RESEARCH

Swiss Finance Institute
Research Paper Series
N°25-45

Real-Time Climate Controversy Detection

David Jaggi
University of Zurich

Nicolas Jamet
RAM Active Investment

Markus Leippold
University of Zurich and Swiss Finance Institute

Tingyu Yu
University of Zurich

Real-Time Climate Controversy Detection*

David Jaggi, Nicolas Jamet, Markus Leippold, and Tingyu Yu

March 2025

Abstract

This study presents ClimateControversyBERT, a novel open-source language model for real-time detection and classification of corporate climate controversies (i.e., brown projects, misinformation, ambiguous actions) from financial news. Validated using RepRisk and Refinitiv metrics, the model effectively identifies inconsistencies between corporate climate commitments and actions as they emerge. We document significant negative market reactions to these controversies: firms experience an immediate average stock price drop of 0.68%, with further declines over subsequent weeks. The impact is intensified by high media visibility and is notably stronger for firms with existing emission reduction commitments, underscoring the market's penalty for perceived environmental failures.

Keywords: Climate controversy, corporate greenwashing, natural language processing.

JEL classification: G14, Q54, C88

*David Jaggi is from the Department of Finance, University of Zurich, and Zurich University of Applied Sciences (david.jaggi@zhaw.ch); Nicolas Jamet is from RAM Active Investment (nja@ram-ai.com); Markus Leippold is from the Department of Finance, University of Zurich, and Swiss Finance Institute (markus.leippold@df.uzh.ch); and Tingyu Yu is from the Department of Finance, University of Zurich (tingyu.yu@df.uzh.ch). We are grateful for comments from seminar participants at the University of Zurich and RAM AI. We thank Tanguy Renaudie for excellent research assistance. This research has been funded by the Innosuisse Project "Investor and Stakeholder Tools for Tracking Companies' Climate Commitments, Greenwashing and ESG Trends" (Innolink: 100.562 IP-SBM) and the Swiss National Science Foundation (SNSF) under Grant Agreement No. 207800. David Jaggi was granted funding from the DIZH in the form of a DIZH Fellowship.

1 Introduction

Climate change poses a financial risk that affects corporate strategies and investor decisions. Firms, in response, increasingly make climate commitments to signal credible transition plans and meet stakeholder expectations. However, many fail to meet these targets or misrepresent their progress. Such discrepancies can result in significant reputation and financial damage.

In October 2021, ExxonMobil's CEO was accused of misleading Congress about the company's understanding of climate science (McGreal, 2021; Supran, Rahmstorf, & Oreskes, 2023). Although ExxonMobil's scientists accurately projected global warming, the company's public statements contradicted these internal findings. Investors who relied on Exxon's external claims without independent verification are likely to misprice the firm's climate-related risks. A similar case emerged in Volkswagen's 2015 emissions scandal. The company secretly installed defeat devices—software designed to cheat emissions tests—misleading regulators, investors, and consumers (Bouzzine & Lueg, 2020). After exposed, Volkswagen's stock price fell over 30%, and the firm incurred billions in legal costs. These examples underscore that climate controversies can materially harm firm value.

Despite its importance, traditional third-party sustainability ratings offer limited visibility into such events. They often rely on self-reported data or information updated with a considerable lag.[1] As a result, important signals about a firm's climate-related behaviors may be either overlooked or recognized only after a significant delay. Investors and regulators thus require timely and objective tools to detect climate incidents as they unfold in financial news, enabling them to respond before reputational harm and market losses escalate. Moreover, in light of continued global interest in climate change—and notwithstanding the political opposition to corporate climate initiatives in some quarters—there is persistent pressure for more transparent and accountable reporting of environmental performance. Firms operating in jurisdictions with strong anti-environmental activism sentiment may face conflicting pressures to de-emphasize climate commitments publicly, yet their financial stakeholders still demand accurate information about environmental risks.

In this paper, we propose a novel approach to address these challenges by introducing and open-sourcing a specialized pipeline based on ClimateBERT (Webersinke et al., 2021) and fine-

[1] Frameworks like ESG Controversies Methodology of MSCI (2024) assess firms' alleged contributions to climate change but do not offer automated tools for real-time detection.

tuned classification layers. This helps us to systematically detect and categorize climate-related controversies from unstructured financial news at scale. Our pipeline follows a multi-stage process. First, a relevance filter identifies climate-related paragraphs from millions of news articles, and an entity extraction module links them to specific companies. Second, a classifier detects whether the paragraph describes a controversy. Third, a category classifier assigns the controversy to one of three predefined types. This architecture ensures both precision and scalability while maintaining domain specificity. By examining paragraphs over a time span of two decades, our method identifies climate controversies worldwide, pinpointing real-world corporate actions that may contrast with public commitments. We focus on controversies in three categories: *brown projects*, *misinformation*, or *ambiguous actions*. By capturing these distinct types of negative events, our approach could uncover misalignment between stated goals and actual behavior in a wide range of corporate contexts.

Given that our pipeline relies predominantly on news articles, the credibility of text sources is essential. As Das (2011) notes, outlets like Dow Jones undergo rigorous fact-checking, ensuring relatively high-quality information that is more likely to influence investor perceptions. Meanwhile, blogs or social media platforms, though providing rapid coverage, often lack professional editorial standards (Jong & van der Linde, 2022; Kleinnijenhuis et al., 2013, 2015; Yu et al., 2023). In light of these insights, our data focus on professionally curated news articles to reduce noise from potentially unreliable outlets.

We analyze more than 15 million news articles covering firms across 88 countries. The time-series variation of detected climate controversies closely aligns with major environmental events and regulatory shifts. A sharp increase in *brown projects* appeared in 2010, coinciding with the Deepwater Horizon oil spill, one of the most severe environmental disasters. *Misinformation* rises notably around 2015, likely driven by the Volkswagen emissions scandal. In 2021, a wave of corporate net-zero pledges followed major climate summits and investor pressure. However, a simultaneous rise in *ambiguous actions* suggests many pledges lacked clear implementation plans, reinforcing concerns about greenwashing. At the industry level, climate controversies are highly concentrated in sectors with direct environmental impacts, like Petroleum and Natural Gas, followed by the Utilities and Automobiles & Trucks sectors.

We validate our detected controversies using several third-party measures, including RepRisk incidents, Refinitiv ESG controversies, and the Media Climate Change Concern index from Ardia et al. (2023). All show strong correlations with our measures, confirming alignment between traditional human-curated ESG monitoring and our automated pipeline. Moreover, our approach offers greater granularity, capturing specific climate controversies and related risks that structured ESG databases often overlook.

A key question is whether and to what extent climate controversies affect firm value. If markets perceive these events as material risks, they should be reflected in stock prices, providing insight into the financial consequences of failing to meet climate expectations. Several economic channels can drive this impact. First, reputational costs may erode consumer trust and weaken demand, particularly given that sustainability concerns have become increasingly important to customers. Firms with damaged environmental credibility may lose partnerships or face investor divestment, further straining their growth prospects. Second, regulatory or legal risks can result in fines, lawsuits, or compliance costs if authorities or activist groups challenge a firm's climate claims. Increased scrutiny may also trigger stricter regulations and force firms to make costly operational adjustments.

To test the value effects, we conduct event studies analyzing stock market reactions to the detected climate controversies. Stock prices decline by 0.68% on average immediately following news coverage, with a t-statistic of 4.7, indicating that investors view these events as financially material and incorporate the information quickly. The negative return deepens to 1.27% over twenty days, significant at the 1% level. Among controversy types, *ambiguous actions* have the weakest impact, with negative coefficients but not significant. *Brown projects* and *misinformation* trigger stronger market responses. The effect is more pronounced for firms with emissions reduction targets, likely because such controversies undermine their credibility, signaling weak governance or poor implementation of related climate commitments.

The economic significance of our approach extends beyond academic interest to practical market applications. For institutional investors, our controversy detection system provides an early warning mechanism for portfolio risk management before traditional ESG metrics capture these events. Our findings suggest that markets price climate controversies with varying efficiency—some trigger immediate price corrections while others reveal gradual valuation impacts. For corporate managers, our results quantify the financial cost of climate-related missteps, offering a concrete business case for stronger environmental governance. Regulatory bodies may leverage similar methodologies to enhance their surveillance capabilities, particularly as disclosure requirements like the EU's Sustainable Finance Disclosure Regulation (SFDR) and the SEC's proposed climate disclosure rules intensify the need for verification mechanisms. By reducing informational asymmetries between firms and stakeholders, tools like ours could improve market efficiency while lowering the cost of capital for genuinely sustainable enterprises. Our research underscores the potential for advanced natural language processing to enhance transparency in corporate climate behavior and facilitate a more reliable transition toward a low-carbon economy.

Related literature: We contribute to the growing body of work using domain-specific language models in finance, particularly for climate-related analysis. While standard BERT architectures (Vaswani et al., 2017) and derivatives like RoBERTa (Liu et al., 2019) have transformed general text classification, specialized models such as ClimateBERT (Webersinke et al., 2021) are increasingly important when analyzing climate-focused corpora. A variety of applications demonstrate its effectiveness, including the exploration of climate disclosures and their influence on credit spreads (Kölbel et al., 2024), the detection of net zero targets (Schimanski et al., 2023), the classification of climate risk factors (Garrido-Merchán, González-Barthe, & Vaca, 2023; Varini et al., 2020), the classification of corporate disclosures (Bingler et al., 2022), and environmental claims (Stammbach et al., 2023). By employing a domain-specific model and tailoring it to detect climate controversies, we further demonstrate how these techniques can uncover actionable insights at the intersection of environmental commitments and financial outcomes. Our study underscores the value of adapting natural language processing methodologies to specialized domains, aligning with the observation that domain-specific models outperform general approaches in climate-related tasks (Webersinke et al., 2021).

We also add to the growing literature on quantifying climate risk from news. Engle et al. (2020) estimate market-level climate risk by analyzing the similarity between Wall Street Journal content and a predefined climate vocabulary. Ardia et al. (2023) develop a Media Climate Change Concern Index based on climate-related reporting in major U.S. newspapers. In parallel, a growing body of research applies textual analysis to assess firm-specific exposure to climate risks (Baz et al., 2023; Bingler et al., 2024; Giglio et al., 2023; Kölbel et al., 2024; Li et al., 2024; Sautner et al., 2023).

Lastly, we relate to the literature by examining the value implication of various climate exposures (Bolton & Kacperczyk, 2021; Hsu, Li, & Tsou, 2023; Ilhan, Sautner, & Vilkov, 2020; Kruttli, Roth Tran, & Watugala, 2023; Pástor, Stambaugh, & Taylor, 2022).

The remainder of this paper is structured as follows. Section 2 describes the data source. Section 3 explains the detection and classification of climate controversies. Section 4 provides validation tests. In Section 5, we empirically show how these controversies affect firm value. Section 6 concludes with implications for corporate governance and investment strategy.

2 Data

Our analysis draws upon multiple complementary data sources. The core of our study is a large corpus of news articles from Dow Jones Newswires, spanning 2004 to 2023. This dataset, comprising over 15 million articles covering companies across 88 countries, serves as the foundation for our climate controversy detection model. It includes rich metadata, such as company identifiers, subject codes, and timestamps. To further refine the attribution of news to specific firms, we apply Named Entity Recognition at the paragraph level to link company mentions more accurately.

Additionally, to validate our controversy measures, we incorporate incident data from RepRisk, ESG Controversy scores from Refinitiv, and the Media Climate Change Concern Index from Ardia et al. (2023). For firm value analysis, we obtain stock price data from CRSP, financial information from Compustat, as well as corporate emission targets from MSCI.

3 Climate Controversy Detection

This section details our sequential natural language processing pipeline designed to detect and classify climate-related controversies in Dow Jones news articles. The pipeline begins by identifying relevant paragraphs (Step 1). It then employs a fine-tuned model to detect the presence of controversies (Step 2), followed by a second fine-tuned model to classify the specific controversy type (Step 3) into brown projects, misinformation, or ambiguous actions. We also present the rationale for the pipeline's two-stage architecture and explain how detected controversies are assigned to specific companies. A detailed account of the annotation data preparation crucial for training these models is given in Appendix A.

3.1 Step 1: Identifying Climate-Related Paragraphs

To more effectively identify relevant news, we exclude articles with fewer than 100 words. We then retain only those articles classified under the "N/ENV" or "N/CO2" subject codes or those containing explicitly predefined climate-related keywords.[2] While this filtering step effectively obtains climate articles, it also inadvertently captures irrelevant content or masks key climate-related information. To address this limitation, we apply ClimateBERT (Webersinke et al., 2021) to classify each paragraph as climate-related or not. Pretrained on over two million climate paragraphs sourced from news articles, research abstracts, and corporate reports, ClimateBERT significantly improves classification accuracy by enabling nuanced, context-aware analysis beyond simple keyword searching. Its domain-adaptive training particularly enhances performance on climate-specific language, reducing errors in various downstream tasks such as text classification and sentiment analysis. By leveraging ClimateBERT in our pipeline, we collect all climate-related paragraphs, ensuring a reliable dataset for subsequent detection of climate controversies.

3.2 Step 2: Detecting Climate Controversies

Next, we implement our ClimateControversyBERT,[3] a fine-tuned model designed to detect controversial content in climate-related paragraphs. This is a crucial step to identify the different types of controversies further, as the occurrence of controversies in news articles is sparse, and directly detecting different types of controversy in the news would leave us with a very unbalanced dataset. More precisely, we fine-tune the ClimateBERT model using 2,199 annotated sentences extracted from online sources, of which 790 mention climate controversies. Given the imbalance in our dataset, we

[2]We employ a comprehensive list of climate-related keywords, including "climate change," "global warm," "carbon footprint," "greenhouse gas," "carbon emission," "sea level," "renewable energy," "fossil fuel," "extreme weather," "air pollution," "melting glaciers," "climate action," "energy efficiency," "sustainable development," "conservation effort," "carbon neutral," "carbon sequestration," "climate policy," "climate adaptation," "climate mitigation," and "oil spill." This keyword-based approach ensures broad coverage while maintaining a focus on climate-relevant content.

[3]The ClimateControversyBERT model and associated code are open-sourced and available on Hugging Face https://huggingface.co/climatebert/ClimateControversyBert.

apply weighted class adjustments to improve classification performance.[4] This approach gives more importance to minority classes during training. ClimateControversyBERT achieves an F1-score of 88%, as shown in Table B1.[5] Applying the model to all climate-related paragraphs from earlier steps, we identify only 38,615 (1.2%) out of 1,408,588 paragraphs as controversial, highlighting both the sparsity of climate controversies in news coverage.

3.3 Step 3: Differentiating Controversy Types

Climate-related controversies in corporate news disclosures are complex and multifaceted. Recognizing distinct categories that differentiate controversies based on corporate intent, behavior, and environmental consequences is crucial. We thus propose a classification with three categories.

First, *brown projects* address controversies where corporations directly contribute to environmental harm or fund environmentally damaging projects. Examples include fossil fuel extraction, deforestation-linked agricultural practices, and infrastructure projects with significant ecological consequences. These cases typically reflect profit-driven environmental costs. Second, *misinformation* covers intentional or misleading disclosures where firms manipulate environmental data to present a more favorable sustainability performance. Examples include inflated fuel efficiency claims, understated emissions, or misleading carbon offset disclosures. As transparency becomes increasingly critical in financial markets, identifying misinformation is essential for preserving data integrity and maintaining investor trust. Third, *ambiguous actions* capture controversies where corporate climate-related efforts are insufficient, inconsistent, or misaligned with stated commitments or international standards. Examples include companies failing to make meaningful progress in reducing greenhouse gas emissions or implementing sustainability initiatives that lack credibility. Given growing concerns over greenwashing, this classification is crucial for assessing whether firms' climate initiatives align with investor expectations and regulatory standards.

To classify these controversy types, we fine-tune a second ClimateBERT model using annotated training sentences categorized by controversy type, following the same fine-tuning procedure as the

[4]Additionally, we obfuscate the organizations in the training sentences using named entity recognition (NER), as Son et al. (2023) show that including financial entity names in pre-trained language models might bias the performance of downstream tasks. We exclude frequently mentioned non-corporate entities, such as the International Energy Agency, United Nations, International Sustainability Standards Board, and Environmental Protection Agency.

[5]Table B1, Panel A, lists the hyperparameters used for fine-tuning.

first model. Panel B of Table B1 shows that the model achieves an F1 score of 83%, demonstrating good classification performance in distinguishing controversy types.

After applying our two fine-tuned ClimateBERT models to each paragraph in news articles, we identify those classified as climate controversies and assign them to specific controversy types. Table 1 shows that *brown projects* are the most common climate controversies (0.62%), followed by *misinformation* (0.37%) and *ambiguous actions* (0.21%, or 6,661 paragraphs). We consider a firm involved in a climate controversy if at least one news article mentioning the firm on that day contains a controversy, as identified by our model. The controversy type is assigned based on the dominant category among paragraphs in the relevant article.

3.4 Model Architecture Rationale

Following the identification of climate-related paragraphs, our pipeline employs a two-stage classification strategy first to detect the presence of controversy and then categorize its specific type. Step 1 utilizes our fine-tuned ClimateControversyBERT model for binary classification, determining whether a paragraph describes a controversy or not, as detailed in Section 3.2. Subsequently, Step 2 applies a second, similarly fine-tuned model exclusively to the paragraphs identified as controversial in the first stage. This second model classifies the controversy into one of three predefined types: brown projects, misinformation, or ambiguous actions, with the process described in Section 3.3.

We carefully considered alternative architectures, such as training a single, integrated 4-class model (incorporating 'No Controversy' alongside the three controversy types). However, we deliberately opted for the sequential, two-stage approach due to several compelling reasons, primarily driven by the nature of the data itself. The most significant factor is the inherent sparsity of climate controversies within financial news text. As evidenced in Table 1, controversial paragraphs represent only a small fraction (1.2%) of potentially relevant climate-related paragraphs. Training a single 4-class model on such a heavily imbalanced dataset poses considerable challenges. Standard training procedures could easily allow the model to achieve high superficial accuracy by simply predicting the overwhelmingly dominant 'No Controversy' class, leading to inadequate performance in identifying the rare but critically important controversy classes. Our two-stage method strategically isolates this difficult and highly imbalanced detection task first, enabling the initial model to focus specifically on recognizing any controversy signal present.

Furthermore, separating the detection and classification tasks allows each model to develop greater specialization and focus. The first model (Step 1) is trained solely to learn the distinguishing features between general climate discussion and any form of controversy. The second model (Step 2) then operates on a pre-filtered, higher-signal dataset containing only controversial paragraphs. This allows it to dedicate its full learning capacity to discerning the potentially subtle semantic differences between the three specific controversy types (brown projects, misinformation, and ambiguous actions). We anticipate that this focused learning yields superior performance on both the initial detection and the subsequent nuanced classification sub-tasks compared to a single model tackling both simultaneously across the highly skewed data distribution. This staged approach also facilitates more controlled training, evaluation, and optimization for each distinct task, allowing for targeted assessment of both controversy detection (recall) and categorization (precision across types). Therefore, our pipeline proceeds by first applying the controversy detection model (Section 3.2) and subsequently applying the controversy classification model (Section 3.3) only to those paragraphs flagged as controversial by the initial stage.

3.5 Assignment of Controversies to Companies

Accurately attributing climate controversies to specific firms presents methodological challenges, as news articles often reference multiple companies while the controversy may pertain to only a subset of them. While Dow Jones Newswire provides company mappings, these are at the article level and might not correspond to the firm mentioned in the controversy-related paragraph. Additionally, this feature was introduced only in 2012, leaving a large portion of our sample period uncovered.

To achieve paragraph-level precision, we apply Named Entity Recognition (NER), using the SequenceTagger from the Flair package, to extract company names from controversy paragraphs. These entities are then matched to public company names using RapidFuzz with a 90% threshold for partial ratio matching.[6] Such a rigorous approach ensures high attribution accuracy by requiring explicit company mentions within controversy paragraphs. However, this strict criterion may exclude

[6]RapidFuzz was selected for our entity matching task due to its superior performance advantages over traditional Levenshtein distance calculations (Akbik et al., 2019). A 90% threshold for partial ratio also provides an optimal balance between accuracy and coverage (Bachmann, 2024).

cases where firms are implied but not explicitly named.

Overall, we detect climate controversy at the article level and link them to the correspondent firms that were mentioned in the relevant paragraphs. The next section provides more details on information captured by these controversies and the validation of our method.

4 Validation Tests

This section examines the validity of our detected climate controversies. First, we provide descriptive analyses of these climate controversies in terms of excerpts from news articles, time-series variation, and industry distribution. Second, we validate our controversy measure using RepRisk data, a widely recognized industry benchmark for ESG-related incidents. Third, we investigate their relationship with existing ESG controversy scores using data from Refinitiv. Fourth, we examine the relationship between detected controversies and the Media Climate Change Concern index.

4.1 Controversy Descriptions

Table 2 presents examples of climate-related controversies identified by our NLP method. *Brown projects* involve companies contributing directly to high carbon emissions through continued investment in fossil fuel infrastructure. For instance, HSBC has been criticized for financing new coal-powered projects, which contradicts its public pledge to phase out coal. Goldman Sachs continues to invest in power plants that emit millions of tons of greenhouse gases. Similarly, Drax operates the UK's single largest CO2 emitter. *Misinformation* controversies involve deceptive or misleading claims about environmental performance. Volkswagen is a notable case, having been investigated by the U.S. Environmental Protection Agency (EPA) for falsifying emission test results. Deutsche Bank provides another example, with the Securities and Exchange Commission (SEC) investigating whether its asset management arm misrepresented its ESG assessments. Additionally, BNY Mellon paid $1.5 million to settle SEC charges that it misrepresented the ESG review of its investments. *Ambiguous actions* reflect mixed signals or inconsistent behavior regarding climate commitments. This is represented by companies such as ExxonMobil, where executives admit the need to refine their climate strategy while facing pressure from shareholders to set more ambitious targets.

Chevron demonstrates similar ambiguity, with mounting pressure from ESG investors to address scope 3 emissions. Amazon further illustrates this category by rejecting shareholder proposals for climate change disclosures while claiming it was already making certain disclosures on the topic. These examples underscore the diverse nature of climate-related controversies and demonstrate the effectiveness of our NLP method in capturing them.

Table B2 presents the time-series variation of climate controversies, closely aligning with major environmental events and regulatory shifts. From 2004 to 2008, global awareness of climate issues remained relatively low. A sharp spike in 2010 (4,972 controversies) coincides with the Deepwater Horizon oil spill, one of history's most severe environmental disasters, which fueled a surge in media coverage of brown projects. A notable rise in misinformation appeared around 2015, likely driven by the Volkswagen emissions scandal, where the company manipulated emissions data to misrepresent its environmental impact. The 2021 peak in controversies corresponds to a wave of corporate net-zero pledges following investor pressure for decarbonization. However, the simultaneous rise in ambiguous actions suggests that many of these pledges lacked clear implementation plans, reinforcing concerns over greenwashing.

At the industry level, climate controversies are highly concentrated in sectors with direct environmental impacts. Table 3 lists the top 20 industries with detected controversies, with Petroleum and Natural Gas leading at 11,438 cases, including 6,499 classified as brown projects. This reflects its role in fossil fuel extraction and high-emission activities. Utilities and Automobiles & Trucks also show significant controversy involvement. At the firm level, BP, ExxonMobil, and Chevron top the list (Table B3), each with over 1,000 controversies. Geographically, Table B4 shows that the United States leads in detected controversies, followed by the United Kingdom and Germany. The lower controversy counts in other countries likely reflect differences in media transparency and climate reporting standards rather than fewer actual incidents.

4.2 Comparison with RepRisk Incidents

In this subsection, we validate the climate controversies we detected using the RepRisk incident data. RepRisk leverages a broad set of global sources, including news reports and regulatory filings, to systematically identify corporate incidents related to environmental, social, and governance risks.

This makes it a valuable benchmark for assessing the robustness of our detection approach. We estimate the following Poisson regression:

$$Incidents_{i,t} = \alpha + \beta ClimateControversy_{i,t} + \gamma X_{i,t} + a_j + b_t + \epsilon_{i,t},$$

where $Incidents_{i,t}$ is the number of RepRisk incidents for firm i in year t. $ClimateControversy_{i,t}$ is the logarithm of climate controversies for firm i in year t. Control variables include the logarithm of all news articles mentioning the firm in the same year, the logarithm of assets, ROA, investment rate, leverage, tangibility, and sales growth. We consider industry and year-fixed effects. Standard errors are clustered by industry.

Table 4, Panel A, presents the results for overall climate controversies. Column (1) considers all RepRisk incidents, showing a statistically significant coefficient of 0.4 (at the 1% level). Interpreted using the Poisson model, this implies that a 100% (doubling) increase in detected controversies corresponds to a 49.2% increase in RepRisk incidents ($\exp(0.4) - 1 \approx 49.2\%$). The subsequent columns analyze more granular categories of RepRisk incidents. The coefficient remains similar when focusing on environmental incidents in Column (2), climate-related incidents in Column (3), and greenhouse gas-related incidents in Column (4). Given that many of our detected climate controversies involve misleading corporate communications, Column (5) examines this type of RepRisk incidents, which are defined as cases where firms present a misleadingly positive image that contradicts their actual actions. Doubling the number of detected climate controversies is associated with a 53.7% increase in such incidents. Column (6) further explores lobbying-related incidents where firms attempt to influence regulators to support corporate interests. The result is similar.

Panel B analyzes the three types of climate controversies. Brown projects show the strongest and most consistent associations across all RepRisk incident types, with statistically significant coefficients ranging from 0.31 to 0.38. This can be explained as direct environmental harm—such as fossil fuel extraction and high-emission industrial activities—which is highly visible and commonly highlighted in corporate risk assessments. Misinformation follows a different pattern, showing stronger links to misleading communication (0.24) and lobbying incidents (0.37) but weaker and often insignificant effects on climate-related incidents. This suggests that greenwashing and deceptive communications are less likely to be flagged in climate impact reports. Our classification effectively isolates communication-based controversies as a distinct dimension of climate-related corporate risk. Ambiguous actions show the weakest associations with RepRisk incidents. This may reflect their lower immediate environmental or financial impact, making them less likely to be recorded as explicit ESG risks. Overall, our approach adds granularity and depth to the analysis of corporate climate risks, extending beyond what structured ESG databases typically capture.

4.3 Comparison with Refinitiv ESG Controversies

Next, we compare our detected controversies with Refinitiv's ESG Controversies scores, which capture company actions that contradict ESG commitments or involve negative events. This serves as another external validation of our measures against a widely recognized third-party benchmark.

Table 5 presents the results. The dependent variable is the Refinitiv ESG Controversies score, ranging from 0 to 1, with higher scores indicating fewer controversies. We use the same independent variables and fixed effects as in Table 4. Column (1) shows a statistically significant negative coefficient on overall controversies. Column (2) indicates that the predictive power is driven by misinformation and ambiguous actions. Despite the strong alignment with Refinitiv scores, our approach offers key advantages. Unlike vendor data, which is updated periodically and may lag behind real-world events, our method enables near real-time detection of emerging controversies.

4.4 Comparison with Media Climate Change Concern Index

To further investigate how corporate controversies influence real-time public discourse on climate issues, we link our measure to the MCCC index, which tracks daily climate change coverage in the media. The index is further broken down into themes such as business impact, environmental impact, societal debate, and research. We regress the daily change in the MCCC index and its subcomponents on the daily change in detected climate controversies, assessing how new controversies shape the tone and focus of climate-related media coverage. Table 6, Panel A, reports the results for overall climate controversies. All coefficients are positive and significant at the 1% level, indicating that controversies are associated with increased media attention. The strongest effects appear in coverage related to business impact—suggesting that corporate climate controversies tend to trigger greater interest in how these events affect financial markets and business operations. Panel B explores the effects of controversy categories. Brown projects are most strongly linked to societal debate, while ambiguous actions are more closely associated with environment-focused coverage. Controversies uncovered by our method are not isolated signals; they appear to influence media focus toward specific concerns within the broader climate conversation.

5 Value Implications

5.1 Baseline Estimates

A key question is whether and how climate-related controversies impact firm values. If markets perceive these events as material risks, they should be reflected in stock prices, providing insight into the financial consequences of failing to meet climate expectations. Several economic channels can drive this impact. First, reputational costs may erode consumer trust and weaken demand, particularly given that sustainability concerns have become increasingly important to customers. Firms with damaged environmental credibility may lose partnerships or face investor divestment, further straining their growth prospects. Second, regulatory or legal risks can result in fines, lawsuits, or compliance costs if authorities or activist groups challenge a firm's climate-related claims. Increased scrutiny may also trigger stricter regulations and force firms to make costly operational adjustments. These adjustments can sometimes disrupt operations if firms must change production processes, abandon existing projects, or navigate supply chain constraints. Hence, we expect a decline in stock prices after the release of these climate controversies, reflecting an expected deterioration in future cash flows or an increased cost of capital.

To explore the value implications, we conduct event studies analyzing stock market reactions to climate controversy disclosures. We define an event as a firm date with at least two news articles related to climate controversies for the firm to reduce noise. When a single article mentions multiple types of controversy, we assign it to the category with the most paragraphs. This approach minimizes overlap and serves to distinguish the return effects associated with different types of controversy. Our dataset encompasses 2,330 events from 2004 to 2023, comprising 1,507 related to brown projects, 884 to misinformation, and 461 pertaining to ambiguous actions. We systematically examine stock price changes across multiple event windows, spanning from ten days prior to the event date to the event date itself, and extending to five, ten, fifteen, and twenty days post-event. To calculate abnormal returns, we employ market-adjusted returns as our primary metric.

Table 7 reports the estimates. In Panel A, Column (1) shows a negative return of 0.68% immediately after the controversy is reported in the news articles, with a t-statistic of 4.7. This indicates that investors view climate controversies as financially material, and the market rapidly incorporates this information. The magnitude of the negative returns continues to grow, reaching 1.27% twenty days later, significant at the 1% level. Panels B–D present results by controversy types. Investors are least concerned about ambiguous climate actions, as all coefficients are negative, but only one is marginally significant at the 10% level ten days after the event. In contrast, the other two types of controversies trigger stronger market reactions. Stock prices dropped 1.02% (0.92%) immediately following controversies related to misinformation (brown projects), with losses extending to 2.09% (1.85%) twenty days later.

5.2 Severity of Controversies

While the previous analysis provides insights into the overall market reaction to climate controversies, it treats all events equally, potentially masking variations in impact based on severity. In reality, more severe controversies are likely to attract greater media attention and broader reporting across multiple sources, leading to wider investor awareness and stronger stock price reactions. To test this, we regress cumulative abnormal returns (CARs) on the number of news articles covering each event. We estimate the following regression:

$$CAR_{i,t} = \alpha + \beta ClimateControversy_{i,t} + \gamma X_{i,t} + a_i + b_t + \epsilon_{i,t},$$

where $CAR_{i,t}$ represents cumulative abnormal returns over different event windows for firm i with climate controversy on time t. $ClimateControversy_{i,t}$ is the logarithm of the number of news articles reporting the controversy on the same date. Control variables include the logarithm of market capitalization, book-to-market ratio, ROA, investment rate, leverage, tangibility, and sales growth. We consider the firm and year-month fixed effect, with standard errors clustered by industry and year. Table 8 presents a negative association between controversy severity and CARs, with statistically significant effects persisting over all the tested intervals. The effect intensifies over longer event windows, suggesting that more severe controversies lead to prolonged market concerns. The effect is strongest for brown projects. In short, firms facing greater media scrutiny experience larger negative returns.

5.3 Role of Emission Reduction Targets

Lastly, we examine whether firms with emission reduction targets or net-zero commitments experience different stock price reactions to climate controversies. Firms that have publicly committed to reducing emissions are generally perceived as more environmentally responsible, attracting investors who prioritize sustainability. However, if such firms become involved in climate controversies, the market reaction may be more severe, as investors could interpret the event as a breach of trust or greenwashing, leading to greater reputational damage. To test this, we introduce an interaction term between the controversy variable and indicators for firms with emission targets. If markets penalize firms more for failing to uphold their commitments, we expect the interaction term to be negative and significant, indicating a stronger stock price decline for firms with climate pledges. Conversely, if investors view these firms as better positioned to manage climate risks, the negative effect of controversies might be mitigated, leading to a less severe market reaction. Specifically, we estimate the following regression:

$$CAR_{i,t} = \alpha + \beta_1 ClimateControversy_{i,t} + \beta_2 ClimateControversy_{i,t} \times EmiTarget_i + \gamma X_{i,t} + a_i + b_t + \epsilon_{i,t}.$$

Table 9 presents the results. $Target$ is a dummy variable equal to one if the firm has set at least one valid emission reduction target according to the MSCI database. $NZTarget$ equals one for firms aiming to achieve net-zero emissions by 2050. Columns (7)–(8) show that firms with emission targets experience a stronger market penalty when facing climate controversies, with the effect becoming more pronounced and statistically significant over longer event windows. However, this pattern does not hold for firms with net-zero targets. One possible explanation is that firms with general emission reduction targets are perceived as committed to sustainability but lacking accountability mechanisms, making them more vulnerable to accusations of greenwashing. Investors react more negatively when these firms are involved in climate controversies because such incidents signal weak governance or ineffective implementation of sustainability policies. On the other hand, the absence of a significant effect for net-zero target firms may indicate that investors view these firms as having longer-term strategic plans that are less sensitive to short-term controversies.

6 Conclusion

In this paper, we introduce a novel natural language processing pipeline, *ClimateControversyBERT*, explicitly designed to identify and categorize climate-related corporate controversies within financial news at scale. By distinguishing explicit, high-impact misconduct (*brown projects*) and subtle forms of greenwashing (*misinformation* and *ambiguous actions*), our approach enables investors, regulators, and policymakers to evaluate the credibility of corporate climate commitments more effectively.

Our methodology demonstrates significant complementarity with established ESG assessment metrics such as RepRisk and Refinitiv, overcoming limitations inherent to traditional approaches that depend primarily on self-reported data or delayed manual assessments. The real-time and granular insights provided by *ClimateControversyBERT* offer a robust early-warning mechanism, supporting proactive management of ESG risks and improved transparency in corporate climate disclosures. Importantly, our findings reveal significant economic consequences of climate controversies. Firms linked to such incidents experience notably negative cumulative abnormal returns, especially in cases involving misinformation or direct brown projects. These results highlight the material financial risks of climate-related misconduct and indicate that investors respond strongly to perceived failures in environmental responsibility. The market penalties are more significant for firms with emission reduction targets, suggesting heightened investor sensitivity to potential greenwashing.

This study helps bridge the gap between NLP innovation and climate finance, promoting greater transparency in transitioning to a low-carbon economy. However, our work has some limitations. Specifically, the fine-tuning process relies on a relatively modest dataset comprising 2,199 sentences, and the clustering of controversies surrounding significant events complicates the resultant trend analyses. Future research endeavors could seek to expand the training corpus, potentially incorporating non-English texts, and further investigate the dynamics of how controversies propagate in real-time across social media platforms. Such advancements would significantly bolster the efficacy of automated controversy detection, thereby supporting sustainable business practices and facilitating informed investment decisions.

References

Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., & Vollgraf, R. (2019). FLAIR: An easy-to-use framework for state-of-the-art NLP. *NAACL 2019, 2019 Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, 54–59.

Ardia, D., Bluteau, K., Boudt, K., & Inghelbrecht, K. (2023). Climate Change Concerns and the Performance of Green vs. Brown Stocks. *Management Science*, *69*(12), 7607–7632.

Bachmann, M. (2024, April). *Rapidfuzz/rapidfuzz: Release 3.8.1* (Version v3.8.1). Zenodo.

Baz, S., Cathcart, L., Michaelides, A., & Zhang, Y. (2023). Firm-level climate regulatory exposure. *Working Paper, SSRN 3873886.*

Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2022). Cheap talk and cherry-picking: What climatebert has to say on corporate climate risk disclosures. *Finance Research Letters*, *47*, 102776.

Bingler, J. A., Kraus, M., Leippold, M., & Webersinke, N. (2024). How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk. *Journal of Banking and Finance*, *164*, 107191.

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

Bouzzine, Y. D., & Lueg, R. (2020). The contagion effect of environmental violations: The case of dieselgate in germany. *Business Strategy and the Environment*, *29*(8), 3187–3202.

Das, S. R. (2011). News analytics: Framework, techniques, and metrics. *The handbook of News Analytics in Finance*, 41–71.

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

Garrido-Merchán, E. C., González-Barthe, C., & Vaca, M. C. (2023). Fine-tuning climatebert transformer with climatext for the disclosure analysis of climate-related financial risks. *arXiv preprint arXiv:2303.13373.*

Giglio, S., Kuchler, T., Stroebel, J., & Zeng, X. (2023). Biodiversity risk. *Working Paper, SSRN 4410107.*

Hsu, P.-h., Li, K., & Tsou, C.-Y. (2023). The pollution premium. *The Journal of Finance*, *78*(3), 1343–1392.

Ilhan, E., Sautner, Z., & Vilkov, G. (2020). Carbon Tail Risk. *The Review of Financial Studies*, *34*(3), 1540–1571.

Jong, W., & van der Linde, V. (2022). Clean diesel and dirty scandal: The echo of Volkswagen's dieselgate in an intra-industry setting. *Public Relations Review*, *48*(1), 102146.

Kleinnijenhuis, J., Schultz, F., Oegema, D., & Van Atteveldt, W. (2013). Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism*, *14*(2), 271–291.

Kleinnijenhuis, J., Schultz, F., Utz, S., & Oegema, D. (2015). The Mediating Role of the News in the BP Oil Spill Crisis 2010: How U.S. News Is Influenced by Public Relations and in Turn Influences Public Awareness, Foreign News, and the Share Price. *Communication Research*, *42*(3), 408–428.

Kölbel, J. F., Leippold, M., Rillaerts, J., & Wang, Q. (2024). Ask BERT: How regulatory disclosure of transition and physical climate risks affects the CDS term structure. *Journal of Financial Econometrics*, *22*(1), 30–69.

Kruttli, M. S., Roth Tran, B., & Watugala, S. W. (2023). Pricing poseidon: Extreme weather uncertainty and firm return dynamics. *Working Paper, SSRN 3284517*.

Li, Q., Shan, H., Tang, Y., & Yao, V. (2024). Corporate climate risk: Measurements and responses. *The Review of Financial Studies*, hhad094.

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

McGreal, C. (2021). Exxon CEO accused of lying about climate science to congressional panel. *The Guardian*.

MSCI. (2024). *MSCI ESG Controversies and Global Normas Methodlogy*.

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, *146*(2), 403–424.

Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023). Firm-level climate change exposure. *The Journal of Finance*, *78*(3), 1449–1498.

Schimanski, T., Bingler, J., Hyslop, C., Kraus, M., & Leippold, M. (2023). ClimateBERT-NetZero: Detecting and assessing net zero and reduction targets. *arXiv preprint arXiv:2310.08096*.

Son, G., Lee, H., Kang, N., & Hahm, M. (2023). Removing non-stationary knowledge from pre-trained language models for entity-level sentiment classification in finance.

Stammbach, D., Webersinke, N., Bingler, J. A., Kraus, M., & Leippold, M. (2023). Environmental claim detection. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*, 1051–1066.

Supran, G., Rahmstorf, S., & Oreskes, N. (2023). Assessing ExxonMobil's global warming projections. *Science*, *379*(6628), eabk0063.

Varini, F. S., Boyd-Graber, J., Ciaramita, M., & Leippold, M. (2020). ClimaText: A dataset for climate change topic detection. *arXiv preprint arXiv:2012.00483*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems (NeurIPS)*, *30*.

Webersinke, N., Kraus, M., Bingler, J. A., & Leippold, M. (2021). ClimateBERT: A pretrained language model for climate-related text. *arXiv preprint arXiv:2110.12010*.

Yu, Z., Wang, M. D., Wei, X., & Lou, J. (2023). News credibility and influence within the financial markets. *Journal of Behavioral Finance*, *24*(2), 238–257.

Variable Definitions

Variables	Definitions	Sources
$ClimateControversy$	The logarithm of the number of climate controversies.	Dow Jones
$BrownProjects$	The logarithm of the number of climate controversies related to brown projects.	Dow Jones
$Misinformation$	The logarithm of the number of climate controversies related to misinformation.	Dow Jones
$AmbiguousActions$	The logarithm of the number of climate controversies related to ambiguous actions.	Dow Jones
$Incidents^{All}$	The number of incidents.	RepRisk
$Incidents^{Env}$	The number of incidents related to environment.	RepRisk
$Incidents^{Climate}$	The number of incidents related to climate change.	RepRisk
$Incidents^{GHG}$	The number of incidents related to greenhouse gas.	RepRisk
$Incidents^{MisCom}$	The number of incidents related to misleading communication.	RepRisk
$Incidents^{Lobby}$	The number of incidents related to lobbying.	RepRisk
$ESGControversies$	The ESG Controversies scores.	Refinitiv
$\Delta MCCC$	The daily change of MCCC index.	Ardia et al. (2023)
$\Delta MCCC^{Bus}$	The daily change of MCCC index related to business impact.	Ardia et al. (2023)
$\Delta MCCC^{Env}$	The daily change of MCCC index related to environmental impact.	Ardia et al. (2023)
$\Delta MCCC^{Soc}$	The daily change of MCCC index related to societal debate.	Ardia et al. (2023)
$\Delta MCCC^{Res}$	The daily change of MCCC index related to research.	Ardia et al. (2023)
$CAR[-10, X]$	Cumulative market-adjusted abnormal returns for the event window ten days before the event date to X days afterward.	CRSP
$log(Asset)$	Logarithm of total assets.	Compustat
$log(Size)$	Logarithm of market capital.	Compustat
$log(B/M)$	Logarithm of book to market ratio.	Compustat
ROA	Operating income before depreciation divided by total assets.	Compustat
$Investment$	Growth rate of total assets.	Compustat
$Leverage$	Total debt divided by total assets.	Compustat
$Tangibility$	Net property, plant, and equipment divided by total assets.	Compustat
$SaleGrowth$	Growth rate of sales.	Compustat

Table 1: Distribution of Climate-Related News Articles & Paragraphs

This table presents the distribution of climate-related news articles and paragraphs through filtering and classification. Panel A shows that only 1.48% of the 15.7 million articles analyzed met our climate relevance criteria after excluding articles with fewer than 100 words and filtering for climate-specific subject codes or keywords. A significant proportion of the excluded news articles were very short and represented news briefs rather than full articles. Panel B details the paragraph-level analysis based on the climate related news articles defined in Panel A where ClimateBERT classified 43.55% of all paragraphs as climate-related. Among these, 1.2% contained controversies, which were further categorized as Brown Projects (0.62%), Misinformation (0.37%), and Ambiguous Actions (0.21%).

Panel A: Article-Level Distribution		
Climate-Related	230,942	1.48%
Total	15,655,478	100%
Panel B: Paragraph-Level Distribution		
Climate-Related	1,415,698	43.55%
No-Controversy	1,364,908	41.99%
Controversy	38,721	1.2%
Brown Projects	20,139	0.62%
Misinformation	11,908	0.37%
Ambiguous Actions	6,674	0.21%
Not Climate-Related	1,834,995	56.45%
Total	3,250,936	100.00%

Table 2: Example of Controversies

The table presents example paragraphs for different types of corporate controversies detected by *ClimateControversyBERT* in news articles, including brown projects, misinformation, and ambiguous actions. It also includes each company's emission reduction commitment and current implementation status based on MSCI data. "Net Zero (2050)" indicates that the firm has pledged to achieve net-zero emissions by 2050; otherwise, we report the emission Scopes for which the company has set reduction targets.

Firm	Controversy Type	Paragraph	Commitment	Status
HSBC	Brown Projects	"HSBC's new coal phase-out policy 'ignores the elephant in the room': the investments its asset-management arm has in companies engaged in new coal-power projects leaving 15 billion tons of CO2, roughly equivalent to what the U.K. emits in 32 years, in play."	Scope 1, 2 and 3	On track with all targets
Goldman Sachs	Brown Projects	"Not everything Goldman does is environmentally savvy, either. The firm makes billions as an owner and a banker in an industry often targeted by green lobbyists: power-generation plants. Goldman owns full or partial stakes in 19 plants around the U.S. that emit millions of tons of gases thought to cause global warming."	Net Zero (2050)	Not on track with any targets
Drax	Brown Projects	"Drax is the U.K.'s largest single CO2 emitter, producing about 20.8 million tons of CO2 a year. The coal-fired power plant accounts for some 7% of the U.K.'s power needs.	Net Zero (2050)	On track with all targets
Volkswagen	Misinformation	"The U.S. Environmental Protection Agency disclosed last week that Volkswagen had admitted to using software in some VW and Audi diesel-powered cars to falsify the results of emissions tests. The company is now subject to an EPA and U.S. criminal investigation and could face fines in the region of $18 billion."	Scope 1, 2 and 3	On track with some targets
Deutsche Bank	Misinformation	"In a bellwether case, the Securities and Exchange Commission is investigating whether Deutsche Bank AG's asset-management arm lived up to claims it made about its ESG investing criteria. A whistleblower and internal emails say that only a fraction of its assets went through a sustainability assessment, contrary to the firm's public statements."	Net Zero (2050)	Not on track with any targets
BNY Mellon	Misinformation	"In May, New York-based mutual fund manager BNY Mellon Investment Adviser agreed to pay $1.5 million to settle SEC charges that it misrepresented the ESG review it made of investments. The bank didn't admit the administrative charges."	Scope 1 and 2	On track with all targets
ExxonMobil	Ambiguous Actions	"Senior executives within the company now believe it needs to act urgently to refine its strategy to navigate the energy transition, and some of Exxon's largest shareholders have told executives recently that they need to set more-ambitious climate-change targets or risk further alienating investors, the people said."	Scope 1 and 2	Not on track with any targets
Chevron	Ambiguous Actions	"Chevron is facing mounting pressure from ESG investors that asked the company to set targets for its indirect emissions as a sign it is addressing climate change and the risks associated with the rise in global temperatures."	Scope 1 and 2	Not on track with any targets
Amazon	Ambiguous Actions	"Earlier this year, Amazon shareholders rejected an investor-led proposal that called for the company to disclose how its business could be disrupted by climate change and how it could reduce dependence on fossil fuels. The company had recommended investors vote against the measure, in part, because it said it was already making certain disclosures on the topic."	Net Zero (2050)	Not on track with any targets

Table 3: Industry Distribution of Climate Controversies

This table presents the top 20 industries with the highest number of climate controversies, using the Fama-French 49 industry classifications.

Fama-French 49 Industries	Climate Controversies	Brown Projects	Misinformation	Ambiguous Actions
Petroleum and Natural Gas	11,438	6,499	2,449	1,169
Utilities	3,231	1,859	468	592
Automobiles and Trucks	1,700	243	1,099	245
Non-Metallic and Industrial Metal Mining	1,011	660	95	119
Banking	448	175	58	167
Chemicals	367	169	90	55
Trading	329	162	46	87
Transportation	307	190	37	41
Electronic Equipment	295	108	86	69
Wholesale	283	199	48	16
Steel Works Etc	263	182	23	24
Retail	251	105	48	68
Coal	246	176	13	41
Food Products	234	118	33	63
Machinery	217	97	67	32
Consumer Goods	207	87	39	46
Computer Software	166	65	43	50
Electrical Equipment	155	83	36	23
Insurance	102	30	12	50
Construction Materials	99	51	20	15

Table 4: Validation Tests: RepRisk Incidents

This table reports estimates from Poisson regressions of RepRisk incident numbers on the logarithm of climate controversy counts at firm-year level. Control variables include the total number of news articles mentioning the firm, total assets, ROA, investment rate, leverage, tangibility, and sales growth. We consider industry and year-fixed effects. t-statistics based on standard errors clustered at the industry level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Panel A: Overall Climate Controversies						
	$Incidents^{All}$	$Incidents^{Env}$	$Incidents^{Climate}$	$Incidents^{GHG}$	$Incidents^{MisCom}$	$Incidents^{Lobby}$
	(1)	(2)	(3)	(4)	(5)	(6)
$ClimateControversy$	0.40***	0.43***	0.38***	0.40***	0.43***	0.29***
	(8.72)	(10.00)	(8.22)	(7.80)	(7.86)	(4.12)
N	35,181	35,181	35,181	35,181	35,112	22,283
Pseudo R-sq	0.52	0.47	0.46	0.44	0.43	0.30
Control	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Panel B: Climate Controversy Types						
	$Incidents^{All}$	$Incidents^{Env}$	$Incidents^{Climate}$	$Incidents^{GHG}$	$Incidents^{MisCom}$	$Incidents^{Lobby}$
	(1)	(2)	(3)	(4)	(5)	(6)
$BrownProjects$	0.31***	0.37***	0.34***	0.38***	0.33***	0.31***
	(5.59)	(5.40)	(5.60)	(6.31)	(4.73)	(4.25)
$Misinformation$	0.29***	0.20**	0.03	0.02	0.24**	0.37***
	(3.97)	(2.19)	(0.34)	(0.34)	(2.36)	(3.16)
$AmbiguousActions$	0.11**	0.06	0.13**	0.10	0.20***	-0.07
	(2.14)	(0.80)	(2.03)	(1.48)	(2.75)	(-0.56)
N	35,181	35,181	35,181	35,181	35,112	22,283
Pseudo R-sq	0.51	0.46	0.46	0.43	0.43	0.30
Control	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Validation Tests: ESG Controversies Scores

This table reports estimates from regression of Refinitiv's ESG Controversies scores on the logarithm of climate controversy counts at firm-year level. Control variables include the total number of news articles mentioning the firm, total assets, ROA, investment rate, leverage, tangibility, and sales growth. We consider industry and year-fixed effects. t-statistics based on standard errors clustered at the industry level are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	$ESGControversies$	
	(1)	(2)
$ClimateControversy$	-0.05***	
	(-3.64)	
$BrownProjects$		-0.01
		(-0.39)
$Misinformation$		-0.08***
		(-4.72)
$AmbiguousActions$		-0.04***
		(-2.73)
N	8,735	8,735
adj. R-sq	0.35	0.35
Controls	Yes	Yes
Industry FE	Yes	Yes
Year FE	Yes	Yes

Table 6: Validation Tests: MCCC Index

This table reports estimates from time-series regressions of daily changes in the MCCC Index on daily changes in aggregate climate controversy counts. t-statistics based on robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Panel A: Overall Climate Controversies					
	$\Delta MCCC$	$\Delta MCCC^{Bus}$	$\Delta MCCC^{Env}$	$\Delta MCCC^{Soc}$	$\Delta MCCC^{Res}$
	(1)	(2)	(3)	(4)	(5)
$\Delta ClimateControversy$	0.34***	0.42***	0.21***	0.19***	0.31***
	(4.82)	(6.28)	(2.87)	(2.71)	(3.63)
N	6,275	6,275	6,275	6,275	6,275

Panel B: Climate Controversy Types					
	$\Delta MCCC$	$\Delta MCCC^{Bus}$	$\Delta MCCC^{Env}$	$\Delta MCCC^{Soc}$	$\Delta MCCC^{Res}$
	(1)	(2)	(3)	(4)	(5)
$\Delta BrownProjects$	0.30***	0.35***	0.15	0.23**	0.22**
	(3.00)	(3.76)	(1.59)	(2.18)	(2.06)
$\Delta Misinformation$	0.45***	0.60***	0.30	0.21	0.40*
	(2.62)	(4.19)	(1.57)	(1.12)	(1.96)
$\Delta AmbiguousActions$	0.69***	0.65**	0.69**	0.07	0.90***
	(2.63)	(2.54)	(2.25)	(0.26)	(3.12)
N	6,275	6,275	6,275	6,275	6,275

Table 7: Stock Market Reaction

This table examines whether the cumulative abnormal returns (CARs) around climate controversies are different from zero. Cumulative market-adjusted abnormal returns are measured from ten days before the event to the event date or five/ten/fifteen/twenty days after. We consider firm-date with at least two news articles related to climate controversies as our events to reduce noise. The sample period is from 2004 to 2023. The sample includes firms with Dow Jones news coverage and Compustat/CRSP data. t-statistics based on standard errors clustered by industry and date are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	[-10,0]	[-10,5]	[-10,10]	[-10,15]	[-10,20]
	(1)	(2)	(3)	(4)	(5)
Panel A: Overall					
CAR	-0.68***	-0.87***	-1.02***	-1.13***	-1.27***
t-stat	(-4.70)	(-4.17)	(-4.43)	(-3.08)	(-3.36)
N	2,330	2,330	2,330	2,330	2,330
Panel B: Brown Projects					
CAR	-0.92***	-1.24***	-1.40***	-1.62**	-1.85**
t-stat	(-3.53)	(-3.65)	(-3.31)	(-2.61)	(-2.63)
N	1,507	1,507	1,507	1,507	1,507
Panel C: Misinformation					
CAR	-1.02***	-1.52***	-1.79***	-1.85***	-2.09***
t-stat	(-6.24)	(-6.48)	(-6.40)	(-4.49)	(-6.03)
N	884	884	884	884	884
Panel D: Ambiguous Actions					
CAR	-0.74	-1.05	-1.02*	-0.96	-0.90
t-stat	(-1.44)	(-1.69)	(-1.97)	(-1.59)	(-1.68)
N	461	461	461	461	461

Table 8: Stock Market Reaction: Severity of Controversies

This table regresses the cumulative abnormal returns (CARs) around climate controversies on the number of news articles mentioning each event, as a proxy for severity. The dependent variables are cumulative market-adjusted abnormal returns measured from ten days before the event to the event date or five/ten/fifteen/twenty days after. We consider firm-date with at least two news articles related to climate controversies as our events to reduce noise. Control variables include the natural logarithm of size, book-to-market ratio, profitability, investment rate, leverage, tangibility, and sales growth. All variables are normalized to a zero mean and one standard deviation. The sample period is from 2004 to 2023. The sample includes firms with Dow Jones news coverage and Compustat/CRSP data. We consider firm and year-month fixed effects. t-statistics based on standard errors clustered by industry and year are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	CAR[-10,0]		*CAR*[-10,5]		*CAR*[-10,10]		*CAR*[-10,15]		*CAR*[-10,20]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
ClimateControversy	-0.99***		-2.24***		-2.90***		-2.42***		-2.35***	
	(-4.11)		(-5.29)		(-5.16)		(-4.84)		(-3.88)	
BrownProjects		-0.79**		-1.54***		-1.67**		-1.60**		-1.58**
		(-2.49)		(-3.30)		(-2.79)		(-2.76)		(-2.77)
Misinformation		-0.34		-0.73***		-0.91***		-0.57**		-0.43*
		(-1.68)		(-3.49)		(-3.95)		(-2.62)		(-1.74)
AmbiguousActions		-0.64		-1.34*		-1.58*		-1.15		-0.85
		(-1.28)		(-1.98)		(-1.95)		(-1.40)		(-0.97)
log(Size)	-2.59***	-2.51***	-3.97***	-3.82***	-3.82**	-3.68**	-3.24*	-3.10*	-3.95*	-3.82*
	(-2.88)	(-2.94)	(-3.27)	(-3.42)	(-2.65)	(-2.67)	(-1.86)	(-1.85)	(-2.06)	(-2.05)
log(B/M)	-2.64***	-2.54**	-2.72**	-2.49*	-1.93	-1.65	-0.83	-0.62	-0.75	-0.57
	(-2.91)	(-2.85)	(-2.27)	(-2.08)	(-1.46)	(-1.20)	(-0.47)	(-0.34)	(-0.34)	(-0.25)
ROA	-0.90**	-0.94**	-0.52	-0.58	-0.36	-0.41	-0.09	-0.16	-0.09	-0.15
	(-2.23)	(-2.37)	(-1.01)	(-1.19)	(-0.56)	(-0.66)	(-0.16)	(-0.28)	(-0.13)	(-0.23)
Investment	-0.45*	-0.45*	-0.63***	-0.63**	-0.49**	-0.47*	-0.14	-0.13	-0.12	-0.10
	(-2.09)	(-1.87)	(-2.90)	(-2.67)	(-2.18)	(-1.88)	(-0.43)	(-0.40)	(-0.43)	(-0.35)
Leverage	-1.41	-1.38	-1.20	-1.16	-0.57	-0.53	-0.03	0.02	-0.19	-0.13
	(-1.32)	(-1.31)	(-1.05)	(-1.01)	(-0.47)	(-0.44)	(-0.03)	(0.02)	(-0.17)	(-0.12)
Tangibility	0.66	0.64	1.04	1.01	0.94	0.89	1.09	1.03	1.10	1.02
	(0.90)	(0.86)	(1.41)	(1.35)	(0.97)	(0.90)	(1.24)	(1.15)	(0.99)	(0.91)
SalesGrowth	0.24	0.22	0.18	0.14	-0.09	-0.13	-0.43	-0.47	-0.06	-0.10
	(0.76)	(0.69)	(0.33)	(0.26)	(-0.15)	(-0.23)	(-1.11)	(-1.22)	(-0.13)	(-0.22)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	2,175	2,175	2,175	2,175	2,175	2,175	2,175	2,175	2,175	2,175
adj. R-sq	0.34	0.34	0.39	0.39	0.42	0.42	0.45	0.45	0.46	0.46

Table 9: Stock Market Reaction: Emission Reduction Targets

This table regresses the cumulative abnormal returns (CARs) around climate controversies on the number of news articles mentioning each event, as a proxy for severity. The dependent variables are cumulative market-adjusted abnormal returns measured from ten days before the event to the event date or five/ten/fifteen/twenty days after. We consider firm-date with at least two news articles related to climate controversies as our events to reduce noise. $Target$ equals one if the firm has set at least one valid emission reduction target. $NZTarget$ equals one for the firm aims to achieve net zero emissions by 2050. Control variables include the natural logarithm of size, book-to-market ratio, profitability, investment rate, leverage, tangibility, and sales growth. All variables are normalized to a zero mean and one standard deviation. The sample period is from 2010 to 2023 (due to the availability of emission target data). The sample includes firms with Dow Jones news coverage and Compustat/CRSP data. We consider firm and year-month fixed effects. t-statistics based on standard errors clustered by industry and year are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

	CAR[-10,0]		CAR[-10,5]		CAR[-10,10]		CAR[-10,15]		CAR[-10,20]	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$ClimateControversy$	-1.06***	-1.11***	-2.46***	-2.42***	-3.19***	-3.07***	-1.97**	-2.44***	-1.82**	-2.32***
	(-3.45)	(-3.74)	(-3.96)	(-4.03)	(-4.35)	(-4.33)	(-2.83)	(-3.57)	(-2.36)	(-3.14)
$ClimateControversy \times Target$	-0.05		0.27		0.30		-0.88**		-0.90**	
	(-0.15)		(0.66)		(0.79)		(-2.68)		(-2.75)	
$ClimateControversy \times NZTarget$		0.69		3.54		1.39		-0.35		0.37
		(0.28)		(1.59)		(0.52)		(-0.08)		(0.10)
$log(Size)$	-2.36*	-2.35*	-4.18**	-4.14**	-4.49*	-4.49*	-3.83	-3.78	-5.56*	-5.50*
	(-1.79)	(-1.78)	(-2.25)	(-2.25)	(-2.03)	(-2.03)	(-1.49)	(-1.47)	(-2.11)	(-2.09)
$log(B/M)$	-2.99**	-2.99**	-3.50*	-3.53*	-2.58	-2.60	-2.24	-2.21	-2.52	-2.49
	(-2.93)	(-2.98)	(-2.02)	(-2.05)	(-1.21)	(-1.21)	(-0.92)	(-0.90)	(-0.85)	(-0.84)
ROA	-0.65	-0.65	-0.15	-0.13	-0.08	-0.07	0.14	0.13	-0.17	-0.19
	(-1.02)	(-1.01)	(-0.20)	(-0.19)	(-0.08)	(-0.08)	(0.17)	(0.15)	(-0.20)	(-0.21)
$Investment$	-0.56	-0.57	-0.61	-0.62	-0.32	-0.32	0.10	0.10	0.27	0.26
	(-1.45)	(-1.48)	(-1.03)	(-1.06)	(-0.56)	(-0.57)	(0.17)	(0.17)	(0.42)	(0.41)
$Leverage$	-1.60	-1.60	-1.26	-1.26	-0.56	-0.55	-0.44	-0.46	-0.95	-0.98
	(-1.20)	(-1.21)	(-0.91)	(-0.91)	(-0.34)	(-0.34)	(-0.30)	(-0.32)	(-0.63)	(-0.64)
$Tangibility$	1.03	1.04	2.15**	2.16**	2.36**	2.36**	2.15**	2.18**	2.46**	2.50**
	(1.33)	(1.35)	(2.25)	(2.27)	(2.38)	(2.32)	(2.35)	(2.19)	(2.37)	(2.24)
$SalesGrowth$	-0.09	-0.09	-0.26	-0.28	-0.50	-0.51	-0.80	-0.78	-0.24	-0.21
	(-0.20)	(-0.20)	(-0.52)	(-0.55)	(-0.72)	(-0.75)	(-1.46)	(-1.43)	(-0.37)	(-0.33)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,695	1,695	1,695	1,695	1,695	1,695	1,695	1,695	1,695	1,695
adj. R-sq	0.36	0.36	0.39	0.39	0.43	0.43	0.45	0.45	0.48	0.48

A Data Preparation and Annotation for Fine-tuning

The fine-tuning of our controversy detection and classification models (Steps 2 and 3) relies on a carefully constructed, multi-stage process to generate a high-quality annotated dataset that reflects real-world climate controversies. Due to the challenge of sourcing sufficient and accurately labeled data, particularly for nuanced classifications, we begin by identifying initial seed examples through reverse-engineering severe climate incidents from the RepRisk database. Specifically, we search for related news coverage for these incidents and manually collect 578 initial sentences. Some of these sentences contain repetitive information or are not directly related to corporate climate commitments. To broaden the linguistic diversity and focus on climate commitment-related controversies, we employ a bootstrapping strategy.

First, we utilize GPT-3.5 to generate synthetically similar sentences based on the manually collected examples to expand the set of positive samples. This augmented dataset, combined with more readily available negative (non-climate-related) samples, is used to fine-tune a preliminary ClimateControversyBERT classifier.

Second, we compile a large pool of sentences by randomly sampling 30% of Dow Jones Newswires articles from each year between 2015 and 2023 (729,663 sentences in total). These articles are filtered using the same criteria as Step 1 of our main pipeline. Sentences that meet the length requirement (5–100 words) and are identified as climate-related with high confidence (>0.95) by a base ClimateBERT model form a candidate pool of approximately 218,000 sentences.

Third, we apply the preliminary controversy classifier to the large candidate pool.[7] Based on its predictions, we strategically sample sentences for final, in-depth annotation: 1,000 sentences classified as 'Controversy' with high confidence (>0.95), 1,000 sentences classified as 'No Controversy' with high confidence (>0.95), and an additional 100 sentences predicted as 'No Controversy' but with lower confidence (<0.8) to deliberately capture challenging boundary examples and potential false negatives. We combine this data with the initial 578 seed sentences for final annotation.

Fourth, we prompt GPT-4 (specifically 'gpt-4-1106-preview') to perform the final annotation

[7] The preliminary classifier, like the final model, is based on the "climatebert/distilroberta-base-climate-sentiment" architecture and uses similar hyperparameters: batch size 32, learning rate 5e-5, and weight decay 0.01. To address class imbalance, we apply inverse class weighting.

and classify each sentence into one of four categories. These include: (1) *Brown projects* – entities that significantly contribute to emissions or high-carbon projects; (2) *Misinformation* – distortion or dissemination of deceptive climate information; (3) *Ambiguous actions* – insufficient, inconsistent, or opaque corporate climate efforts relative to stated goals or standards; (4) *No controversy* – no identifiable firm-specific climate controversy. To enhance annotation reliability, each sentence is annotated twice by GPT-4 using different prompts with varied examples and temperature settings. For the 406 sentences where the two outputs disagreed, the research team manually review and finalize the labels. This meticulous process yield a high-quality annotated dataset: 249 *Brown projects*, 314 *Misinformation*, 227 *Ambiguous actions*, and 1,409 *No controversy.* This labeled dataset serves as the foundation for training and evaluating our two-stage controversy detection and classification models.

B Additional Tables

Table B1: ClimateBERT Fine-Tuning

This table presents our choice of hyperparameter for model fine-tuning in Panel A and model performance for controversy detection and classification in Panel B.

Panel A: Choice of Hyperparameter		
Hyperparameter	Value	
Epochs	50	
Batch Size	32	
Gradient Accumulation	16	
Warmup steps	500	
Learning rate	5e-5	
Patience	4	
Panel B: Model Performance		
Metric	Detection	Types
Accuracy (std.)	0.877 (0.013)	0.828 (0.011)
F1-Score (std.)	0.877 (0.012)	0.827 (0.013)
Precision (std.)	0.878 (0.013)	0.830 (0.017)
Recall (std.)	0.877 (0.013)	0.828 (0.012)

Table B2: Time-Series Variation of Climate Controversies

This table presents the yearly frequency of our detected climate controversies.

Year	Climate Controversies	Brown Projects	Misinformation	Ambiguous Actions
2004	355	184	65	72
2005	577	360	59	73
2006	1,039	592	197	120
2007	1,180	625	231	155
2008	1,180	726	144	185
2009	1,178	651	211	226
2010	4,972	2,876	1,151	388
2011	2,507	1,386	645	208
2012	1,694	1,055	364	133
2013	1,034	527	265	93
2014	541	279	132	56
2015	704	275	325	44
2016	779	275	364	83
2017	473	144	147	103
2018	499	193	154	86
2019	1,011	414	215	282
2020	768	358	111	229
2021	1,124	501	142	402
2022	914	424	155	232
2023	488	193	57	191

Table B3: Top Firms with Climate Controversies

This table lists the top 20 firms with the highest number of climate controversies.

Firms	Climate Controversies	Brown Projects	Misinformation	Ambiguous Actions
BP plc	5,251	2,815	1,383	428
Exxon Mobil Corp	1,216	607	223	244
Chevron Corp	1,080	623	183	131
Transocean Ltd	653	358	193	20
Volkswagen AG	583	11	550	11
ConocoPhillips	445	292	42	53
TotalEnergies SE	435	305	27	73
Halliburton Co	299	101	149	8
Toyota Motor Corp	267	63	84	83
BHP Group Ltd	266	150	28	46
American Electric Power Co	253	153	27	37
Anadarko Petroleum Corp	247	137	84	9
PG&E Corp	244	120	63	45
RWE AG	237	121	29	53
Rio Tinto Group	222	155	12	28
Petroleo Brasileiro SA – PETR	188	150	12	14
Southern Co	176	102	18	25
CNOOC Ltd	169	134	17	6
E.ON SE	164	85	11	54
Eni SpA	163	108	14	25

Table B4: Geographical Coverage of Climate Controversies

This table presents the geographical distribution of our detected climate controversies.

Firms	Climate Controversies	Brown Projects	Misinformation	Ambiguous Actions
United States	12,154	6,393	2,514	1,889
United Kingdom	3,913	2,102	886	467
Germany	1,549	401	830	203
Japan	944	440	254	164
Switzerland	835	444	217	63
France	613	345	90	124
Australia	416	290	31	57
Brazil	249	178	25	15
Spain	245	151	20	57
Italy	215	127	19	46
China	186	129	28	12
India	186	142	18	18
Russia	162	110	12	13
Hong Kong	127	94	11	13
South Korea	114	57	39	8
Norway	99	55	7	23
Jersey	94	54	5	24
Netherlands	69	27	16	21
Chile	52	31	5	11
Sweden	46	24	8	7

Swiss Finance Institute

Swiss Finance Institute (SFI) is the national center for fundamental research, doctoral training, knowledge exchange, and continuing education in the fields of banking and finance. SFI's mission is to grow knowledge capital for the Swiss financial marketplace. Created in 2006 as a public–private partnership, SFI is a common initiative of the Swiss finance industry, leading Swiss universities, and the Swiss Confederation.

swiss:finance:institute

c/o University of Geneva, Bd. Du Pont d'Arve 42, CH-1211 Geneva 4
T +41 22 379 84 71, rps@sfi.ch, www.sfi.ch