

Measuring Climate Patents with LLMs: Evidence from Patent Texts

Yukun Shi*

Junyao Wang[†]

Yupu Zhang[‡]

December 2025

Abstract

This paper develops a new measure of firm-level climate innovation using generative AI applied to U.S. patent abstracts and invention descriptions. The model screens all patents based on meaning rather than fixed codes such as Y02. It captures climate patents missed by Y02 and removes Y02-tagged patents that are not truly climate-related. We validate the measure using economic tests. Generative-AI climate patents are more strongly related to patent value and forward citations than Y02-based patents. The measure also better tracks stock market reactions to major climate policy events. Firms facing stronger environmental enforcement later produce more generative-AI climate patents, while the Y02-based measure shows weaker responses. We also find real effects through supply chains: firms whose suppliers generate more generative-AI climate patents show larger later declines in Scope 1 emissions. Using this measure, we find that climate innovation is negatively related to Tobin's Q , with the effect driven by mitigation technologies and no clear valuation effect for adaptation. Overall, the results show that generative AI provides a reliable and useful measure of climate innovation.

*University of Glasgow

[†]University of Edinburgh

[‡]University of Glasgow

1 Introduction

Climate-related technological innovation has become a central element of the global transition towards a low-carbon economy. Policy assessments emphasise that progress towards climate mitigation goals depends critically on sustained advances in science and technology, rather than on regulatory measures alone (OECD, 2025). Within this policy framework, climate-related patents are increasingly viewed as an important indicator of technological effort, as they document firms’ development of new solutions aimed at reducing emissions, improving energy efficiency, and adapting production processes to environmental constraints. By capturing early-stage innovation before its effects are fully reflected in output or emissions outcomes, patent data provide a forward-looking perspective on firms’ engagement in climate-oriented technological change.¹

The most widely adopted approach identifies climate-related patents using the Y02 classification developed by the European Patent Office (EPO). Introduced in 2010 and incorporated into the Cooperative Patent Classification (CPC) system in 2013, Y02 was designed to capture technologies related to climate change mitigation and adaptation across areas such as renewable energy, energy efficiency, carbon capture, transportation, industrial processes, and waste management. While this framework provides a standardised way to track climate-oriented innovation across countries, it remains an imperfect proxy. Because Y02 assigns climate relevance based on predefined and relatively broad technology classes, it may label patents as “green” even when their contribution to climate mitigation is limited, while failing to capture inventions whose descriptions clearly relate to mitigation or adaptation but fall outside the established classification scheme (Lerner & Seru, 2022).

This paper addresses these gaps by developing a new measure of firm-level climate innovation using large language models (LLMs) applied to U.S. patent abstracts and the brief description of the invention. The analysis follows a two-step approach. Given the large

¹https://www.oecd.org/en/publications/oecd-science-technology-and-innovation-outlook-2025_5fe57b90-en.html. OECD (2025), *OECD Science, Technology and Innovation Outlook 2025*, OECD Publishing.

number of patents, we first use abstracts to screen whether a patent is potentially related to climate issues. We then use the brief description of the invention to refine this classification and to distinguish more clearly among different types of climate-related technologies. Instead of relying on fixed technology classes such as Y02, the LLM reads the text of each patent and assesses whether the invention relates to climate goals based on its meaning. By using information from both the abstract and the invention description, the method captures climate-related aspects that are described in the text but not covered by existing classification codes. As a result, the LLM-based measure offers a broader and more informative view of firms' climate innovation activity.

To establish the validity of the proposed measure, we conduct a series of tests following the approach commonly used in the textual-analysis literature Hassan et al. (2019) and Sautner, Van Lent, et al. (2023a). We begin by comparing climate-related patents identified by the LLM with those identified using the Y02 classification. We assess whether the LLM-based measure is economically meaningful by examining its association with patent value and forward citations. Patents classified as climate-related by the LLM model exhibit significantly higher market-based value and receive more forward citations than patents identified using Y02. These results suggest that the LLM classifier is more closely aligned with the economic substance of climate-related innovation rather than purely formal or technological classifications. Next, we conduct a market-based validation by analysing investor responses to major climate policy events. Firms with higher levels of LLM-identified climate innovation experience more negative stock price reactions following climate policy rollbacks and more positive reactions following policy support. In contrast, the Y02-based measure shows little systematic response to these events. This evidence indicates that financial markets view the LLM-based measure as more informative about firms' exposure to climate-related risks and opportunities. We further examine whether the new measure predicts firms' responses to environmental shocks and regulatory pressure. Firms subject to more frequent environmental inspections subsequently generate significantly more climate-related patents

when innovation is measured using the LLM approach, while the Y02 measure exhibits little response. In addition, we explore real environmental outcomes and supply-chain channels. Firms whose suppliers generate more LLM-identified climate patents exhibit larger reductions in Scope 1 emissions over time, providing further evidence that the measure captures meaningful climate-related activity.

Having established the validity and economic relevance of the LLM-based climate innovation measure, we examine how climate-related innovation is reflected in firm valuation and real corporate outcomes. Market valuation is measured using Tobin’s Q , which allows us to assess whether equity markets treat climate-related innovation as a source of long-term value or as a near-term cost associated with adjustment and uncertainty. We also examine heterogeneity in this relationship by’ carbon exposure of firms, testing whether climate innovation is priced differently for firms facing a greater risk of transition. In addition, we study whether valuation effects differ across types of climate-related technologies. From an asset-pricing perspective, this analysis sheds light on whether markets apply a uniform discount to climate innovation or distinguish across technologies based on their economic implications. The negative association between climate patenting and Tobin’s Q is concentrated in mitigation-related innovations, while adaptation-oriented patents show no systematic relation with firm value. This pattern is consistent with markets pricing climate innovation primarily through expected cash-flow effects and investment risk, as mitigation technologies are more likely to involve large upfront costs, delayed payoffs, and higher uncertainty. Moreover, we distinguish between cash-flow and discount-rate channels to clarify the underlying mechanisms. This allows us to assess whether valuation effects reflect changes in expected operating performance or shifts in required returns.

We then turn to the real effects of climate patenting. Specifically, we examine whether firms that engage more intensively in climate-related innovation subsequently reduce their carbon emissions and improve their environmental and ESG-related performance. By linking climate patenting to emissions growth and external ESG assessments, this analysis evaluates

whether the innovation captured by the LLM-based measure reflects substantive technological and organisational change rather than symbolic activity. Through this sequence of analyses, the paper connects a new measure of climate innovation to firm valuation, corporate adjustment, and environmental outcomes, offering a unified perspective on how climate-related innovation is recognised by financial markets and translated into real economic effects.

This paper contributes to the literature on textual analysis in finance by introducing a new LLM-based approach to identify climate-related patents. Prior studies largely rely on predefined taxonomies or dictionary-based methods to extract information from text (Hoberg & Phillips, 2016; Loughran & McDonald, 2011). In the context of climate innovation, the dominant approach is to use CPC Y02 classifications to label patents as green. While widely adopted, this method assigns climate relevance at a broad class level and does not fully reflect the actual content of individual inventions. As a result, it may include patents with limited environmental relevance and overlook innovations that clearly address climate issues but fall outside the designated classes. Recent work has begun to relax the use of fixed dictionaries by adopting machine learning techniques, such as Word2Vec or keyword discovery algorithms, to expand relevant vocabularies (Sautner, van Lent, et al., 2023). However, these approaches remain largely word-based and do not fully capture meaning in context. Building on emerging studies that apply large language models to measure economic concepts from text (Harford et al., 2023), we apply an LLM to patent abstracts to construct a firm-level climate innovation measure. By reading and interpreting the abstract and brief description of the invention section of patent documents rather than relying on predefined codes or keywords, the model identifies climate-related inventions based on their semantic content. This approach provides a more flexible and fine-grained measure of climate innovation and reduces measurement error inherent in existing classifications. Through this approach, the paper extends the use of LLMs from narrative disclosures to technological innovation and illustrates how context-aware language models can generate economically meaningful measures for research on climate-related technological change.

Moreover, our study contributes to the literature on green innovation by improving how climate-related patents are measured in empirical research (Block et al., 2025; Fusillo, 2023; Gianluca Orsatti et al., 2020; Green et al., 2024; Hoberg & Phillips, 2016). Most existing studies identify green patents using broad expert-based classification systems, such as IPC or CPC codes, including the widely used Y02 classes. These methods allow researchers to work with large samples, but they assume that all patents within a given class are equally relevant for environmental goals. Recent research has raised doubts about this assumption and shows that many patents labelled as green under standard classifications do not reflect meaningful environmental innovation. We add to this literature by showing that how green patents are measured matters for empirical results. In particular, recent evidence suggests that the economic benefits of green patents are driven by a much smaller set of patents that are truly green and technologically novel (Santarlaschi et al., 2025). By refining the identification of climate-related patents and focusing on their technological content rather than broad labels, our study addresses concerns about misclassification in the green innovation literature. In this sense, our contribution supports calls for more careful measurement of environmental innovation, as imprecise definitions may blur both the true economic value of green technologies and the role of strategic or symbolic patenting.

Besides, this paper contributes to the literature on climate innovation and firm valuation by providing new asset-pricing evidence using a text-based measure of climate-related innovation. We apply a large language model to patent texts to identify climate innovation based on technological content, rather than relying on predefined patent classes. This approach helps capture differences across climate technologies that standard classification schemes may miss. Our results complement and extend prior evidence on the weak or mixed valuation effects of climate innovation. While existing studies find limited investor valuation of green patents (Asimakopoulous et al., 2023) or show that climate innovation mainly offsets valuation losses linked to carbon risk (Kuang & Liang, 2024), we find that climate innovation is associated with lower firm valuation when measured more precisely. At the same time, our

findings align with evidence that valuation effects vary across climate technologies (Briere et al., 2025; Dechezleprêtre et al., 2021).

The remainder of the paper is organised as follows. Section 2 introduces the research setting and reviews the related literature. Section 3 develops the testable hypotheses. Section 4 describes the data, and Sections 5 and 6 present the construction and validation of the LLM-based climate patent measure. Section 7 examines how climate patenting is priced by capital markets, while Section 8 analyses the real effects of climate patenting on emissions and environmental performance. Section 9 concludes.

2 Literature Review

2.1 Green innovation

2.1.1 Drivers of Green innovation

Empirical research highlights the central roles of regulatory pressure, market incentives, and firm-specific conditions in shaping green innovation. On the regulatory side, environmental rules constitute a primary source of external discipline. J. Lin et al. (2024) show that regulatory pressure imposed on downstream firms is transmitted through supply-chain linkages, prompting upstream suppliers to strengthen their engagement in green innovation. He and Qiu (2025) find that formal enforcement actions by the U.S. Environmental Protection Agency lead to a rise in green innovation by sanctioned firms, a pattern explained by improvements in innovation efficiency and the redeployment of environmentally skilled labour. Evidence from emissions taxes indicates a similar mechanism: such policies not only promote cleaner technologies but also enhance firms’ capacity to absorb external knowledge (Brown et al., 2022). By contrast, uncertainty surrounding environmental policies—particularly unstable subsidy regimes driven by shifts in air pollution—reduces firms’ willingness to invest in green R&D (Wang et al., 2026). These findings suggest that credible and stable regulatory

signals are essential for sustaining green innovation incentives.

Market forces provide an additional set of motivations. Aghion et al. (2023) show that rising environmental awareness among consumers, combined with stronger product-market competition, encourages firms to adopt “clean innovation” as a means of avoiding price-based rivalry and meeting the expectations of socially responsible consumers. Consistent with this view, Dai et al. (2025) document that competition amplifies firms’ incentives to pursue green innovation when regulatory pressure tightens, enabling firms to meet compliance requirements while differentiating through environmental technologies. Supply chain relationships further contribute to this process: climate related innovation by suppliers especially product-oriented innovation reduces the carbon emissions of their customers, and high emission customers systematically prefer suppliers with stronger climate-patent portfolios Hege et al. (2023). In international settings, firms operating in countries with stricter environmental rules commit more resources to green innovation, thereby improving their long-term value and global competitiveness (Kim et al., 2021). Taken together, these studies indicate that green innovation emerges from the combined influence of regulation, competition, and social preference, rather than from regulatory instruments alone.

2.1.2 Consequences of Green Innovation

The consequences of green innovation extend across environmental performance, firm outcomes, and capital-market reactions. In the environmental domain, Li et al. (2022) show that high-pollution firms adopting high-quality green innovation achieve meaningful reductions in future toxic emissions and contribute to advances in environmental technology. The type of innovation plays a central role: only pollution-prevention innovation improves both environmental and financial performance, while pollution-control innovation reduces emissions but has limited financial impact (Cheng et al., 2025). At the same time, green technological progress does not always lead to net emission reductions, as rebound effects and shifts in emissions across activities may offset expected gains (Bolton et al., 2022). Green innova-

tion also carries important implications for capital markets. F. Chen et al. (2024) show that although carbon-trading policies are, on average, associated with negative stock-market reactions, firms with stronger green innovation capabilities earn more favourable return outcomes, suggesting that green innovation operates as a key channel through which carbon regulation affects asset pricing. Hege et al. (2024) find that lucky grants of climate-related patents, despite their random component convey credible certification signals that raise stock valuations and reduce financing costs. Reza (2024) further demonstrate that the valuation of green innovation is highly sensitive to policy direction: it increases under stronger environmental commitments, such as the Paris Agreement, and declines during periods of policy uncertainty, such as the Trump election. Using a text-based measure constructed with ClimateBERT and GPT-3, Leippold and Yu (2025) show that firms with stronger green innovation, even within high-carbon industries, achieve superior environmental performance and lower expected returns, consistent with a risk-based interpretation of market pricing. Evidence from Cohen et al. (2020) also shows that traditional energy firms—often excluded by ESG investment screens—are leading producers of high-quality green patents and play a foundational role in the development of environmental technologies. Overall, the literature indicates that green innovation has systematic effects that reach beyond environmental outcomes. It influences firm performance, shapes market valuation, and generates spillover effects throughout supply-chain networks. These insights establish green innovation as a central construct in research on sustainable finance, corporate strategy, and environmental policy.

2.2 Artificial intelligence as a measurement tool in economics

2.2.1 The role of big data in ESG research

Recent studies in ESG and climate-related research increasingly rely on methods from natural language processing and machine learning to extract information from large volumes of unstructured text. Over time, there has been a clear progression in the techniques used,

beginning with simple word-count methods and advancing toward more complex language models and generative AI frameworks. Early contributions often employed bag-of-words approaches, where ESG exposure is measured by counting the frequency of selected keywords in firm disclosures or communications. For example, Sautner, Van Lent, et al. (2023b) measure firm-level climate change exposure using a custom-built climate dictionary and classify risk into physical, transition, and regulatory categories. Similarly, Nagar and Schoenfeld (2024) count references to the word “weather” in annual reports to construct a firm-level weather exposure measure. Other studies apply this method to environmental and social themes, such as Balakrishnan et al. (2023), who examine racial diversity using word frequencies in corporate filings. To better capture latent themes and reduce reliance on predefined dictionaries, subsequent work applies topic modelling methods such as Latent Dirichlet Allocation. This allows researchers to uncover underlying ESG-related topics in large corpora without manual classification. Ardia et al. (2023) construct a daily Climate Change Concerns Index based on news tagged as climate-related. Using correlated topic modelling, they identify 30 topics that are grouped into broader themes, and show that both transition and physical risks influence how stock prices respond to climate news. Building on topic models, more recent studies make use of word embedding techniques, which represent words as continuous vectors based on their context. These models allow for the measurement of semantic similarity and improve the identification of ESG-relevant content. Y. Lin et al. (2024) use word embeddings to quantify environmental and social disclosure from over 200,000 annual reports, documenting long-term global trends in disclosure quality. Briscoe-Tran (2025) applies similar techniques to employee reviews, constructing firm-level ESG measures based on employee perceptions, and finds that these scores are predictive of financial outcomes. Later studies incorporate transformer-based models such as BERT to enable more context-sensitive analysis. These models go beyond static word representations and can identify the meaning of terms based on sentence-level context. Kölbel et al. (2024) use BERT to classify climate risk disclosures into transition and physical risks, showing how these are reflected in

credit default swap spreads during key political events. Hu et al. (2023) combine LDA and FinBERT to assign firms’ climate-related disclosures to seven interpretable risk categories. Sautner, van Lent, et al. (2023) use BERT to detect climate-related content in earnings calls and show that investors react more strongly to financial implications than to ethical considerations. Most recently, generative AI models have been introduced to ESG research. These models, based on large-scale language generation architectures can summarise, classify, and produce textual insights across various formats. They offer new possibilities for extracting structured ESG information from unstructured sources, enabling more flexible and scalable analysis across firms, time, and risk categories.

2.2.2 Large language models as a general measurement tool

Recent advances in artificial intelligence have led to growing interest in the use of generative models for economic and financial research. Within this emerging literature, Mo and Ouyang (2025) provide a comprehensive review of the use of generative artificial intelligence in financial economics and classify existing studies into four main applications: prediction, information extraction and semantic analysis, task automation, and data generation. The authors show that large language models outperform traditional text-based methods in capturing economic meaning from unstructured data and in reducing information frictions in empirical research. Overall, the paper argues that generative AI is evolving from a supportive analytical tool into a key technology for constructing new economic data and measures, with important implications for research design in financial economics.

Jha et al. (2025) employ ChatGPT to derive forward-looking indicators from earnings call transcripts, and show that these expectation measures predict future macroeconomic activity and firm outcomes. Cao et al. (2025) apply Google Bard to identify product market peers for US listed firms and find that AI-based peers closely match expert benchmarks while outperforming traditional text-similarity methods in explaining return and accounting correlations. Bartik et al. (2024) use large language models to classify US municipal

housing regulations, achieving accuracy levels comparable to human coding. At a broader scale, Breitung and Müller (2025) build time-varying global business networks using GPT-3 applied to historical firm descriptions across many countries, and show that these networks improve predictions of stock return dynamics and merger activity relative to word-based approaches. Fetzner et al. (2024) construct a global production network by combining LLM-generated links with a systematic filtering step, revealing a shift in world trade toward more upstream and central products. Using a different application, Dyck et al. (2019) apply a retrieval-augmented LLM to measure controlling families’ environmental preferences from public information, and find that such preferences translate into lower emissions only when environmental concern is strong and abatement costs are low.

2.2.3 Patent texts as a source of innovation information

Recent studies use patent texts to extract information that is not captured by standard patent classifications or citation-based measures. Early contributions use textual similarity and language patterns in patent documents to study technological change and innovation dynamics. For example, Kelly et al. (2021) analyse the full textual content of patents to measure technological importance based on similarity to past and future inventions. Their approach identifies breakthrough innovations and tracks long-run waves of technological change across sectors and time.

Subsequent studies employ patent texts to examine firm-level economic outcomes. Harford et al. (2025) combine patent texts with product filings to construct a measure of patent utilisation, capturing the extent to which patented technologies are applied in new products. Using text-based matching between patents and products, they show that higher patent utilisation is associated with stronger market reactions to product launches and improved firm performance. Recent work further applies advanced text-embedding and machine learning methods to map technological relationships across patents. Kelly et al. (2021) use embedding-based representations of patent texts to measure patent-to-patent technological

similarity and construct large-scale innovation networks. Their framework allows researchers to study knowledge flows, technological distance, and patent quality using the semantic content of patent documents rather than predefined classes.

Another strand of the literature uses patent texts to identify specific types of innovation relevant for policy and firm outcomes. M. A. Chen and Wang (2024) apply large language models to patent texts to classify artificial intelligence innovations into functional categories. Using this text-based classification, they study how different types of AI innovation affect employment, productivity, and firm value. Their findings show that patent texts can be used to distinguish economically meaningful forms of innovation within the same broad technology domain. Large language models have also been applied to study changes in intellectual property protection. Acikalin et al. (2025) use LLMs to analyse patent portfolios and identify firms exposed to the weakening of patent protection following the Alice Supreme Court decision. Their text-based measures allow them to examine heterogeneous firm responses in innovation, competition, acquisitions, and employment, demonstrating the value of patent texts in capturing legal and economic exposure that is not observable from patent counts alone. Zheng (2025) uses patent texts to examine how investors value innovation. Focusing on the brief description of the invention section of patents, Zheng (2025) applies a BERT-based model to extract contextual information from patent texts and shows that this information explains a large share of the variation in patent-level stock market valuation. The study further finds that investors underreact to information contained in patent texts and that this underreaction declines after mandatory pre-grant disclosure is introduced. These results highlight patent texts as an important information source for understanding firms' internally developed intangible assets.

These studies demonstrate that patent texts have been used for a wide range of research purposes, including measuring technological similarity, identifying breakthrough innovation, linking patents to products, classifying policy-relevant technologies, assessing legal exposure, and explaining market valuation. They also show that different sections of patent documents

serve different roles. While full texts and abstracts are often used for large-scale similarity and classification tasks, more descriptive sections that combine technical content with human judgement are particularly informative for studying the economic value of innovation. This literature motivates our use of patent texts to identify climate-related inventions beyond existing classification schemes.

3 Hypotheses

Climate-related innovation can be interpreted by capital markets in fundamentally different ways. One view emphasises the potential of climate innovation to enhance firms' long-term growth opportunities and competitive position Cheng et al. (2024). Climate-related patents may reflect the development of new technologies that allow firms to adapt to stricter environmental regulation, benefit from policy support, and gain access to emerging low-carbon markets. From this perspective, climate innovation represents an investment in future productive capacity rather than a short-term expense. Firms that successfully develop climate-related technologies may be better positioned to exploit new demand, improve efficiency, and secure first-mover advantages as the economy transitions towards cleaner production. In addition, climate patents can strengthen firms' intangible asset base by signalling technological capability and organisational commitment to innovation. Such signals may be valued by investors who expect climate-related technologies to generate long-run economic benefits. Under this view, equity markets should capitalise the expected future gains from climate innovation into firm value, even if the immediate financial returns are limited Sautner, van Lent, et al. (2023). As a result, firms with greater climate-related patenting activity should command higher valuation multiples, as reflected in a higher Tobin's q . This argument is consistent with a growth-option interpretation of climate innovation, whereby patents embody opportunities for future expansion that are not yet fully realised in current cash flows. If this mechanism dominates, climate-related patenting should be associated with higher market

valuations over time.

H1a (Value-Enhancing Innovation Hypothesis). Firms with greater climate-related patenting activity exhibit higher future market valuation.

An alternative view stresses the costs and risks associated with climate-related innovation. Climate patents often arise in response to regulatory pressure, environmental constraints, or the need to comply with evolving standards rather than purely profit-driven motives. Developing climate-related technologies typically requires substantial upfront investment in research, capital equipment, and organisational restructuring. These investments may divert resources away from core operations and generate limited short- to medium-term cash flow. Moreover, the economic returns to climate innovation are highly uncertain. The commercial success of climate technologies depends on future regulation, policy enforcement, technological standards, and market adoption, all of which are difficult to predict. As a result, climate-related patenting may increase earnings volatility and heighten uncertainty about future profitability. From a valuation perspective, investors may view climate innovation as a costly transition activity that imposes financial burdens before its benefits materialise. Even if climate technologies yield long-run advantages, markets may discount firms that engage heavily in such innovation due to concerns about delayed pay-offs, execution risk, and policy uncertainty. Under this interpretation, climate-related patenting serves as a signal of higher adjustment costs and risk exposure rather than immediate value creation. Consequently, equity markets may apply a valuation discount to firms with greater climate patenting activity. This view implies that climate innovation is priced negatively, at least over the medium term, reflecting the perceived costs and risks of the transition process.

H1b (Cost-and-Uncertainty Hypothesis). Firms with greater climate-related patenting activity exhibit lower future market valuation.

4 Data sources

4.1 Data on climate patents

We obtain firm-level utility patent and citation data from Kogan et al. (2017), which provides a comprehensive patent–firm matched dataset updated through 2023². This dataset includes patent identifiers, grant dates, measures of patent value, forward citations, and full Cooperative Patent Classification (CPC) codes, as well as CRSP permanent identifiers (PERMNO) that allow patents to be linked to publicly listed firms. We use this dataset as the backbone of our patent sample and follow its firm–patent matching throughout the analysis. To construct the text corpus used in the LLM-based classification, we obtain patent full-text data from the United States Patent and Trademark Office (USPTO)³. Specifically, we download patent grant full-text files in XML format from the *Patent Grant Full-Text Data* collection and recover the complete textual content of granted patents, excluding figures and drawings. These sources allow us to construct a consistent patent text corpus for the full sample period. Because the Kogan et al. (2017) dataset does not report patent application filing dates, we supplement the patent-level information using data from PatentsView⁴. This

²The extended dataset of Kogan et al. (2017) is available at <https://github.com/KPSS2017/Technological-Innovation-Resource-Allocation-and-Growth-Extended-Data>. It provides updated KPSS values and forward citations, a patent number to CRSP permno match, and a patent number to CPC class match.

³Patent Grant Full-Text Data in XML format is available at <https://data.uspto.gov/bulkdata/datasets/PTGRXML>. This source provides the concatenated full text of weekly patent grant documents associated with applications filed from 2001 onward. For patents granted prior to 2001, we rely on the Patent Grant Full-Text Data–APS files, available at <https://data.uspto.gov/bulkdata/datasets/ptgraps>, which provide full-text coverage of granted patents from 1976 onward.

⁴PatentsView data are available at <https://patentsview.org>. We use this source to obtain patent application filing dates and merge them to our main patent sample using patent identifiers.

step allows us to align patent timing more precisely and to construct measures based on application-year activity when required. In this study, we measure climate-related innovation using a large language model (LLM)-based classifier, which serves as our primary measure. The LLM approach identifies climate-related patents based on the textual content of patent documents, allowing us to capture climate innovation beyond predefined classification rules. For comparison and validation purposes, We also construct a conventional measure of green innovation based on the Y02 cooperative patent classification scheme. Under this approach, a patent is classified as green if it is assigned to at least one Y02 subclass, and Y02-tagged patents are aggregated to the firm level following standard practice in the literature. Using the Y02-based measure as a benchmark enables us to evaluate the extent to which the LLM-based classifier aligns with, and extends beyond, the traditional classification-based definition of climate-related patents. This comparison provides a consistent framework for assessing the incremental information captured by the text-based LLM measure relative to the established classification scheme.

4.2 Data on Carbon Emissions

We obtain firm-level carbon emission data from S&P Trucost, following the data construction approach used in prior studies. Trucost reports corporate emissions in accordance with the Greenhouse Gas Protocol, which provides widely accepted standards for measuring greenhouse gases. Our analysis focuses on Scope 1 emissions, which capture direct emissions from a firm’s own operations, and Scope 2 emissions, which reflect indirect emissions from purchased electricity, steam, heating, or cooling. We use the Trucost identifiers to link these measures to our firm sample and retain the observations that have consistent reporting across the years covered by our study.

4.3 Data on Supply Chain

We construct our supply-chain dataset by following the screening and sample-building procedures used in Schiller (2017) and Dai et al. (2021). We combine information from two major sources: the Compustat Customer Segment data, which captures relationships disclosed in firms’ 10-K filings, and the FactSet Revere Supply Chain database, which provides a broader set of supplier–customer links gathered from public disclosures, corporate filings, and other market sources. Each record contains the names and identifiers of the supplier and customer, along with the dates of the relationship and, when available, sales information. Consistent with the prior literature, we restrict both suppliers and customers to firms in the CRSP–Compustat universe and treat a supplier–customer link as active from the first to the last year in which the relationship appears. We use SegLink to perform exact name-based matching across the two datasets to ensure more precise identification of firm pairs. All other filtering steps and sample definitions follow the established approach in the literature.

4.4 Data on Environmental Enforcement

Information on environmental enforcement activity is obtained from the U.S. Environmental Protection Agency’s Enforcement and Compliance History Online (ECHO) database. The ECHO data provide records of environmental inspections conducted at the regional level, which we aggregate to the state–county–year level. We use these aggregated inspection counts as a measure of regional regulatory enforcement intensity faced by firms operating in those locations.

To assign this regional enforcement information to individual firms, we first obtain firm location data at the state and county level from the Dun & Bradstreet (D&B) database. These location identifiers allow us to map firms to the corresponding geographic units in the ECHO data. Because firm names in Compustat reflect only current identifiers and do not consistently track historical name changes, we augment firm identifiers using historical company names collected from SEC filings available through WRDS, including Form 10-

K and Form 8-K disclosures. We then use fuzzy name matching to link firms across the Compustat, D&B, and ECHO datasets, ensuring consistent firm identification over time.

4.5 Financial Statement Data

Firm-level financial data are obtained from Compustat North America and are used to construct standard accounting and market-based variables, including firm size (log assets), Tobin’s Q , asset tangibility, and leverage. Observations with missing industry classifications are excluded. To focus on non-financial operating firms, we remove companies in regulated utilities and the financial sector, identified by two-digit SIC codes between 4900–4999 and 6000–6999.

4.6 Summary Statistics

Our analysis covers granted U.S. patents from 2000 to 2023. Using the CPC Y02 classification, we identify 90,710 climate-related patents over this period, whereas the LLM-based classifier identifies 27,018 patents from the same universe. The difference in scale reflects the more selective nature of the LLM approach, which focuses on the substantive climate function of inventions rather than formal technology codes.

Figure 1 illustrates the cumulative geographical distribution of climate-related patents across U.S. states. Under the Y02 patent classification scheme, climate patenting is highly concentrated, with California accounting for a disproportionately large share of total patents, followed by a small group of large industrial states. In contrast, the LLM-based measure produces a less concentrated pattern, with a stronger presence of Midwestern states such as Michigan and Ohio, while California remains important but no longer dominates. This comparison suggests that the two measures overlap in identifying major innovation hubs, but they differ in the regional emphasis they place on climate-related technological activity.

Industry-level differences are shown in Figure 2, which compares green patent proportions across major U.S. industries under the two classification methods. Climate-related

patenting is most prominent in heavy industrial and energy-related sectors, including industrial chemicals, engines and turbines, and petroleum refining. While both measures highlight broadly similar industries, the relative green shares vary across sectors. In several manufacturing-intensive industries, the LLM-based classifier assigns a higher proportion of patents to climate-related activity, indicating that it captures climate-relevant innovations embedded within conventional industrial technologies that may not be explicitly classified as green under CPC codes.

Time-series patterns are reported in Figure 4. Both classification methods show rising climate-related patenting from the early 2000s into the 2010s, consistent with the growing role of climate and energy technologies. However, the Y02-based series increases more sharply and reaches a substantially higher peak before declining in the later years of the sample. By contrast, the LLM-based series remains lower throughout and follows a smoother trajectory over time, reflecting a narrower and more stable set of climate-relevant inventions identified through semantic screening. The technology-level trends within the LLM measure further show that variation over time is driven mainly by mitigation-related patents, particularly those related to emission reduction.

Besides, Figure 5 summarises the composition of climate-related patents across hierarchical technology categories. Mitigation accounts for the vast majority of identified climate patents, with adaptation representing only a small share. Within mitigation, emission reduction technologies dominate carbon dioxide removal. Emission reduction activity is mainly split between renewable substitution and energy-efficiency improvements, while the relatively small group of carbon dioxide removal patents is concentrated in direct air carbon capture and storage, with smaller shares in afforestation and reforestation and very limited activity in BECCS. Taken together, these descriptive statistics show that the LLM-based measure not only yields a smaller and more targeted set of climate patents than Y02, but also provides a distinct view of the geographic, industrial, and technological structure of climate-related innovation.

5 Methodology : Using Generative AI to identify climate patents

The identification of climate-related patents plays a central role in empirical research on green innovation, technological change, and climate policy. Earlier work in finance and economics has relied heavily on traditional text-analysis tools, such as text filters or word-embedding methods, which classify documents based on the frequency or co-occurrence of selected terms. These approaches are transparent and computationally simple, but they capture only a narrow portion of the semantic information contained in technical documents. As a result, they often struggle to distinguish inventions that meaningfully contribute to climate mitigation or adaptation from those that merely contain related terminology.

Recent advances in generative AI have altered the way researchers process textual information. Large Language Models (LLMs), including ChatGPT, Gemini, and others, allow documents to be interpreted through direct queries framed around concepts rather than tokens. Trained on large and diverse corpora, these models possess strong semantic understanding and can evaluate technical content with greater precision than rule-based NLP systems. This advantage is particularly relevant for patent analysis, where abstracts often contain dense and domain-specific language.

The most widely adopted method for identifying climate-related patents has been the Y02 classification developed by the European Patent Office (EPO). Introduced in 2010 and incorporated into the Cooperative Patent Classification (CPC) in 2013, Y02 was designed to capture technologies that contribute to climate-change mitigation across areas such as renewable energy, energy efficiency, carbon capture, transportation, industrial processes, and waste management. The system provides a standardised and transparent framework for tracking climate-oriented innovation across jurisdictions. Despite its prominence, Y02 is an imperfect proxy. Because it relies on a fixed set of codes, it may include inventions with limited relevance to climate mitigation or fail to capture emerging or cross-domain technolo-

gies that contribute to mitigation or adaptation but do not fit the established classification scheme. These limitations are well recognised in the literature and are particularly relevant for modern innovation, where climate technologies increasingly integrate digital systems, materials science, and industrial engineering.

To address these limitations, we employ a controlled and fully replicable LLM-based framework to identify climate-related patents. Our approach is implemented using a locally deployed instance of Qwen3-32B, a 32-billion-parameter large language model released by Alibaba under an Apache 2.0 open-source license in April 2025. Qwen3-32B is trained on extensive multilingual corpora that include technical and scientific material, making it well suited for interpreting patent language. The model supports long-context inference and includes a dedicated reasoning mode, which enhances its ability to evaluate technical descriptions in climate-related innovations. Running the model locally ensures full control over computation, software versions, model weights, and data security, which is important given the reproducibility concerns associated with closed-source models whose parameters evolve over time.

Given the large scale of the patent universe and the computational cost of model inference, our identification strategy proceeds in two stages that align naturally with the structure of patent documents. In the first stage, we rely exclusively on patent abstracts to determine whether an invention is climate related. Abstracts provide a concise and standardised summary of the invention’s core purpose and technical contribution, making them suitable for large-scale screening across jurisdictions and filing years. Using abstracts at this stage allows us to efficiently construct a binary indicator of climate-related innovation while preserving the semantic advantages of the LLM-based approach.

A key challenge in using LLMs for classification is the potential for stochastic output. To eliminate randomness, all parameters governing stochasticity, including temperature, top-k, and top-p, are set to zero, ensuring that identical inputs yield identical outputs. To further minimise ambiguity in interpretation, we use a standardised and concise instruction

for every observation. The model is instructed to evaluate only the patent abstract and to return a categorical “yes” or “no” answer in a fixed JSON format. Climate-related patents are defined as those involving technologies that reduce, control, or prevent anthropogenic greenhouse-gas emissions, consistent with frameworks such as the Kyoto Protocol and the Paris Agreement, as well as technologies that support adaptation to climate risks. The classification is generated using the following prompt:

Prompt: You are an expert in climate-related patents. Based only on the following abstract, determine whether the invention qualifies as a climate-related patent. Climate-related patents include technologies that control, reduce, or prevent greenhouse-gas emissions, and technologies that support adaptation to climate change. Read the abstract carefully and respond with just YES or NO.

This procedure yields a consistent binary measure of whether a patent is climate related, which we use throughout the analysis. The LLM-based indicator complements the Y02 classification by capturing climate-relevant content that formal CPC categories may overlook. The appendix provides illustrative examples where Y02 tags and LLM assessments diverge, including patents classified as climate related under Y02 that lack substantive climate relevance, as well as patents without Y02 tags that are identified by the LLM as climate focused.

In the second stage, we restrict the sample to patents that are identified as climate related in the abstract-screening step and assign them to more detailed technology categories. We do not rely on abstracts for this step, as they are intentionally brief and often lack the technical detail needed to distinguish between closely related climate functions. Instead, we use the *Brief Description of the Invention* section, which typically describes the technical goal of the invention, its main mechanism, and the problem it is intended to solve. This information provides a clearer basis for identifying the patent’s primary climate role. Using this section, we apply a hierarchical classification based on IPCC definitions of mitigation and adaptation. The detailed prompt and classification rules used in this stage are reported in the appendix.

Mitigation patents are those whose main purpose is to reduce, control, or avoid greenhouse-gas emissions, or to remove greenhouse gases from the atmosphere. Adaptation patents are those designed to reduce exposure or vulnerability to climate risks and to improve resilience to climate impacts. Within mitigation, we further distinguish between emission reduction technologies, which lower emissions from energy use, production, transport, buildings, or industrial processes, and carbon dioxide removal (CDR) technologies, which remove CO₂ from the atmosphere and store it durably. We then group mitigation patents into specific technology classes that reflect common transition pathways discussed by the IPCC, including energy efficiency, renewable substitution, bioenergy with carbon capture and storage, direct air carbon capture and storage, and afforestation and reforestation. Throughout this step, the model is instructed to assign a single label based on the main technical contribution described in the invention text. This approach preserves efficiency in the initial screening stage while allowing a more precise and economically meaningful classification of climate innovation in the second stage.

6 Validation of the LLM-Based Measure of Climate Innovation

6.1 Internal Validity

A central step in validating the LLM-based climate patent measure is to demonstrate that it captures the same underlying construct as an established benchmark. Table 2 reports regressions that relate the LLM measure to the widely used Y02 climate patent classification. If the two measures reflect a common dimension of climate-related innovation, we should observe a strong and stable correlation after accounting for systematic differences across patent families and years.

The evidence is consistent with this expectation. Columns (1) and (2) of Table 2 show a

strong and stable association between the two measures. After controlling for firm characteristics, including size, Tobin’s q, leverage, tangibility, capital expenditure, R&D, operating cash flow profitability as well as year and patent-family fixed effects, the coefficient on y02 remains large and highly significant across specifications. Patents labeled as Y02 receive LLM climate scores that are 13–14 percentage points higher than other patents. This pattern holds whether we absorb industry heterogeneity or firm-level unobserved differences, indicating that the LLM classifier captures the core technological features that underpin the Y02 classification.

At the same time, the explanatory power of the regressions is modest, with R-squared values between 0.14 and 0.17. This limited fit suggests that the LLM measure is not simply a restatement of the Y02 label. Instead, the LLM classifier appears to extract additional information from patent text that the coarse CPC system does not capture. In particular, the LLM score varies even within Y02-tagged patents and assigns meaningful climate relevance to patents outside the Y02 classes, pointing to a broader and more nuanced interpretation of climate-related innovation.

Taken together, these results provide two important insights. First, the LLM measure aligns well with the established Y02 system, confirming that it captures genuine climate-related technological content rather than superficial wording. Second, the LLM-based measure captures additional variation that is not captured by the Y02 classification. This incremental content is the basis for our later findings that the LLM-based measure has stronger predictive power for firms’ environmental actions and market responses than the traditional patent-classification approach.

[Insert Table 2 about here]

6.2 External Validation: Patent-Level Predictive Validity

Table 3 assesses whether the climate patent indicators capture economically meaningful variation at the patent level. Panel A relates each measure to a market-based index of

patent value, while Panel B examines their association with forward citations. All regressions include year and patent-family fixed effects, and standard errors are clustered at the firm level. Across both panels, the LLM-based indicator shows a markedly stronger connection to economically meaningful outcomes than the traditional Y02 tag. In the value regressions, the LLM coefficient is 0.120 in the baseline specification and remains positive and significant once firm fixed effects are introduced (0.024). The corresponding Y02 estimates are smaller in magnitude and lose statistical significance when firm fixed effects are absorbed. The citation regressions reveal an even sharper contrast. Patents classified as climate-related by the LLM measure receive substantially more forward citations, with coefficients of 1.028 and 0.898, whereas the Y02 indicator exhibits no meaningful relationship with citations in either specification. Taken together, the results indicate that the LLM-based measure identifies patents that are more valuable and more central in subsequent technological developments, suggesting that it captures economically relevant aspects of climate innovation that the Y02 classification does not.

[Insert Table 3 about here]

6.3 Market-Based Validation

Climate policy announcements alter expectations about future regulatory costs and technology demand, and firms differ in how exposed they are to such changes. This structure helps explain why climate-related events generate heterogeneous stock-price reactions. Trump’s victory in 2016 signalled a shift toward weaker environmental regulation, including withdrawal from the Paris Agreement, reduced limits on fossil-fuel extraction, and the dismantling of the Clean Power Plan. These plans lowered expected compliance costs for firms with limited engagement in green technologies. For firms that had invested more heavily in climate-related innovation, the policy shift implied weaker future demand for their technologies and reduced support for transition efforts. Investors adjusted valuations accordingly, producing negative abnormal returns for firms with stronger climate-innovation portfolios.

In this setting, the LLM-based patent measure offers a more informative view of firms’ exposure to regulatory relaxation, generating sharper and statistically stronger return responses than the traditional Y02 scheme.

The Inflation Reduction Act presents the opposite regime. The IRA commits substantial and long-term federal support to clean technology through subsidies, tax incentives, and explicit emissions-reduction goals. These provisions strengthen the transition path and improve expected cash flows for firms whose technologies complement a low-carbon economy. Firms with deeper climate-innovation capabilities therefore experience positive price adjustments when policy credibility increases.

We implement a regression model for each event, using cumulative abnormal returns (CARs) as dependent variables to capture how firms react to new climate-related information:

$$CAR_i = \alpha + \beta \text{Climate patent}_i + \gamma X_i + a_j + \varepsilon_i,$$

where $CAR[0, 1]$, $CAR[0, 5]$, and $CAR[0, 10]$ measure abnormal returns over one, five, and ten trading days after the event date. All variables, including the log-transformed patent measures, are standardized to zero mean and unit variance. Industry fixed effects are included, and t -statistics are clustered by industry.

Panel A of Table 4 shows the results for Trump’s election. The LLM measure yields a clear and negative response. A one-standard-deviation increase in the LLM measure is linked to a 0.9% decline in one-day CARs, a 0.7% decline over five days, and a 1.2% decline over ten days, with the first and tenth-day effects statistically significant. These results are consistent with investor views at the time. Trump’s policy stance withdrawal from the Paris Agreement, support for fossil-fuel expansion, and the rollback of federal climate rules reduced the expected burden on carbon-intensive firms. For firms with stronger climate-related innovation, the shift signaled weaker demand for their technologies and fewer incentives for firms to pursue low-carbon projects. Investors marked down these firms accordingly. The

Y02 measure, in contrast, produces much smaller and mostly insignificant estimates. This suggests that the narrower Y02 category does not capture the aspects of climate innovation that investors considered relevant when rules were expected to loosen.

Panel B of Table 4 reports the results for the IRA. Here the pattern reverses. A one-standard-deviation increase in the LLM measure predicts a 1.7% rise in one-day CARs, 2.1% over five days, and 2.5% over ten days, with all estimates statistically significant. The IRA set out long-term federal support for clean technology through subsidies, tax credits, and ambitious emissions targets. These commitments raise expected cash flows for firms whose innovation efforts align with a low carbon transition. Investors reacted by assigning higher valuations to firms with stronger climate-innovation portfolios. Again, the Y02 measure shows little systematic response, with small coefficients and no statistical significance.

Taken together, these results show a consistent economic mechanism. Policy signals that weaken climate rules lower the market value of firms with strong climate patents, while policy signals that strengthen regulation increase it. The LLM measure tracks these shifts closely and produces clear, intuitive, and stable patterns across both events. The Y02 measure, by comparison, shows limited explanatory power. The evidence indicates that the LLM approach captures a broader and more economically relevant dimension of climate innovation—one that aligns more closely with how investors interpret major changes in climate policy and adjust their expectations about regulatory costs, technology demand, and long-run competitive outcomes

[Insert Table 4 about here]

6.4 Validity of Climate Innovation Responses to Regulatory and Informational Signals

This section examines whether firms adjust their climate-related innovation in response to both external and internal pressures, and whether such adjustments are more clearly reflected

by the LLM-based innovation measure than by Y02. We consider two complementary sources of pressure. The first is external and institutional, captured by regulatory inspections that reflect heightened oversight and enforcement attention. The second is internal and forward-looking, captured by climate-related language in earnings conference calls that reveals firms' own expectations, concerns, and strategic priorities. Together, these analyses assess whether firms increase climate innovation when facing stronger regulatory scrutiny or heightened internal focus on climate issues, and whether the LLM-based measure is better suited to capture these behavioural responses than traditional patent classifications.

We rely on variation in regulatory inspection intensity obtained from the U.S. Environmental Protection Agency's (EPA) Enforcement and Compliance History Online (ECHO) database. Using records of firms' environmental compliance inspections, we aggregate the total number of inspections faced by a firm at a given location in a given year at the firm–county–year level and construct the variable *lninspect_tot*. This measure captures the intensity of regulatory oversight faced by a firm in a specific time and place, rather than local air quality conditions or the firm's actual pollution emissions.

In economic terms, although these inspections mainly target conventional air pollutants and do not directly regulate carbon dioxide emissions or climate policy objectives, they reflect the degree of regulatory attention and enforcement effort in a given region and period. More frequent inspections convey a credible signal that regulatory scrutiny is increasing, thereby shaping firms' expectations about future regulatory uncertainty, compliance requirements, and potential enforcement risks. Accordingly, *lninspect_tot* is not interpreted as a direct climate policy instrument, but as a proxy for exogenous regulatory attention faced by firms. Through this signalling channel, regulatory oversight influences firms' assessments of the future regulatory environment and, in turn, their long-term strategic and innovation decisions, including investments in climate-related technologies.

Table 5 examines whether such regulatory oversight is reflected in firms' subsequent climate-related innovation. This mechanism builds on the long-standing view that external

monitoring alters firms’ compliance incentives and innovation choices. Theoretical work in environmental economics suggests that firms subject to more frequent inspections internalize the risk of future regulatory actions and therefore invest in technologies that lower compliance costs or reduce exposure to environmental sanctions. Consistent with this view, Columns (1) and (3) show that higher inspection intensity in year t is followed by a clear increase in the LLM-based climate innovation measure in years $t+1$ and $t+2$. The effect remains stable across specifications that control for a broad set of firm characteristics and include firm, county, and year fixed effects, indicating that the results are not driven by persistent firm traits or local regulatory conditions. These patterns suggest that inspections act as a credible signal of regulatory attention, prompting firms to undertake additional climate-oriented innovation.

By contrast, the Y02-based innovation measures reported in Columns (2) and (4) show no statistically reliable response to inspection intensity. Although the estimated coefficients have the same sign as those based on the LLM measure, they are economically small and lose precision once standard controls and fixed effects are included. This contrast indicates that the LLM-based measure is more sensitive to regulatory motivations for innovation, capturing forms of climate-related innovation that firms undertake in response to regulatory pressure but that fall outside the traditional technology classifications embedded in the Y02 classification. Overall, the evidence suggests that firms adjust their climate-related innovation following regulatory scrutiny, and that this adjustment is revealed only when climate innovation is measured using the LLM-based approach.

Table 5 examines whether climate-related language in earnings conference calls is systematically associated with firms’ climate innovation activity, and whether this association is more clearly captured by the LLM-based measure than by Y02. The analysis uses firm-level measures of climate-related language from (Sautner, Van Lent, et al., 2023a) ⁵. The esti-

⁵The earnings-call climate language measures are from Sautner et al. (2023) and are constructed from earnings conference call transcripts. The data quantify financial analysts’ attention to firms’ climate change exposure and distinguish between opportunity, physical, and regulatory dimensions of climate risk. The dataset covers more than 10,000 firms from 34 countries over the period 2002–2021 and is publicly available at <https://www.df.uzh.ch/en/sustainable-finance/data-center/data-firm-level-climate-change-exposure.html>.

mates indicate a systematic association between climate-related language in earnings calls and subsequent climate innovation when innovation is measured using the LLM-based indicator. Firms with greater climate-related discussion in earnings calls tend to exhibit higher levels of climate innovation as captured by the LLM measure. In contrast, we find no robust evidence of a comparable relationship when climate innovation is measured using the Y02 classification.

The difference is especially pronounced for climate-related regulatory language. Mentions of regulatory issues are associated with higher climate innovation under both measures, but the estimated effects are larger and more precisely measured for the LLM-based indicator. This pattern indicates that regulatory discussion in earnings calls coincides with firms' adjustments in climate-related innovation, and that such adjustments are captured more effectively by the LLM-based measure. Overall, the evidence shows that earnings-call climate language, particularly on regulatory topics, is closely linked to firms' climate innovation activity, and that this link is measured more clearly by the LLM approach than by Y02.

[Insert Table 5 about here]

6.5 Validation through supply-chain environmental outcomes

In this section, we examine whether climate innovation by upstream suppliers shapes the direct carbon emissions of downstream firms. This question speaks to a broader point about how environmental technologies travel across production networks. Many climate-related technologies are embodied in intermediate goods rather than sold as final products. As a result, the customers who purchase and integrate these goods into their own production processes are often the parties most affected by innovation undertaken upstream. This view is consistent with the patterns observed in patent data, where a large share of climate patents relate to product improvements that become part of a broader supply chain. Our analysis focuses on Scope 1 emissions, which capture emissions generated directly by firms themselves. These emissions should be the first to respond when suppliers introduce technologies that

alter the carbon content of inputs or improve production efficiency. While other emission categories may also change, our current data only allow us to study the direct operational channel.

To evaluate this mechanism, we construct a customer-year panel by matching supplier-customer links from the FactSet Revere Supply Chain database with sales data from the Compustat Segment files. The two datasets are complementary: FactSet Revere provides detailed information on the structure of production networks, while Compustat identifies major customers and the value of sales flowing through each link. Their intersection allows us to observe not only whether a firm is connected to a supplier, but also the relative economic importance of that supplier. For each supplier-year, we measure climate innovation as the share of climate-related patents in total filings. At the customer level, we compute a sales-weighted average across all suppliers, producing a variable w_{cpr} that reflects the intensity of climate-oriented innovation embodied in the inputs a firm receives.

We estimate the following specification:

$$\Delta_{t,t+k} \ln(\text{Emissions}_i) = \beta \text{SupplierClimatePatentRatio}_{i,t} + \gamma \mathbf{X}_{i,t} + \delta_{\text{industry} \times t} + \varepsilon_{i,t}, \quad (1)$$

where

$$\Delta_{t,t+k} = \ln(\text{Emissions}_{i,t+k}) - \ln(\text{Emissions}_{i,t}) \quad (2)$$

captures the forward-looking change in firm i 's Scope 1 emissions over horizons of one to five years. Because the dependent variable is expressed in differences, firm fixed effects are not included. All regressions contain industry-by-year fixed effects to absorb sectoral decarbonization trends, and standard errors are clustered at the firm level.

Panel A of Table 6 shows the results obtained using the climate patent classification generated by the LLM method. The coefficients for horizons $t + 1$ and $t + 2$ are negative but imprecise, which is consistent with a gradual diffusion mechanism. Climate technology does not immediately reshape downstream production structures; its effects materialize only

once customers adopt the inputs in which the new technologies are embedded. Beginning in year $t + 3$, the coefficients become large and statistically significant. The estimate of -1.094 in column (3) implies that a 0.1 increase in w_{cpr}^{LLM} is associated with roughly a 10% reduction in a customer’s Scope 1 emissions three years ahead. The effects remain strong in years $t + 4$ and $t + 5$, pointing to a persistent decline in direct emissions once the technology has been fully integrated into downstream operations. These patterns suggest that climate innovation propagates through supply chains in a slow but meaningful way, with measurable reductions in emissions emerging only after firms have had time to adjust their production processes. Table 6 Panel B provides the corresponding results using the Y02 classification. In sharp contrast to the LLM measure, the coefficients on w_{cpr}^{Y02} are small and statistically insignificant across all horizons. The Y02 scheme captures a narrower subset of climate technologies, and the results suggest that it does not pick up the forms of innovation that matter for downstream emissions. The stronger and more stable performance of the LLM-based measure therefore points to the importance of using broader and more flexible classification methods when studying the environmental impact of innovation.

These findings show that climate innovation undertaken by suppliers can significantly reduce the direct emissions of downstream firms, but only when the innovation is measured in a way that captures its full technological scope. The delayed but persistent effects across multiple years are consistent with a story in which climate technologies diffuse gradually through supply chains and alter production practices once adoption is complete. While our analysis is limited to Scope 1 emissions and does not capture changes in energy purchases or broader value-chain emissions, the evidence highlights the role of supply-chain channels in shaping firm-level environmental outcomes and underscores the value of LLM-based classification tools for identifying the technologies that drive these changes.

[Insert Table 6 about here]

7 Market Valuation of Climate Innovation

An important question in the climate finance literature is how capital markets value firms’ climate-related activities. While climate innovation may enhance firms’ long-term competitiveness, it often requires substantial upfront investment and is associated with technological and regulatory uncertainty. As a result, it is not clear whether equity markets perceive climate innovation as value enhancing or as a short-term cost burden.

$$\text{Tobin's } q_{i,t+k} = \beta \cdot \text{Climate_LLM}_{i,t} + \gamma \cdot X_{i,t-1} + \zeta_{j,t} + \varepsilon_{i,t}, \quad (3)$$

To examine this issue, we study the relationship between firms’ climate innovation and their market valuation, measured by Tobin’s q . Tobin’s q , defined as the ratio of a firm’s market value to the replacement value of its assets, is widely used to capture the market valuation of both tangible and intangible assets. Table 7 reports the baseline results. We regress Tobin’s q in future years on the LLM-based climate innovation measure, controlling for firm size, asset tangibility, leverage, capital expenditures, R&D intensity, operating cash flow, and profitability. All specifications include firm and year fixed effects, which allow us to exploit within-firm variation over time. Columns (1) to (5) examine Tobin’s q over increasing horizons, up to five years ahead.

Across all specifications, the coefficient on the climate innovation measure is negative and statistically significant. The magnitude of the estimated effect is stable across horizons and becomes slightly larger at longer horizons. This pattern indicates that the valuation effect does not reverse over time. Economically, the estimates suggest that firms with greater climate innovation exposure are valued at a discount relative to otherwise similar firms. These findings are consistent with the view that equity markets associate climate innovation with higher adjustment costs and uncertainty in the medium term. Although climate-related technologies may generate long-run benefits, they often require substantial investment and face uncertain payoffs. As a result, markets appear to price climate innovation as a costly

transition activity rather than immediately rewarding it with higher valuations. Overall, the evidence suggests that capital markets incorporate climate innovation into firm valuation, primarily through a valuation discount that reflects the perceived costs and risks of the transition process.

We further examine whether the valuation effects of climate innovation vary across different types of climate-related technologies. Moving beyond the aggregate climate patent measure, we progressively refine the classification of climate innovation to identify the sources of the valuation discount. A clear distinction emerges between adaptation and mitigation innovations. Adaptation-related patents refer to technologies that help firms adjust to the physical impacts of climate change, such as extreme weather or changing environmental conditions, without directly reducing emissions. Mitigation-related patents capture technologies aimed at lowering greenhouse gas emissions or shifting production processes toward lower-carbon alternatives. While adaptation-related patents are not significantly associated with future firm valuation, mitigation-related patents exhibit a consistently negative relationship with Tobin's q across all horizons. This pattern indicates that the valuation discount observed for climate innovation is not general, but is concentrated in technologies directly linked to emissions reduction and the transition of production processes. Additional disaggregation within mitigation technologies reveals substantial heterogeneity in market responses. Patents related to emission reduction are associated with an immediate and persistent valuation discount, whereas patents related to carbon dioxide removal display little association with firm value in the short run, with a negative effect emerging only at longer horizons. This difference suggests that markets distinguish between mitigation pathways based on the timing and uncertainty of expected payoffs, particularly the reliance of removal technologies on future scale-up and policy support. A more granular breakdown by specific mitigation technologies reinforces this interpretation. Innovations related to energy efficiency are not associated with significant valuation effects, consistent with their incremental nature and relatively predictable benefits. Renewable substitution technologies are linked to a modest

valuation discount that weakens over time. By contrast, technologies that require substantial upfront investment and face greater implementation and regulatory uncertainty, such as DACCS and afforestation and reforestation, are associated with pronounced valuation discounts at longer horizons. BECCS exhibits a short-term positive association with firm value, which does not persist as the investment horizon lengthens. These results show that equity markets do not uniformly discount climate innovation. Instead, market valuation depends on the technological characteristics of climate-related activities. Innovations that involve higher capital intensity, longer deployment horizons, and greater uncertainty are priced at a discount, whereas more incremental or efficiency-enhancing innovations are not.

[Insert Table 7 about here]

To further examine whether the valuation effects of climate patenting depend on firms' underlying carbon exposure, we introduce interaction terms between climate patenting activity and Scope 1 carbon emissions in Table 8. This specification allows us to assess whether the market response to climate-related innovation varies systematically with firms' operational emissions, holding constant firm-specific and time-specific factors.

The results reveal a clear and economically meaningful pattern. In Column (1), the interaction between climate patenting and absolute Scope 1 emissions is positive and statistically significant, with an estimated coefficient of 0.004 at the 1% level. Column (2) reports a similar result when carbon exposure is measured using emissions intensity, where the interaction coefficient equals 0.003 and is significant at the 10% level. These estimates indicate that the marginal effect of climate patenting on firm valuation becomes more favourable as firms' operational carbon exposure increases.

Importantly, these interaction effects should be interpreted in the context of the negative main effect of climate patenting documented earlier. On average, climate patenting is associated with lower Tobin's Q, reflecting the near- and medium-term costs and uncertainty surrounding climate-related innovation. The positive interaction terms do not overturn this

finding. Rather, they indicate that the valuation discount associated with climate patenting is attenuated for firms with higher Scope 1 emissions. In other words, while climate patenting is generally priced unfavourably by the market, this negative valuation effect is systematically weaker among firms that are more carbon intensive. This pattern is consistent with a differential interpretation of climate patenting across firms. For firms with high operational emissions, climate-related innovation may be perceived as more directly linked to binding constraints, such as regulatory pressure, transition risk, or future compliance costs. In such settings, climate patents can be interpreted as investments that mitigate exposure to carbon-related risks or reduce the likelihood of costly future adjustments. As a result, investors may view climate patenting by high-emission firms as more relevant for preserving firm value, even if the short-run financial returns remain uncertain. By contrast, for firms with relatively low carbon exposure, climate patenting may be perceived as less urgent or less tightly connected to the firm’s core operations. In these cases, climate-related innovation may appear more discretionary and may therefore accentuate concerns about upfront costs, delayed payoffs, and uncertain commercialization prospects. This difference in interpretation provides a natural explanation for why the valuation penalty associated with climate patenting is more pronounced among low-emission firms and less severe among high-emission firms, as the perceived value of climate innovation depends on how directly it addresses firms’ underlying environmental exposure.

[Insert Table 8 about here]

Table 9 provides evidence on the cash-flow channel through which climate-related innovation is priced by capital markets. Table 9 reports the dynamic relationship between climate-related patenting and firms’ future operating cash flow. The estimates show that climate patenting is followed by a statistically significant decline in operating cash flow from $t+1$ through $t+5$, indicating a persistent reduction in firms’ operating performance after climate-related innovation.

The persistence of this pattern suggests that climate patenting is associated with sustained adjustment costs rather than transitory shocks. In corporate investment theory, strategic transformation often requires the reallocation of internal resources, the adoption of new production technologies, and organisational restructuring, all of which can impose real operating costs before productivity gains are realised. Climate-related innovation fits this profile, as it typically involves upfront investment and operational changes whose economic benefits are uncertain and realised only over longer horizons.

These cash-flow dynamics have direct implications for firm valuation. In standard valuation models, firm value reflects the present value of expected future cash flows. When climate patenting is systematically associated with lower operating cash flow over the medium term, rational investors may revise downward their valuation of the firm, even if they anticipate potential long-run benefits from innovation. Under this interpretation, the negative association between climate patenting and Tobin's Q reflects the intertemporal trade-off between near-term costs and delayed payoffs, rather than a negative assessment of the underlying innovation itself. The operating cash flow results support a cash-flow channel through which climate-related patenting affects firm valuation. By documenting a persistent decline in operating performance following climate innovation, the evidence indicates that markets price climate patenting primarily through its impact on expected cash flows during the transition period.

[Insert Table 9 about here]

In addition to the cash-flow channel, changes in firm valuation may also reflect variations in discount rates. In standard asset pricing frameworks, higher perceived risk or required returns would lead to lower firm valuation even if expected cash flows remain unchanged. To examine whether the valuation effects of climate patenting operate through such a discount-rate channel, we focus on idiosyncratic volatility as a commonly used proxy for firms' risk perceptions and cost of capital.

Specifically, we examine the relationship between climate patenting activity and idiosyncratic volatility in Table 14. We find that the estimated effects are economically small and statistically insignificant. This evidence suggests that climate patenting is not systematically associated with higher perceived risk or required returns as reflected in firm-level volatility. The lack of robust effects along this dimension indicates that the observed valuation discount is unlikely to be driven primarily by changes in discount rates. Instead, when considered alongside the persistent decline in operating cash flows documented earlier, the results point to a cash-flow channel as the dominant mechanism through which climate-related innovation is priced by capital markets..

8 Real Effects of Climate Patenting

In this section, we study the relationship between climate patenting activity and firms' subsequent emissions growth, as well as their environmental and ESG-related innovation scores. We begin by analysing emissions outcomes. Scope 1 and Scope 2 emissions capture, respectively, direct emissions from firms' own operations and indirect emissions associated with purchased energy. Together, they provide a comprehensive measure of firms' operational carbon footprint. If climate patenting reflects meaningful technological or process innovation, one would expect it to be followed by a slower expansion of emissions, as new technologies are adopted and production processes become less carbon intensive. Conversely, if climate patents mainly serve reputational or disclosure purposes, there should be little systematic relationship with subsequent emissions dynamics.

The results reported in Table 10 indicate a clear negative association between climate patenting activity and future emissions growth. Firms with higher levels of climate patenting experience significantly lower growth in Scope 1 emissions over one-, two-, and three-year horizons. A similar pattern emerges for Scope 2 emissions, with coefficients that are negative and statistically significant at longer horizons. Importantly, the magnitude of the estimated

effects increases with the forecast horizon, suggesting that the impact of climate patenting on emissions unfolds gradually rather than immediately. This temporal pattern is consistent with the nature of technological and process innovation, which typically requires time to be implemented and scaled within firms' operations.

[Insert Table 10 about here]

We next turn to broader measures of environmental and innovation performance. While emissions outcomes provide a direct indicator of environmental impact, they capture only one dimension of firms' responses to climate challenges. Environmental pillar scores and ESG innovation scores aggregate information on a wider set of activities, including environmental management practices, investments in sustainable technologies, and innovation related to environmental objectives. These measures therefore offer a complementary perspective on whether climate patenting is associated with broader organisational and strategic changes.

Table 11 reports a positive and statistically significant relationship between climate patenting activity and both environmental scores and ESG innovation scores. Firms with greater climate patenting activity tend to receive higher environmental evaluations and are rated more favourably in terms of ESG-related innovation. The consistency of these results across specifications suggests that climate patenting is closely linked to firms' environmental strategies and innovative capacity, as perceived by external stakeholders. These findings reinforce the interpretation that climate patenting reflects substantive engagement with environmental challenges rather than purely symbolic actions.

Taken together, the emissions and ESG score results provide strong evidence of real effects associated with climate patenting. Firms that engage more actively in climate-related innovation exhibit slower growth in their operational emissions and are recognised as more environmentally oriented and innovative by third-party evaluators. This combination of outcomes suggests that climate patenting plays a meaningful role in shaping firms' environmental performance and innovation trajectories. In this sense, climate patents appear to contribute to tangible progress towards environmental objectives.

[Insert Table 11 about here]

9 Conclusion

This paper develops a new measure of firm-level climate innovation by applying a large language model to the text of patent abstracts and the brief description of the invention. The method allows us to classify climate-related innovation based on the meaning of the patent narrative rather than fixed labels, and it offers a richer view of how firms develop technologies that help address climate challenges. We compare the LLM measure with the widely used Y02 classification and show that, while the two measures are related, the overlap is far from complete. This comparison motivates an examination of whether the LLM-based measure captures economically relevant information beyond existing classifications. We therefore test whether the measure is associated with indicators of patent value. Patents that the LLM model identifies as climate-related have higher market-based value and receive more forward citations. These results hold in settings where the Y02 measure shows weaker or no explanatory power, which suggests that the LLM classification is more closely tied to the economic importance of climate innovation. Next, we study how the measure relates to firm behavior. Firms that face more environmental inspections produce more climate-related patents in later years when innovation is measured with the LLM method. The Y02 measure shows little response. A similar pattern appears when we use firm-level climate risk exposure: firms with greater exposure generate more climate innovations in the future, but again only when innovation is measured with the LLM approach. These findings show that the LLM measure has predictive content for how firms adjust their innovation strategies after facing climate-related shocks.

Using this measure, we examine how climate-related innovation is reflected in firm valuation and real corporate outcomes. On average, climate patenting activity is associated with lower Tobin's Q , indicating that equity markets tend to price climate innovation as a near-

term cost linked to adjustment needs and uncertainty rather than as an immediate source of value. This valuation effect is not uniform. It is weaker for firms with higher operational carbon exposure and is concentrated in mitigation-oriented technologies, while adaptation-related innovations show no systematic relation with firm value. These patterns indicate that markets differentiate climate innovation based on its economic role within the firm, rather than applying a uniform valuation response. An analysis of underlying mechanisms suggests that valuation effects operate primarily through expected cash flows rather than changes in required returns. Climate-related innovation is associated with adjustment costs, delayed payoffs, and uncertainty in operating performance, particularly for mitigation technologies that involve substantial investment and long deployment horizons. There is limited evidence that the observed valuation effects are driven by changes in discount rates, consistent with a cash-flow-based interpretation of market pricing. Finally, we examine the real effects of climate patenting. Firms that engage more intensively in climate-related innovation subsequently reduce their carbon emissions and improve environmental and ESG-related performance. This evidence indicates that the innovation captured by the LLM-based measure reflects substantive technological and organisational change rather than symbolic or cosmetic activity.

This paper applies LLMs to both patent abstracts and the brief description of the invention to construct text-based variables. Rather than relying on predefined keyword lists or rigid classification rules, the models extract semantic information from contextual meaning. This approach is particularly well suited to patent texts, where technological language is often abstract and varies across fields. By jointly drawing on the high-level overview provided by abstracts and the more detailed technical narrative contained in the brief description of the invention, the resulting variables more effectively capture the technological content and economic relevance embedded in patents.

References

- Acikalin, U. U., Caskurlu, T., Hoberg, G., & Phillips, G. M. (2025). @@@Intellectual Property Protection Lost and Competition: An Examination Using Large Language Models.
- Aghion, P., Bénabou, R., Martin, R., & Roulet, A. (2023). ENVIRONMENTAL PREFERENCES AND TECHNOLOGICAL CHOICES: IS MARKET COMPETITION CLEAN OR DIRTY?
- 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.
- Asimakopoulos, P., Asimakopoulos, S., & Li, X. (2023). The role of environmental, social, and governance rating on corporate debt structure. *Journal of Corporate Finance*, 83, 102488.
- Balakrishnan, K., Copat, R., De La Parra, D., & Ramesh, K. (2023). Racial Diversity Exposure and Firm Responses Following the Murder of George Floyd. *Journal of Accounting Research*, 61(3), 737–804.
- Block, J., Lambrecht, D., Willeke, T., Cucculelli, M., & Meloni, D. (2025). (Research Policy) Green patents and green trademarks as indicators of green innovation. *Research Policy*, 54(1), 105138.
- Bolton, P., Kacperczyk, M. T., & Wiedemann, M. (2022). The CO2 Question: Technical Progress and the Climate Crisis. *SSRN Electronic Journal*.
- Breitung, C., & Müller, S. (2025). Global Business Networks. *Journal of Financial Economics*, 166, 104007.
- Briere, M., Nuriyev, M., & Pouget, S. (2025). Market Valuation of Climate Patents: What are the Most Valuable Innovations ?
- Briscoe-Tran, H. (2025). Do Employees Have Useful Information.

- Brown, J. R., Martinsson, G., & Thomann, C. (2022). Can Environmental Policy Encourage Technical Change? Emissions Taxes and R&D Investment in Polluting Firms (D. Denis, Ed.). *The Review of Financial Studies*, 35(10), 4518–4560.
- Cao, Y., Chen, L., Tucker, J. W., & Wan, C. (2025). Can generative AI help identify peer firms? *Review of Accounting Studies*.
- Chen, F., Chen, Z., & Zhang, X. (2024). Belated stock returns for green innovation under carbon emissions trading market. *Journal of Corporate Finance*, 85, 102558.
- Chen, M. A., & Wang, J. (2024). @@@Displacement or Augmentation? The Effects of AI Innovation on Workforce Dynamics and Firm Value.
- Cheng, Q., Lin, A.-P., & Yang, M. (2024). Green innovation and firms' financial and environmental performance: The roles of pollution prevention versus control. *Journal of Accounting and Economics*, 101706.
- Cheng, Q., Lin, A.-P., & Yang, M. (2025). Green innovation and firms' financial and environmental performance: The roles of pollution prevention versus control. *Journal of Accounting and Economics*, 79(1), 101706.
- Cohen, L., Gurun, U. G., & Nguyen, Q. H. (2020). The ESG-Innovation Disconnect: Evidence from Green Patenting.
- Dai, R., Duan, R., & Ng, L. (2025). Innovating Green: Competition Meets Regulation(MS). *Management Science*, mnscl.2023.00600.
- Dai, R., Liang, H., & Ng, L. (2021). Socially responsible corporate customers. *Journal of Financial Economics*, 142(2), 598–626.
- Dechezleprêtre, A., Muckley, C. B., & Neelakantan, P. (2021). Is firm-level clean or dirty innovation valued more? *The European Journal of Finance*, 27(1-2), 31–61.

- Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3), 693–714.
- Fetzer, T., Lambert, P. J., Feld, B., & Garg, P. (2024). AI-Generated Production Networks: Measurement and Applications to Global Trade.
- Fusillo, F. (2023). (Research Policy) Green Technologies and diversity in the knowledge search and output phases: Evidence from European Patents. *Research Policy*, 52(4), 104727.
- Gianluca Orsatti, Quatraro, F., & Pezzoni, M. (2020). (Research Policy) The antecedents of green technologies: The role of team-level recombinant capabilities. *Research Policy*, 49(3), 103919.
- Green, J., Hand, J. R. M., & Sikochi, A. (2024). The asymmetric mispricing information in analysts' target prices. *Review of Accounting Studies*, 29(1), 889–915.
- Harford, J., He, Q., & Qiu, B. (2023). Firm-Level Labor-Shortage Exposure. *SSRN Electronic Journal*.
- Harford, J., He, Q., & Qiu, B. (2025). The Value of Corporate Patent Utilization.
- Hassan, T. A., Hollander, S., Van Lent, L., & Tahoun, A. (2019). Firm-Level Political Risk: Measurement and Effects*. *The Quarterly Journal of Economics*, 134(4), 2135–2202.
- He, Q., & Qiu, B. (2025). Environmental enforcement actions and corporate green innovation. *Journal of Corporate Finance*, 91, 102711.
- Hege, U., Li, K., & Zhang, Y. (2023). Climate Innovation and Carbon Emissions: Evidence from Supply Chain Networks. *SSRN Electronic Journal*.
- Hege, U., Pouget, S., & Zhang, Y. (2024). Climate Patents and Financial Markets.

- Hoberg, G., & Phillips, G. (2016). Text-Based Network Industries and Endogenous Product Differentiation.
- Hu, J., Li, S., & Shevlin, T. (2023). How does the market for corporate control impact tax avoidance? Evidence from international M&A laws. *Review of Accounting Studies*, 28(1), 340–383.
- Jha, M., Liu, H., & Manela, A. (2025). Does Finance Benefit Society? A Language Embedding Approach. *The Review of Financial Studies*, 0(0).
- Kelly, B., Papanikolaou, D., Seru, A., & Taddy, M. (2021). Measuring Technological Innovation over the Long Run. *American Economic Review: Insights*, 3(3), 303–320.
- Kim, I., Pantzalis, C., & Zhang, Z. (2021). Multinationality and the value of green innovation (JCF. *Journal of Corporate Finance*, 69, 101996.
- Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth*. *The Quarterly Journal of Economics*, 132(2), 665–712.
- 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.
- Kuang, H., & Liang, B. (2024). Do Investors Care about Climate Innovations?
- Leippold, M., & Yu, T. (2025). Firm-Level Green Innovation Beyond Patents RFS).
- Lerner, J., & Seru, A. (2022). The Use and Misuse of Patent Data: Issues for Finance and Beyond (A. Karolyi, Ed.). *The Review of Financial Studies*, 35(6), 2667–2704.
- Li, W., Neupane, S., & Tan, K. J. K. (2022). Toxic Emissions and Corporate Green Innovation.

- Lin, J., Cao, X., Dong, X., & An, Y. (2024). Environmental regulations, supply chain relationships, and green technological innovation. *Journal of Corporate Finance*, 88, 102645.
- Lin, Y., Shen, R., Wang, J., & Julia Yu, Y. (2024). Global Evolution of Environmental and Social Disclosure in Annual Reports. *Journal of Accounting Research*, 62(5), 1941–1988.
- Loughran, T., & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01625.x>]. *The Journal of Finance*, 66(1), 35–65.
- Mo, H., & Ouyang, S. (2025). (Generative) AI in Financial Economics.
- Nagar, V., & Schoenfeld, J. (2024). Measuring weather exposure with annual reports. *Review of Accounting Studies*, 29(1), 1–32.
- Reza, S. W. (2024). The Value of Green Innovation.
- Santarlaschi, L., Rungi, A., & Zinilli, A. (2025, July). @@@Seeing Through Green: Text-Based Classification and the Firm’s Returns from Green Patents [arXiv:2507.02287 [econ]].
- Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023a). Firm-Level Climate Change Exposure [_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.13219>]. *The Journal of Finance*, 78(3), 1449–1498.
- Sautner, Z., Van Lent, L., Vilkov, G., & Zhang, R. (2023b). Firm-Level Climate Change Exposure. *The Journal of Finance*, 78(3), 1449–1498.
- Sautner, Z., van Lent, L., Vilkov, G., & Zhang, R. (2023). Climate Value and Values Discovery.
- Schiller, C. (2017). Global Supply-Chain Networks and Corporate Social Responsibility. *SSRN Electronic Journal*.

Wang, M., Wurgler, J., & Zhang, H. (2026). Policy uncertainty reduces green innovation. *Journal of Financial Economics*, 175, 104189.

Zheng, Y. (2025). **Can Investors Learn from Patent Documents? — Evidence from Textual Analysis.** *SSRN Electronic Journal*.

Notes: This figure shows the cumulative number of climate-related patents across U.S. states over the period 2000–2023. Panel A reports climate patents identified using the Y02 green patent classification, while Panel B reports climate patents identified using the LLM-based classifier.

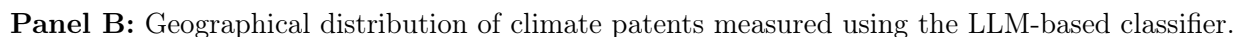


Figure 2: Industry Distribution of Green Patent Proportions

Notes: This figure illustrates the distribution of green patenting activities across major U.S. industries. The chart compares industry-level green patent shares based on two classification methods: the Y02 green patent classification and an LLM-derived classifier. The top industries by green patent proportion include: (1) Industrial Inorganic Chemicals, (2) Engines & Turbines, (3) Petroleum Refining, (4) Industrial Organic Chemicals.

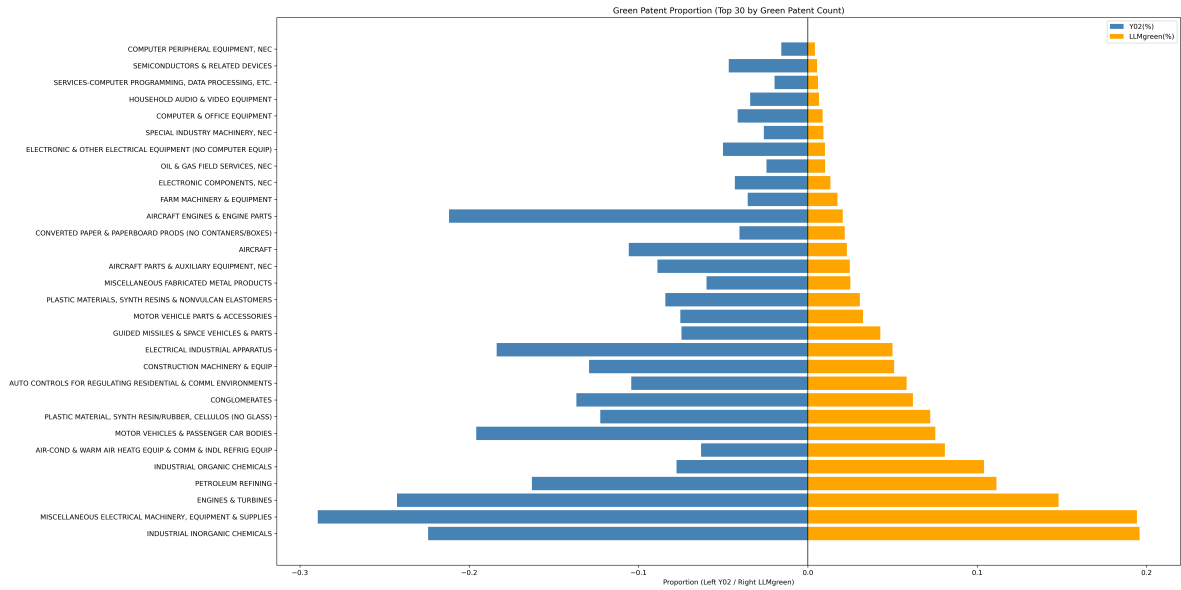


Table 1: Summary Statistics

This table reports summary statistics for the main variables used in the analysis, including climate innovation measures, firm characteristics, and financial performance indicators. The sample spans all firm-year observations used in the baseline regressions.

VARIABLE	N	Mean	SD	p5	p50	p95
Climate_y02	26,830	2.162	8.808	0	0	9
Climate_LLM	26,830	0.559	2.510	0	0	3
Logasset	26,818	6.821	2.321	3.212	6.657	10.900
Tobin_q	25,109	2.561	2.052	0.890	1.858	6.738
Tangibility	26,816	0.180	0.163	0.014	0.130	0.545
Lev	26,674	0.192	0.197	0	0.154	0.557
Capexta	26,818	0.038	0.037	0.003	0.027	0.115
Rdex	26,818	0.113	0.157	0	0.058	0.435
Ocf	26,790	0.004	0.276	-0.583	0.095	0.257
Roa	24,339	-0.022	0.280	-0.618	0.061	0.248

Table 2: Validation of Climate Innovation Measures

This table validates the relationship between the LLM-based climate patent measure and the traditional Y02 climate patent classification. All regressions include year and patent family fixed effects. Standard errors are clustered at the patent level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	(1) LLMgreen	(2) LLMgreen
Y02	0.140*** (11.783)	0.130*** (11.211)
Constant	0.021*** (2.647)	0.018 (1.591)
Observations	1,569,964	1,569,321
R-squared	0.142	0.174
Control variables	YES	YES
Industry FE	YES	NO
Firm FE	NO	YES
Patent family FE	YES	YES
Year FE	YES	YES

Table 3: Patent Value and Patent Citations

This table reports regression results examining (i) the association between climate patent indicators and patent value (Panel A), and (ii) whether green patent indicators predict forward patent citations (Panel B). All regressions include year and patent family fixed effects. Standard errors are clustered at the firm level (Panel B) or patent level (Panel A). The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Patent Value at the Patent Level				
VARIABLES	(1) Xi real	(2) Xi real	(3) Xi real	(4) Xi real
LLMgreen	0.120** (2.107)		0.024* (1.739)	
Y02		0.086* (1.911)		0.033 (1.427)
Constant	-1.950*** (-3.165)	-1.949*** (-3.164)	-6.294*** (-6.633)	-6.293*** (-6.631)
Observations	1,574,018	1,574,018	1,573,369	1,573,369
R-squared	0.659	0.659	0.929	0.929
Control variables	YES	YES	YES	YES
Industry FE	YES	YES	NO	NO
Firm FE	NO	NO	YES	YES
Patent family FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Panel B: Patent Citations				
VARIABLES	(1) Cites	(2) Cites	(3) Cites	(4) Cites
LLMgreen	1.028** (2.242)		0.898*** (2.921)	
Y02		0.278 (0.413)		0.768 (1.570)
Constant	16.341*** (3.189)	16.361*** (3.201)	71.513*** (4.650)	71.546*** (4.652)
Observations	1,574,018	1,574,018	1,573,369	1,573,369
R-squared	0.067	0.067	0.178	0.178
Control variables	YES	YES	YES	YES
Industry FE	YES	YES	NO	NO
Firm FE	NO	NO	YES	YES
Patent family FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 4: Stock Market Reaction to Major U.S. Climate Policy Events

This table examines firms' cumulative abnormal returns (CARs) around major U.S. political events related to climate policy. Panel A reports stock market reactions to Trump's election victory, and Panel B reports reactions to the Inflation Reduction Act (IRA) announcement. CARs are calculated over $[0, 1]$, $[0, 5]$, and $[0, 10]$ event windows. The main independent variable, *climate_LLM*, captures firm-level climate exposure, with *climate_y02* included as an alternative measure. All specifications control for firm characteristics and include industry and year fixed effects. Robust *t*-statistics are reported in parentheses. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Stock Market Reaction to the IRA Announcement						
VARIABLES	(1) CAR[0,1]	(2) CAR[0,1]	(3) CAR[0,5]	(4) CAR[0,5]	(5) CAR[0,10]	(6) CAR[0,10]
Climate_LLM	0.017*** (3.205)		0.021*** (2.629)		0.025** (2.534)	
Climate_y02		0.004 (0.915)		0.002 (0.397)		-0.007 (-0.944)
Constant	-0.005 (-0.881)	-0.006 (-1.195)	0.043*** (2.599)	0.041** (2.461)	0.053*** (2.788)	0.050** (2.579)
Observations	2,944	2,944	2,944	2,944	2,944	2,944
R-squared	0.218	0.215	0.164	0.163	0.149	0.148
Control variables	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Panel B: Stock Market Reaction to Trump's Election Victory						
VARIABLES	(1) CAR[0,1]	(2) CAR[0,1]	(3) CAR[0,5]	(4) CAR[0,5]	(5) CAR[0,10]	(6) CAR[0,10]
Climate_LLM	-0.009*** (-3.158)		-0.007** (-2.347)		-0.012*** (-2.873)	
Climate_y02		-0.005*** (-3.896)		-0.004 (-1.637)		-0.005 (-1.470)
Constant	-0.003 (-0.473)	-0.004 (-0.572)	0.053*** (4.824)	0.053*** (4.804)	0.077*** (6.282)	0.077*** (6.072)
Observations	2,362	2,362	2,362	2,362	2,363	2,363
R-squared	0.249	0.248	0.288	0.288	0.240	0.239
Control variables	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Table 5: Regulatory and Climate Exposure Effects on Climate Innovation

This table examines how environmental regulatory pressure and climate-related exposure affect firms' subsequent climate innovation outcomes. Panel A studies the effect of environmental enforcement intensity on future climate innovation. Panel B examines the role of climate change exposure and regulatory climate exposure. The dependent variables are LLM-based and Y02-based climate innovation factors measured at $t+1$ and $t+2$. All regressions include the fixed effects and control variables specified in each panel. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Environmental Regulatory Pressure and Climate Innovation Responses				
VARIABLES	(1) Climate_LLM $_{t+1}$	(2) Climate_y02 $_{t+1}$	(3) Climate_LLM $_{t+2}$	(4) Climate_y02 $_{t+2}$
Lninspect_tot	0.025* (1.675)	0.092 (1.551)	0.033* (1.880)	0.107 (1.640)
Constant	-52.700 (-1.432)	-201.297 (-1.476)	-64.548 (-1.329)	-202.868 (-1.297)
Observations	259,386	259,386	242,219	242,219
R-squared	0.770	0.753	0.789	0.770
Control variables	YES	YES	YES	YES
Firm FE	YES	YES	YES	YES
County FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Panel B: Climate Exposure and Future Green Innovation Output				
VARIABLES	(1) Climate_LLM $_{t+1}$	(2) Climate_y02 $_{t+1}$	(3) Climate_LLM $_{t+1}$	(4) Climate_y02 $_{t+1}$
Ln_cc_expo	0.011** (2.318)	0.007 (1.351)		
Ln_rg_expo			0.061*** (2.982)	0.042** (1.994)
Constant	-0.286*** (-8.877)	-0.658*** (-11.592)	-0.277*** (-8.604)	-0.652*** (-11.421)
Observations	54,286	54,286	54,286	54,286
R-squared	0.212	0.260	0.210	0.260
Control variables	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

Table 6: Supplier's Climate Patents and Customer's CO2 Emission Changes

This table reports regression results examining whether suppliers' climate innovation affects downstream firms' future Scope 1 carbon emissions. Panel A uses the LLM-based green patent measure, and Panel B uses the Y02-based green patent classification. The dependent variable is the log of Scope 1 emissions from year $t + 1$ to $t + 5$. All regressions include industry and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Supplier LLM-based Green Patents					
VARIABLES	(1) Emissions Scope 1 t+1	(2) Emissions Scope 1 t+2	(3) Emissions Scope 1 t+3	(4) Emissions Scope 1 t+4	(5) Emissions Scope 1 t+5
w_cpr_LLM	-0.082 (-0.601)	-0.304 (-1.513)	-1.161*** (-3.911)	-1.064*** (-2.964)	-0.915** (-2.281)
Supplier Patents	0.001 (0.651)	0.008** (1.996)	0.014** (2.431)	0.015** (2.015)	0.015 (1.580)
Constant	-0.001 (-0.044)	-0.048 (-1.198)	-0.054 (-0.889)	-0.078 (-1.016)	-0.110 (-1.164)
Observations	12,696	12,525	11,348	10,098	8,942
R-squared	0.043	0.060	0.079	0.088	0.100
Control variables	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Panel B: Supplier Y02-based Green Patents					
VARIABLES	(1) Emissions Scope 1 t+1	(2) Emissions Scope 1 t+2	(3) Emissions Scope 1 t+3	(4) Emissions Scope 1 t+4	(5) Emissions Scope 1 t+5
w_cpr_y02	-0.079 (-0.982)	-0.010 (-0.088)	0.051 (0.344)	0.152 (0.917)	0.206 (1.029)
Supplier Patents	0.002 (0.730)	0.008* (1.932)	0.013** (2.241)	0.014* (1.831)	0.014 (1.413)
Constant	0.000 (0.007)	-0.051 (-1.276)	-0.069 (-1.138)	-0.095 (-1.239)	-0.129 (-1.359)
Observations	12,696	12,525	11,348	10,098	8,942
R-squared	0.043	0.060	0.077	0.087	0.100
Control variables	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Year FE	YES	YES	55 YES	YES	YES

Table 7: Pricing of climate patents

This table reports the relationship between firms' climate patenting activity and future Tobin's Q. The dependent variables are Tobin's Q measured from one to five years ahead. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

VARIABLES	(1) Tobin_Q _{t+1}	(2) Tobin_Q _{t+2}	(3) Tobin_Q _{t+3}	(4) Tobin_Q _{t+4}	(5) Tobin_Q _{t+5}
Climate_LLM	-0.003** (-1.974)	-0.003** (-2.262)	-0.004*** (-2.658)	-0.004*** (-2.960)	-0.004** (-2.492)
Constant	4.993*** (12.888)	5.379*** (13.348)	4.991*** (11.692)	4.566*** (10.058)	4.309*** (8.918)
Observations	16,972	15,183	13,551	12,082	10,784
R-squared	0.655	0.656	0.663	0.677	0.679
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Table 8: Climate patents, carbon emissions, and firm valuation

This table reports the relationship between firms' climate patenting activity and firm valuation. The dependent variable is Tobin's Q. Column (1) focuses on climate patents interacted with Scope 1 absolute carbon emissions, while Column (2) considers climate patents interacted with Scope 1 carbon emission intensity. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

VARIABLES	(1) Tobin_Q	(2) Tobin_Q
Scope1AbsoluteClimate_LLM	0.004*** (3.142)	
Scope1IntensityClimate_LLM		0.003* (1.914)
ln_AbsoluteGHGScope1	-0.055 (-1.456)	
ln_IntensityGHGScope1		-0.041 (-0.998)
Climate_LLM	-0.056*** (-3.063)	-0.009* (-1.905)
Constant	7.090*** (6.635)	7.015*** (6.644)
Observations	10,443	10,443
R-squared	0.683	0.682
Firm FE	YES	YES
Year FE	YES	YES

Table 9: Effects of Climate Exposure on Future Operating Cash Flow

This table reports the relationship between climate exposure measured by an LLM-based indicator and firms' future operating cash flow. The dependent variable is operating cash flow measured from $t+1$ to $t+5$. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

VARIABLES	(1) OCF _{$t+1$}	(2) OCF _{$t+2$}	(3) OCF _{$t+3$}	(4) OCF _{$t+4$}	(5) OCF _{$t+5$}
Climate_LLM	-0.00018*** (-3.234)	-0.00025*** (-3.100)	-0.00027** (-2.486)	-0.00036*** (-2.641)	-0.00038*** (-2.589)
Constant	-0.01618 (-0.534)	0.03141 (0.881)	0.05855 (1.532)	0.07084* (1.783)	0.09994** (2.277)
Observations	16,991	15,254	13,622	12,134	10,803
R-squared	0.837	0.819	0.813	0.809	0.797
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Table 10: Effects of Climate Exposure on Scope 1 and Scope 2 Emissions Growth

This table examines the relationship between climate exposure measured by an LLM-based indicator and firms' future emissions growth. The dependent variables are Scope 1 and Scope 2 total emissions growth measured at different horizons. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

	Scope 1 Emissions Growth			Scope 2 Emissions Growth		
VARIABLES	$t+1$	$t+2$	$t+3$	$t+1$	$t+2$	$t+3$
Climate_LLM	-0.001** (-2.077)	-0.002* (-1.921)	-0.003* (-1.880)	-0.001 (-1.531)	-0.002** (-2.365)	-0.003* (-1.765)
Constant	0.830*** (4.298)	2.027*** (5.908)	2.808*** (5.620)	0.653*** (3.615)	1.948*** (7.501)	2.752*** (7.991)
Observations	10,092	9,513	8,610	10,092	9,513	8,610
R-squared	0.198	0.297	0.364	0.242	0.335	0.404
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Table 11: Effects of Climate Exposure on Environmental and ESG Innovation Scores

This table reports the relationship between climate exposure measured by an LLM-based indicator and firms' environmental pillar scores and ESG innovation scores. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

VARIABLES	(1) Environment Score	(2) Environment Score	(3) ESG Innovation Score	(4) ESG Innovation Score
Climate_LLM	0.002** (2.520)	0.001* (1.725)	0.003*** (2.873)	0.002*** (2.626)
Constant	-0.686*** (-20.196)	-0.105 (-1.100)	-0.354*** (-9.081)	0.288*** (2.678)
Observations	7,298	7,105	7,298	7,105
R-squared	0.659	0.862	0.503	0.762
Firm FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES

10 Appendix

Part A

This appendix presents two types of mismatches between the Y02 scheme and the LLM-based classification. Some patents are labeled as climate-related under Y02 but, when evaluated using the LLM, do not contain substantive climate content. Other patents show the opposite pattern: the LLM identifies them as climate-related even though they fall outside the Y02 categories.

Example 1:

Patent Number: US 11,449,349

CPC Climate Classification (Y02): Y02D 30/50

Patent Abstract:

An automatic hold switch is disclosed. The automatic hold switch provides a means to automatically turn on and off a hold feature. When the hold feature is on, one or more input devices of a portable electronic device are disabled or prevented from providing input signals. When the hold feature is off, one or more input devices of a portable electronic device are enabled or allowed to provide input signals. Because the user no longer has to manually control the hold feature, the number of actions that the user needs to take is reduced. In one example, the automatic hold switch is embodied with light sensors that detect when the device is in a dark environment and when the device is in a light environment. A dark environment indicates to the portable electronic device that the user wishes not to input, and therefore the hold feature is turned on. A lighted environment indicates to the portable electronic device that the user wishes to input, and therefore the hold feature is turned off.

LLM-Based Assessment:

No, this is not a climate-related patent because the invention merely concerns an automatic

hold switch for portable electronic devices using light sensors to manage input locking, which is designed for user interface automation and convenience rather than to reduce greenhouse gas emissions, improve energy efficiency, support renewable energy, or achieve any other climate-mitigation objective.

Example 2:

Patent Number: US 9,977,963

CPC Climate Classification: G01P5/001; G06K9/0063; G01N21/01; G05D1/0011
B64C-2201/021; G01C11/02; G01J5/0014

Patent Abstract:

A system and method for tracking the growth of large-area wildland fires. The technique includes monitoring wind conditions in and around a wildfire using near-surface-sited sensors deployed by a high-altitude long endurance (HALE) unmanned aerial vehicle (UAV), such as the Global Hawk. The deployed sensors measure a localized wind vector at multiple locations within and surrounding the wildfire, and transmit the wind data back to the UAV for relaying to a command center operated by fire management authorities, where the wind data are used to assist decision-making, including as input into wildfire growth prediction models. The sensors may provide additional data such as local altitude/elevation, pressure, temperature and relative humidity. The UAV may also carry sensors that provide additional data to be used by the fire management authorities or the wildfire growth prediction models, such as infrared images defining actively flaming areas, and visual images which indicate vegetation type and density.

LLM-Based Assessment:

Yes, this is a climate-related patent, because the invention focuses on monitoring, predicting, and managing large-scale wildland fires, which are directly related to climate-change adaptation. It uses sensor networks, UAVs, and environmental data to support wildfire prediction

and emergency response, key components of climate resilience technologies.

Part B: Institutional Background and Text Selection

This appendix provides institutional background on U.S. patent disclosures and explains the rationale for the selection of patent text and the construction of text-based variables using large language models (LLMs).

Under the U.S. patent system, a patent granted by the United States Patent and Trademark Office (USPTO) must satisfy the requirements of novelty and usefulness, and the invention must be disclosed in a structured and standardised textual form during the examination process. This institutional design ensures that patent documents contain information closely related to firms' technological activities and the potential economic value of their innovations. The disclosed content is reviewed by specialised patent examiners, which enhances its accuracy and credibility. Firms' decisions to patent an invention and to disclose its technical details reflect a deliberate trade-off between the benefits of legal protection and the costs of public disclosure. As a result, granted patent texts provide a reliable and economically meaningful source of data for analysing firms' internal research and development outcomes.

Given the large volume of patent data, the study adopts a two-stage text selection strategy. In the initial screening stage, patent abstracts are used to identify whether a patent is climate-related. Abstracts provide a concise summary of the invention and allow for efficient processing when handling a large number of documents. This step is designed to reduce the sample size while retaining patents that are potentially relevant to climate-related innovation.

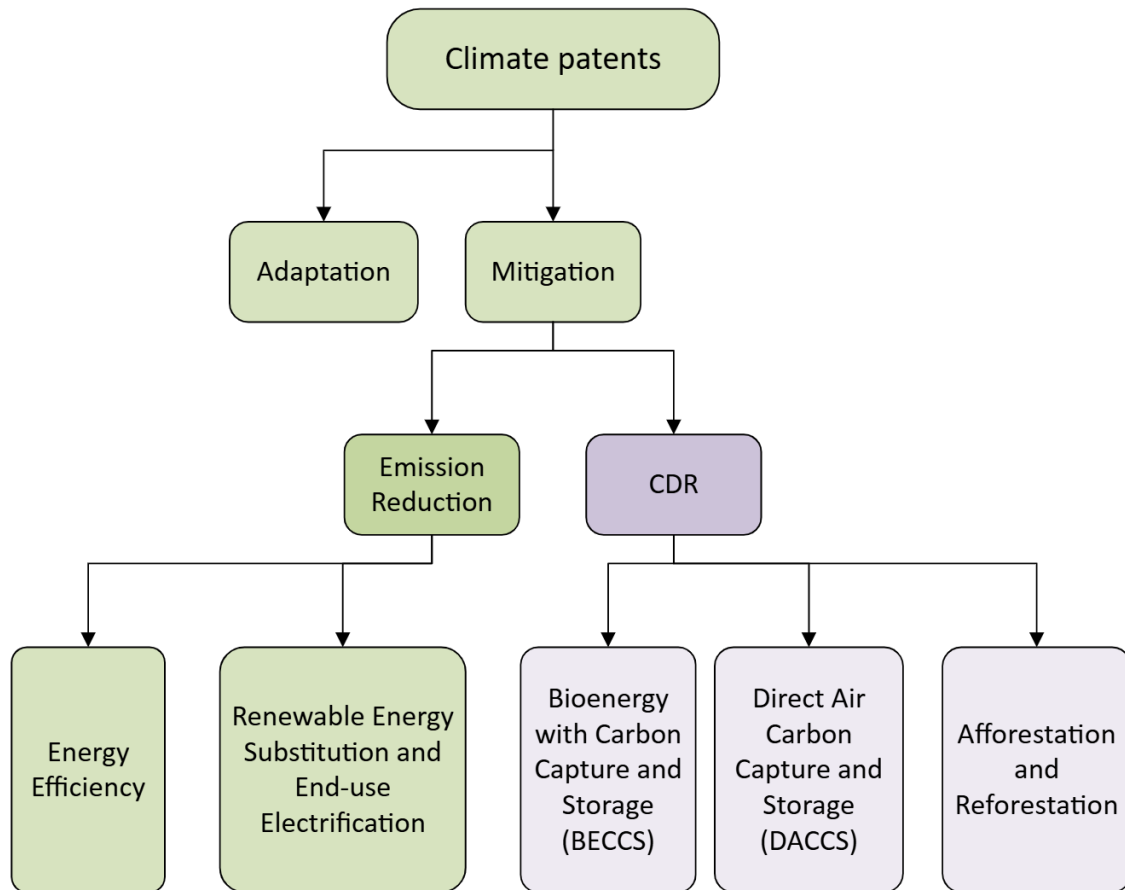
In the subsequent stage, the analysis focuses on the brief description of the invention section of patent documents. This section summarises the technological background, the objective of the invention, its main innovative features, and its intended applications. Compared with the Detailed Description section, which is highly technical and extensive, the brief description presents the core elements of the invention in a more concise and interpretable form. It therefore provides a more suitable textual basis for refining the classification of

climate-related patents and for distinguishing among different types of climate technologies.

This study applies LLMs to both patent abstracts and the brief description of the invention to construct text-based variables. Rather than relying on predefined keyword lists or rigid classification rules, the models extract semantic information from contextual meaning. This approach is particularly well suited to patent texts, where technological language is often abstract and varies across fields. By jointly drawing on the high-level overview provided by abstracts and the more detailed technical narrative contained in the brief description of the invention, the resulting variables more effectively capture the technological content and economic relevance embedded in patents.

Figure 3: Climate-related patent classification framework

Notes: This figure illustrates an IPCC-consistent hierarchical classification framework for climate-related patents. At the first level, patents are classified into Adaptation and Mitigation according to their primary climate function. Mitigation patents are further divided into Emission Reduction and Carbon Dioxide Removal (CDR). Emission Reduction technologies include Energy Efficiency and Renewable Energy Substitution and End-use Electrification, while CDR technologies comprise Bioenergy with Carbon Capture and Storage (BECCS), Direct Air Carbon Capture and Storage (DACCS), and Afforestation and Reforestation.



Part C: Prompt for hierarchical classification of climate-related patents

This section reports the instructions provided to the generative AI model for classifying climate-related patents under a hierarchical classification scheme. The model must follow all definitions and rules exactly and assign labels solely based on the primary climate function of each patent.

You are given the text of a climate-related patent. Your task is to classify the patent using the hierarchical classification scheme below. Each patent must be assigned a single label at each applicable level. Do not provide explanations. Do not output any text other than numeric labels (0 or 1).

Level 1: Adaptation vs. Mitigation

Definitions follow the official IPCC terminology.

Adaptation refers to actions that anticipate the adverse effects of climate change and aim to prevent or minimise associated damage, or to take advantage of potential opportunities.

Mitigation refers to actions that make climate change less severe by preventing or reducing the emission of greenhouse gases into the atmosphere.

Task (mandatory)

Return two binary variables: Adaptation Mitigation

Rules:

- If the patent's primary purpose corresponds to adaptation: Adaptation = 1, Mitigation = 0.
- If the patent's primary purpose corresponds to mitigation: Adaptation = 0, Mitigation = 1.
- Exactly one of these two variables must equal 1.

Level 2: Emission Reduction vs. Carbon Dioxide Removal

This level applies only if Mitigation = 1.

Carbon Dioxide Removal (CDR) refers to anthropogenic activities that remove CO₂ from the atmosphere and durably store it in geological, terrestrial, ocean reservoirs, or in products.

Emission reduction refers to mitigation technologies that reduce greenhouse gas emissions at the source and do not remove CO₂ already present in the atmosphere. Any mitigation technology that does not meet the IPCC definition of CDR must be classified as emission reduction.

Task (mandatory if Mitigation = 1)

Return two binary variables: Emission_reduction CDR

Rules:

- If the patent removes atmospheric CO₂ and provides durable storage: Emission_reduction = 0, CDR = 1.
- If the patent mitigates climate change without removing atmospheric CO₂: Emission_reduction = 1, CDR = 0.
- Exactly one of these two variables must equal 1.
- If Mitigation = 0: Emission_reduction = 0 and CDR = 0.

Level 3A: Emission Reduction Categories

This level applies only if Emission_reduction = 1.

Energy efficiency includes technologies that reduce energy consumption while delivering the same level of energy service.

Renewable substitution or electrification includes technologies that replace fossil fuels with renewable or low-carbon energy sources, or electrify end-use energy demand.

Task (mandatory if Emission_reduction = 1)

Return two binary variables: Energy_efficiency Renewable_substitution_electrification

Rules:

- Exactly one of these two variables must equal 1.
- If Emission_reduction = 0: Energy_efficiency = 0 and Renewable_substitution_electrification = 0.

Level 3B: Carbon Dioxide Removal Approaches

This level applies only if CDR = 1.

Recognised approaches include: BECCS (bioenergy with carbon capture and storage), DACCS (direct air carbon capture and storage), Afforestation or reforestation.

Task (mandatory if CDR = 1)

Return three binary variables: BECCS DACCS Afforestation_Reforestation

Rules:

- Exactly one of these three variables must equal 1.
- If CDR = 0: BECCS = 0, DACCS = 0, Afforestation_Reforestation = 0.

Output Format

Return the result only in the following JSON format:

```
{  "Adaptation": 0,
  "Mitigation": 0,
  "Emission_reduction": 0,
  "CDR": 0,
```

“Energy_efficiency”: 0,
“Renewable_substitution_electrification”: 0,
“BECCS”: 0,
“DACCS”: 0,
“Afforestation_Reforestation”: 0}

Table 12: Classification of Green Innovation in the Literature

Paper	Data	Green Innovation Categories
Leippold & Yu (2025), <i>Review of Finance</i>	Earnings call transcripts; green patent abstracts (used for model training)	Green innovation is classified along the innovation lifecycle into green invention (creation of new green technologies or strategies) and green adoption (implementation and integration of existing green technologies into business operations).
Cheng, Lin & Yang (2025), <i>Journal of Accounting and Economics</i>	USPTO patent data (KPSS); CPC-based green patent classification	Green patents are classified by technological function into pollution prevention innovation (source-reducing, process- or product-based technologies) and pollution control innovation (end-of-pipe abatement and emission treatment technologies).
Bolton, Kacperczyk & Wiedemann (2023), working paper	Global patent filings (CPC/IPC); firm financial data; Trucost emissions data	Patents are classified into green innovation (carbon-free or low-carbon technologies), brown efficiency innovation (efficiency improvements in fossil fuel-based technologies), and other non-environmental patents , explicitly separating green substitution from fossil-fuel efficiency improvements.
Aghion et al. (2023)	EPO PATSTAT patent families; CPC and Y02 climate mitigation classifications	Green innovation is classified based on technological trajectory into clean innovations (non-fossil-fuel-based propulsion technologies such as electric, hydrogen, and fuel-cell vehicles), dirty innovations (internal combustion engine technologies), and grey innovations (efficiency-improving fossil-fuel technologies). Grey patents are excluded from the main analysis, focusing on clean versus dirty technological substitution.

Figure 4: Climate-related Patenting Trends: Y02 vs. LLM Classifications

Notes: This figure plots annual counts of climate-related patents based on two classification systems: (i) the Y02 classification from the EPO’s climate-change mitigation patent classes, and (ii) an LLM-derived classifier that identifies climate-related innovations from patent abstracts.

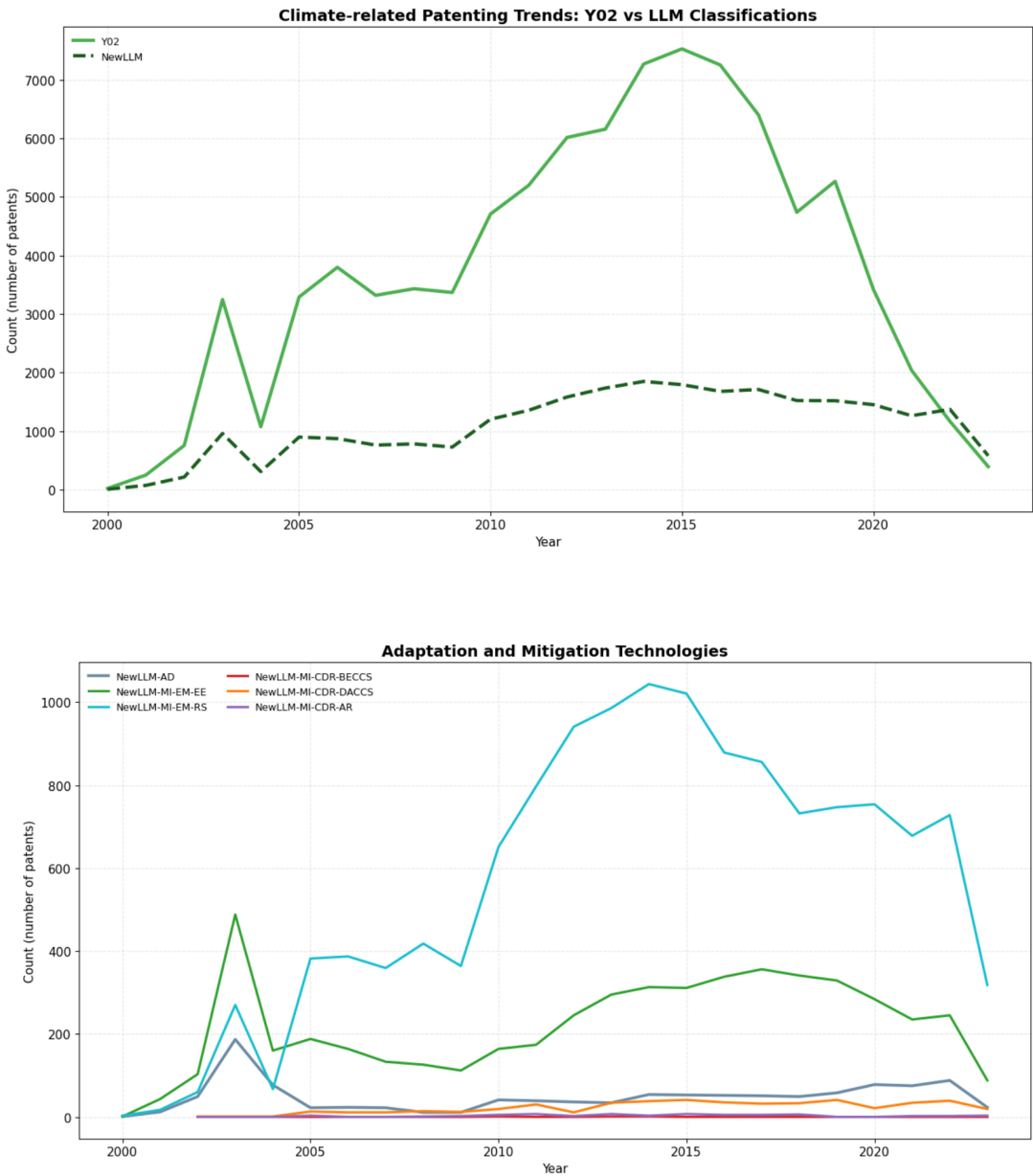


Figure 5: Composition of Climate-related Patents by Adaptation, Mitigation, and Technology Channels

Notes: This figure illustrates the distribution of climate-related patents across hierarchical climate technology categories. The top-left panel compares Adaptation and Mitigation patents based on the Y02 classification. The top-right panel further decomposes Mitigation patents into Emission Reduction and Carbon Dioxide Removal (CDR). The bottom-left panel shows the relative importance of emission reduction channels, distinguishing between energy efficiency improvements and renewable substitution. The bottom-right panel reports the composition of CDR technologies, including Direct Air Carbon Capture and Storage (DACCS), afforestation and reforestation, and bioenergy with carbon capture and storage (BECCS). Percentages are calculated based on the number of patents within each relevant category.



Table 13: Pricing of Climate-Related Patents by Type and Technology

This table reports the relationship between different types of climate-related patents and firms' future market valuation measured by Tobin's Q. Panel A distinguishes between adaptation- and mitigation-related climate patents. Panel B further separates mitigation patents into emission reduction and carbon dioxide removal (CDR) categories. Panel C disaggregates mitigation patents by detailed technology classes. The dependent variable is Tobin's Q measured at different future horizons. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Panel A: Adaptation vs. Mitigation Patents

VARIABLES	(1) TobinQ _{t+1}	(2) TobinQ _{t+3}	(3) TobinQ _{t+5}	(4) TobinQ _{t+1}	(5) TobinQ _{t+3}	(6) TobinQ _{t+5}
Adaptation_LLM	-0.014 (-0.744)	-0.033 (-1.564)	-0.018 (-0.661)			
Mitigation_LLM				-0.003* (-1.885)	-0.003*** (-2.602)	-0.004** (-2.507)
Constant	4.990*** (12.877)	4.980*** (11.670)	4.288*** (8.873)	4.993*** (12.888)	4.990*** (11.689)	4.308*** (8.916)
Observations	16,972	13,551	10,784	16,972	13,551	10,784
R-squared	0.655	0.663	0.679	0.655	0.663	0.679
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Panel B: Emission Reduction vs. Carbon Dioxide Removal Patents

VARIABLES	(1) TobinQ _{t+1}	(2) TobinQ _{t+3}	(3) TobinQ _{t+5}	(4) TobinQ _{t+1}	(5) TobinQ _{t+3}	(6) TobinQ _{t+5}
Emission_Reduction_LLM	-0.003* (-1.900)	-0.003** (-2.553)	-0.004** (-2.435)			
CDR_LLM				-0.008 (-0.303)	-0.041 (-1.462)	-0.060** (-2.101)
Constant	4.993*** (12.887)	4.990*** (11.688)	4.308*** (8.915)	4.990*** (12.878)	4.980*** (11.670)	4.290*** (8.875)
Observations	16,972	13,551	10,784	16,972	13,551	10,784
R-squared	0.655	0.663	0.679	0.655	0.663	0.679
Firm FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES

Table 13 (continued)

Panel C: Disaggregated Mitigation Technology Categories

VARIABLES	(1) TobinQ _{t+1}	(2) TobinQ _{t+2}	(3) TobinQ _{t+3}	(4) TobinQ _{t+4}	(5) TobinQ _{t+5}
Energy_Efficiency_LLM	0.005 (0.759)	-0.002 (-0.414)	-0.007 (-1.122)	-0.005 (-0.845)	-0.007 (-1.208)
Renewable_Substitution_LLM	-0.007** (-2.406)	-0.004 (-1.646)	-0.003 (-1.179)	-0.004 (-1.273)	-0.003 (-0.935)
BECCS_LLM	-0.931 (-0.955)	0.307*** (4.183)	-0.478 (-0.664)	0.092 (0.418)	-0.152 (-0.921)
DACCS_LLM	0.002 (0.067)	-0.018 (-0.581)	-0.021 (-0.596)	-0.043 (-1.577)	-0.034 (-0.941)
Afforestation_Reforestation_LLM	0.064 (1.347)	0.036 (1.294)	-0.061 (-1.433)	-0.100** (-1.998)	-0.135* (-1.692)
Constant	4.996*** (12.891)	5.379*** (13.346)	4.989*** (11.686)	4.561*** (10.045)	4.304*** (8.905)
Observations	16,972	15,183	13,551	12,082	10,784
R-squared	0.655	0.656	0.663	0.677	0.679
Firm FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Table 14: Climate Innovation and Future Idiosyncratic Volatility

This table examines the relationship between climate-related innovation measured by the LLM-based indicator and firms' future idiosyncratic volatility. The dependent variable is idiosyncratic volatility measured at horizons $t + 1$, $t + 2$, and $t + 3$. All regressions include firm and year fixed effects. Standard errors are clustered at the firm level. The symbols *, **, and *** denote significance at the 10%, 5%, and 1% levels.

VARIABLES	(1) Idiosyncratic Volatility $_{t+1}$	(2) Idiosyncratic Volatility $_{t+2}$	(3) Idiosyncratic Volatility $_{t+3}$
climate_LLM	0.000084 (0.447)	0.000089 (0.498)	0.000042 (0.227)
Constant	0.7318*** (16.146)	0.5730*** (12.039)	0.5495*** (10.770)
Observations	18,174	16,285	14,530
R-squared	0.672	0.642	0.646
Firm FE	YES	YES	YES
Year FE	YES	YES	YES

Table 15: Chapter 4: Variable Definitions

Variable	Definition	Data Source
Independent Variables		
Climate_Y02	Number of climate-related patents granted to firm i in year t , identified based on the Y02 climate change mitigation classification of the Cooperative Patent Classification (CPC) system. The Y02 classification is defined by the European Patent Office (EPO) and the United States Patent and Trademark Office (USPTO) to capture technologies related to climate change mitigation and adaptation.	USPTO / CPC
Climate_LLM	Number of climate-related patents granted to firm i in year t , identified using a large language model (LLM)-based classification approach that evaluates patent texts and flags patents with climate-related content.	Authors' calculations
CAR[0,1] / CAR[0,5] / CAR[0,10]	Cumulative market-adjusted abnormal returns from the event date to one, five, or ten trading days thereafter.	CRSP
GHG Emissions(Scope 1)	Natural logarithm of total Scope 1 greenhouse gas emissions (in tons) emitted by firm i in year t .	S&P Trucost
Dependent Variables		

Variable		Definition	Data Source
Supplier's Climate Patent Ratio		Sales-weighted climate patent ratio of all suppliers selling products or services to a given customer in year t . The climate patent ratio is defined as the number of newly filed climate patents divided by the total number of patents filed in year t .	PatentsView
Supplier's Number of General Patents		Sales-weighted number of general patents of all suppliers selling products or services to a given customer in year t , measured as the total number of ultimately granted patents filed in year t using the $\ln(1+x)$ transformation.	PatentsView
Patent level Variables			
Y02		Indicator variable equal to one if patent p is classified under the Y02 climate change mitigation category of the CPC system, and zero otherwise.	USPTO / CPC
LLMgreen		Indicator variable equal to one if patent p is identified as climate-related by the large language model based on patent text, and zero otherwise.	LLM measures data
Xi_Real		Patent-level real economic value, constructed following (Kogan et al., 2017), based on stock market reactions to patent grant announcements and adjusted for inflation.	Data from (Kogan et al., 2017)

Variable	Definition	Data Source
Cites	Number of forward citations received by patent p , excluding self-citations, measured over a fixed citation window.	USPTO
Firm Characteristics		
Size	Natural logarithm of total assets.	Compustat
Tobin's Q	Total assets plus market value of equity minus book value of equity minus deferred taxes, divided by total assets.	Compustat
Leverage	Total debt divided by total assets.	Compustat
Tangibility	Net property, plant, and equipment divided by total assets.	Compustat
R&D	Research and development expenditures divided by total assets.	Compustat
CapEx	Capital expenditures divided by total assets.	Compustat
OCF	Operating cash flow, proxied by operating income before depreciation divided by total assets.	Compustat
Supplier's Number of General Patents	Sales-weighted number of general patents of all suppliers selling products or services to a given customer in year t , measured as the total number of ultimately granted patents filed in year t using the $\ln(1+x)$ transformation.	PatentsView