Refinitiv's estimate of each firm's total emissions in Tonnes. Column 2 instead takes emissions divided by revenue (in millions of dollars). Column 3 restricts the sample to only those firms who directly report their emissions. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	Mean	Median	sd
Emissions In Tonnes (Estimated)	1,141,675	17,392	6,212,592
Emissions scaled by Revenue (In Millions)	188.58	17.83	1364.52
Reported Emissions	2,898,664	164,018	9,876,934

This table presents summary statistics for the emissions data from Refinitiv employed in the cross sectional regressions.

TABLE A4. Summary Statistics

C.2. Predictive Power

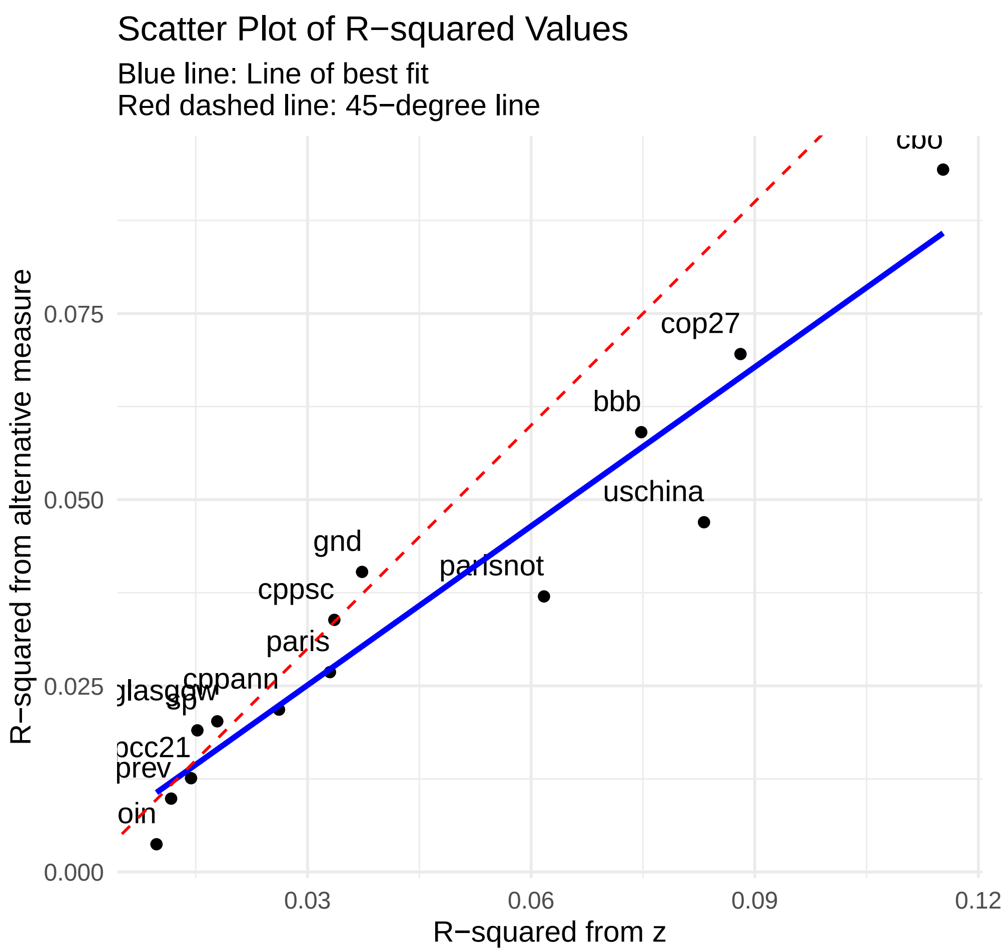

FIGURE A1

This figure shows the R-squared values from a regression of abnormal returns on exposure score, market cap, and industry. The x axis shows the R squared from the regression using my exposure score z. The y axis uses the climate regulatory risk exposure score from Sautner et al. (2023)

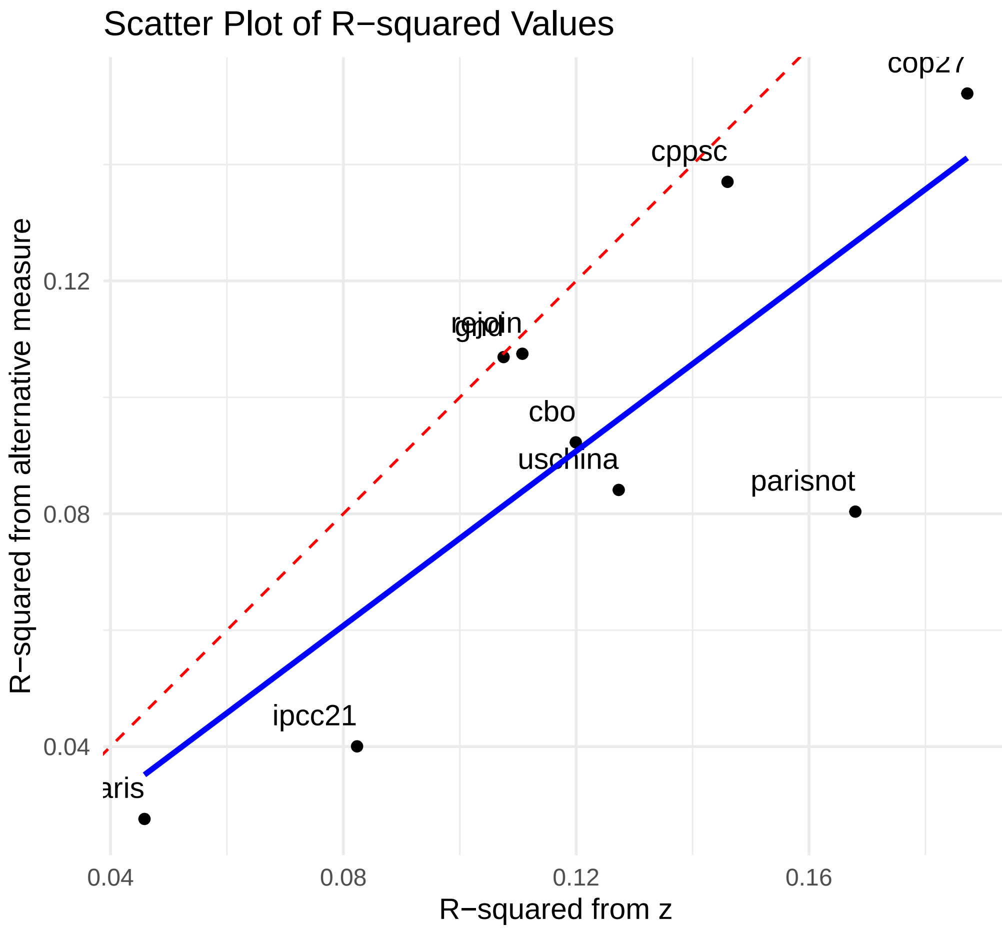

FIGURE A2. Predictive power of exposure score trained on a subset of the announcement dates.

This figure shows the R-squared values from a regression of abnormal returns on exposure score, market cap, and industry. The x axis shows the R squared from the regression using my exposure score z trained on a randomly selected subset of the climate policy announcement dates and tested on the other half. The y axis uses the climate regulatory risk exposure score from Sautner et al. (2023).

C.3. Sign Restriction SVAR

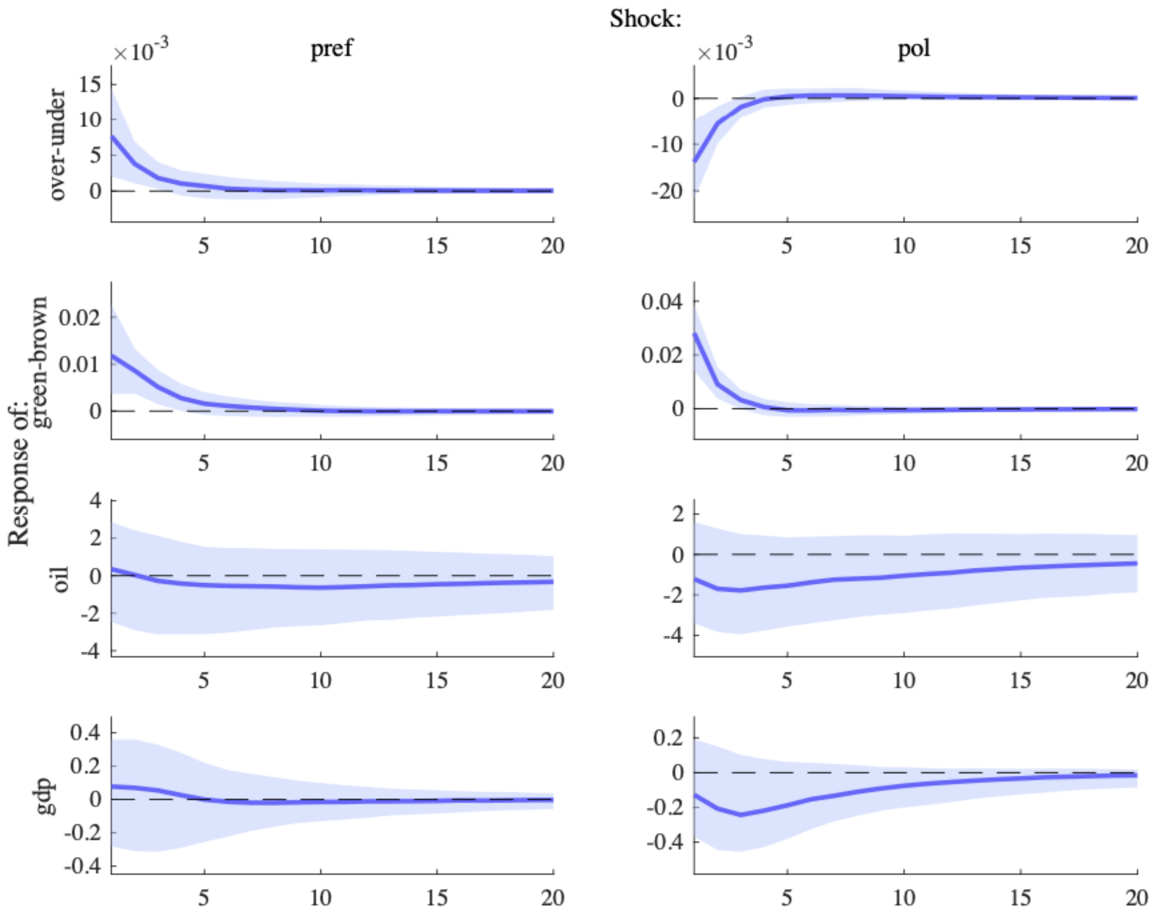

FIGURE A3. SVAR IRFs

Note: This figure presents the IRFS from the policy and preference shocks identified via the sign restriction SVAR

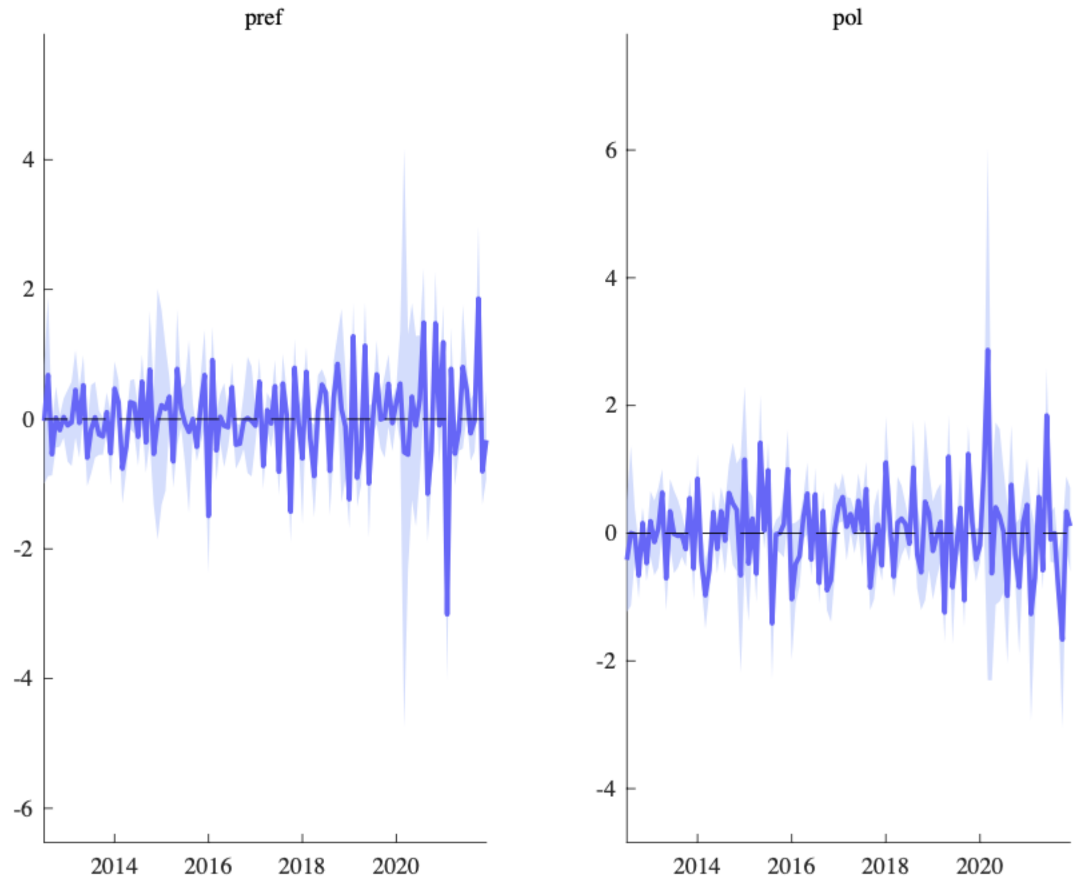

FIGURE A4. Structural Policy and Preference Shocks

This figure presents the structural policy and preference shocks identified by the sign restriction VAR