Are Transition and Physical Climate Risks Priced? Evidence from Directional News-Based Measures

Clément Toussaint *

This version: March 17, 2026

Abstract

I build directional climate risk indicators from over 100,000 news articles using a large language model. The measures distinguish transition from physical risk and, unlike existing indices, code whether news signals increasing or decreasing risk, as well as intensity and horizon. Transition-risk innovations are priced in the cross section, carrying a significantly negative risk premium, while higher transition risk predicts a deterioration in future investment opportunities, consistent with the ICAPM. This joint pricing–predictability pattern disappears when risk direction is ignored. Physical risk predicts aggregate returns but is not priced cross-sectionally, consistent with its localized nature.

Keywords: Climate risk, Large Language Models, Textual analysis, Cross-section of stock returns, ICAPM.

JEL Classification: G12, Q54.

*Centre d'Économie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne,
E-mail: clement.toussaint@univ-paris1.fr

1 Introduction

Climate change is increasingly recognized as a source of persistent and economy-wide risk, with consequences for production, investment, and long-term growth. Beyond its physical impacts, climate change also induces profound transition dynamics through regulatory uncertainty, technological shifts, and changes in consumer preferences. These forces affect firms' future cash flows and discount rates, thereby shaping investors' portfolio decisions (Krueger et al., 2020). Yet, climate-related risks—especially transition risk—are inherently forward-looking and difficult to observe directly. As a result, market participants must infer changes in the climate risk environment from the flow of information embedded in public discourse, news coverage, and market narratives. Whether these fluctuations in climate-related information—reflecting perceptions of climate risk—are priced in asset returns and are associated with changes in the investment opportunity set remains an open question.

To address this question, a growing literature studies the measurement and pricing of climate-related risks (see Giglio et al. (2021) for a comprehensive review). Several papers rely on climate-related newspaper articles to extract climate risk signals using traditional natural language processing (NLP) techniques, such as dictionary-based approaches (Engle et al., 2020; Gavriilidis, 2021; Ardia et al., 2023; Bua et al., 2024; Bessec and Fouquau, 2024; De Nard et al., 2024). Other studies employ topic-modeling techniques, such as Latent Dirichlet Allocation (LDA) (Blei et al., 2003), to identify climate-related themes in news coverage (Faccini et al., 2023).

This literature provides growing evidence that climate-related information is relevant for asset pricing. In particular, Faccini et al. (2023) show that the *U.S. Climate Policy* dimension is priced in the cross section of stock returns, while Treepongkaruna et al. (2023) find that the *Climate Policy Uncertainty* indicator (Gavriilidis, 2021) is also priced. These findings are commonly interpreted within an intertemporal asset pricing framework in the spirit of Merton (1973).

Complementary evidence based on firm-level carbon emissions also points to climate

related risk premia in equity markets (Bolton and Kacperczyk, 2021, 2023). At the same time, recent work emphasizes that emissions are disclosed with substantial delays and may embed forward-looking firm information, complicating the interpretation of carbon-return spreads as risk premia (Zhang, 2025). This motivates the use of text-based indicators that track climate-related information in real time and are not subject to emissions disclosure lags.

However, existing findings remain subject to several important limitations. First, the empirical evidence on the pricing of transition risk remains fragmented, as existing studies typically focus on specific dimensions of climate-related information rather than on a single measure of transition risk.

Second, most text-based climate risk indicators are unidirectional and do not distinguish whether climate-related information reflects an increase or a decrease in perceived risk, which complicates their economic interpretation. For instance, Engle et al. (2020) interpret increases in climate-related media attention as signaling rising transition risk, whereas Faccini et al. (2023) adopt the opposite interpretation. As a result, existing transition risk indicators rely on strong and time-specific assumptions regarding the political and regulatory environment, which limits their generalizability across regimes. In particular, the interpretation adopted by Faccini et al. (2023) may be plausible during periods of regulatory rollback, but is less likely to apply during phases of intensified climate policy. More broadly, existing textual measures abstract from key dimensions of climate risk—such as its direction, horizon, and intensity—thereby complicating their use in asset pricing applications.

Finally, the theoretical interpretation of existing empirical findings within an intertemporal asset pricing framework often remains implicit. In intertemporal asset pricing, state variables capture changes in the investment opportunity set and generate hedging demands, with implications for return predictability (Merton, 1973; Campbell, 1996). While several studies (e.g., Treepongkaruna et al., 2023) motivate their results using the ICAPM, the model's testable restrictions are rarely assessed explicitly. As a result, interpreting climate-

related risk factors as state variables may rely on assumptions that are not formally tested, echoing long-standing critiques of ICAPM implementations in empirical asset pricing (Fama, 1991). In this spirit, a common concern is that the ICAPM can become a *fishing license* (Fama, 1991) when virtually any variable can be labeled a state variable without confronting the model's joint restrictions, such as the equilibrium consistency conditions emphasized by Maio and Santa-Clara (2012). More generally, recent work shows that linking news-based factors to intertemporal restrictions is empirically feasible (e.g., Bybee et al., 2023), but this discipline has yet to be applied to transition and physical risk indicators.

This paper addresses these limitations in two steps. First, I construct multidimensional climate risk indicators from climate-related newspaper articles. The indicators distinguish (i) the type of risk (transition versus physical), (ii) direction (whether news signal an increase or a decrease in perceived risk) (iii) intensity, and (iv) the horizon at which risks are expected to materialize. The direction dimension clarifies whether climate news reflect a deterioration or an improvement in perceived risk, addressing the interpretive ambiguity of existing climate risk proxies. The additional dimensions—absent from existing text-based measures—allow me to filter and weight climate-related information in a way that sharpens the signal and reduces measurement error in the resulting indicators.

To implement this measurement strategy at scale, I use a large language model (LLM) as a text classifier. Recent evidence suggests that LLMs perform at a level comparable to human annotators on a range of social-science classification tasks, including relatively complex coding schemes (Gilardi et al., 2023; Mellon et al., 2024; Ziems et al., 2024). I apply this framework to more than 100,000 climate-related newspaper articles, producing directional, horizon-specific, and intensity-weighted transition and physical risk indicators.

Second, I study the asset-pricing implications of the transition and physical risk indicators in an intertemporal framework and assess whether the evidence satisfies the testable ICAPM implications discussed by Maio and Santa-Clara (2012). Specifically, I test the joint consistency between (i) cross-sectional risk premia, (ii) aggregate return predictability, and

(iii) the economic plausibility of the estimated market risk price when interpreted as an implied coefficient of relative risk aversion. This approach helps evaluate whether climate risks behave as ICAPM state variables that shift the investment opportunity set.

The empirical analysis yields three key findings on the asset-pricing implications of climate risk innovations in the U.S. equity market over the 2014–2024 sample period.

First, I find that transition risk (TRI) innovations are significantly priced in the cross-section of stock returns. Portfolio sorts on TRI innovation betas deliver a positive and significant alpha for the high-minus-low portfolio, largely driven by negative abnormal returns for the low-beta portfolio. Fama–MacBeth regressions (Fama and MacBeth, 1973) show a significantly negative risk premium associated with exposure to TRI innovations. This result supports an insurance interpretation: assets that covary positively with transition-risk innovations provide hedging benefits and therefore command lower expected excess returns, even though they may earn positive abnormal returns in periods when transition-risk shocks materialize. This wedge between expected and realized returns is consistent with evidence that unexpected shifts in climate-related concerns can generate ex post outperformance even when green (hedging) assets command lower expected returns ex ante (Pástor et al., 2022).

Second, the results are consistent with an intertemporal hedging interpretation in the ICAPM framework. Predictive regressions show that the TRI in levels forecasts negative aggregate market returns at multiple horizons, indicating that increases in transition risk are associated with a deterioration in future investment opportunities. Combined with the cross-sectional evidence of a negative price of TRI innovation risk, this pattern matches the testable ICAPM restrictions emphasized by Maio and Santa-Clara (2012): state variables that predict worsening market conditions should command negative prices of risk, as investors value assets that hedge against adverse shifts in the investment opportunity set. In line with this mechanism, TRI-sensitive stocks earn lower expected excess returns because they provide hedging benefits when transition risk rises. Finally, the estimated market risk price is economically plausible, satisfying the third ICAPM restriction emphasized by Maio and

Santa-Clara (2012).

Third, in contrast to transition risk, I find that physical-risk innovations are not priced in the cross section of U.S. stock returns. Although the PRI in levels predicts future aggregate market returns—suggesting a (macroeconomic) risk component—portfolios sorted on exposure to PRI innovations show no significant abnormal returns. Furthermore, exposure to PRI innovations does not earn a statistically significant risk premium in the Fama–MacBeth regressions. This decoupling between time-series predictability and cross-sectional pricing suggests that physical risk may be harder to capture with aggregate, national news-based measures, given its inherently local and spatial nature (Baldauf et al., 2020).

Overall, the evidence for the transition risk indicator is consistent with an intertemporal hedging mechanism in the spirit of Merton (1973) and recent climate asset-pricing models (Pástor et al., 2021; Barnett, 2023).

This paper makes three contributions to the green finance literature and empirical asset pricing. First, I provide cross-sectional evidence that transition risk innovations are priced in U.S. stock returns. Second, to the best of my knowledge, I provide the first joint evidence linking cross-sectional pricing to time-series return predictability in a way that is consistent with testable ICAPM implications, supporting a state-variable interpretation of transition risk. Third, I develop new transition and physical risk indicators that explicitly incorporate risk direction (increase versus decrease), as well as horizon and intensity, using a LLM-based text classification methodology. Importantly, incorporating direction is essential for reconciling cross-sectional pricing with the intertemporal (ICAPM) restrictions that motivate a state-variable interpretation of transition risk. In fact, when direction is removed and the indicator is reduced to a unidirectional attention measure, the joint pricing-and-predictability evidence no longer supports the ICAPM-based hedging interpretation.

The remainder of the paper is organized as follows. section 2 describes the construction of the transition and physical climate risk indicators. section 3 presents the data and summary statistics. section 4 reports the asset pricing methods and results. section 5 concludes.

2 Construction of the Climate Risk Indicators

The objective of this section is to construct textual indicators that proxy transition and physical climate risks using newspaper articles. These indicators are designed to capture multiple dimensions of risk: its type, direction, intensity, and horizon. The construction proceeds in three steps. First, each article is analyzed using a large language model (LLM) to extract the relevant risk dimensions. Second, these article-level signals are aggregated into daily transition and physical risk time series. Finally, topic modelling is used to characterise the thematic content underlying the textual risk indicators.

2.1 LLM-Based Climate Risk Annotation

2.1.1 Annotation Framework and Classification Validation

I use a large language model (LLM), namely *GPT-5-mini*, to analyze the content of each news article. This approach aligns with a growing methodological literature documenting how generative AI can support social-science measurement by enabling structured text coding (Bail, 2024). It is also consistent with recent empirical applications that use LLMs to classify economically meaningful policy content in text data (Cook et al., 2023; Jha et al., 2024).

Operationally, I implement this measurement strategy using a structured prompting approach. For each article, I provide the LLM with a prompt that asks it to answer five multiple-choice questions, following the workflow depicted in Figure 1. The prompt includes concise definitions and decision rules. The questions are reproduced below (without the accompanying instructions and definitions; the full prompt is provided in Appendix A):

Q.1 *Is the following article related to climate change?* Yes / No,

Q.2 *What type of climate risk does it mainly refer to?* Transition / Physical / Both / Unclear,

Q.3 *From the general point of view of a polluting firm reasoning in the short run and at its own level (not macro), does the article suggest that the risk is increasing, decreasing,*

or remaining neutral? Increase / Decrease / Neutral.

Q.4 *If you have not answered "Neutral" previously, what is the intensity*[1] *of the risk described in the article?* Weak / Medium / Strong / N.A.,

Q.5 *What is the time horizon*[2] *of the identified climate risk in the article?* Short / Medium / Long / N.A.

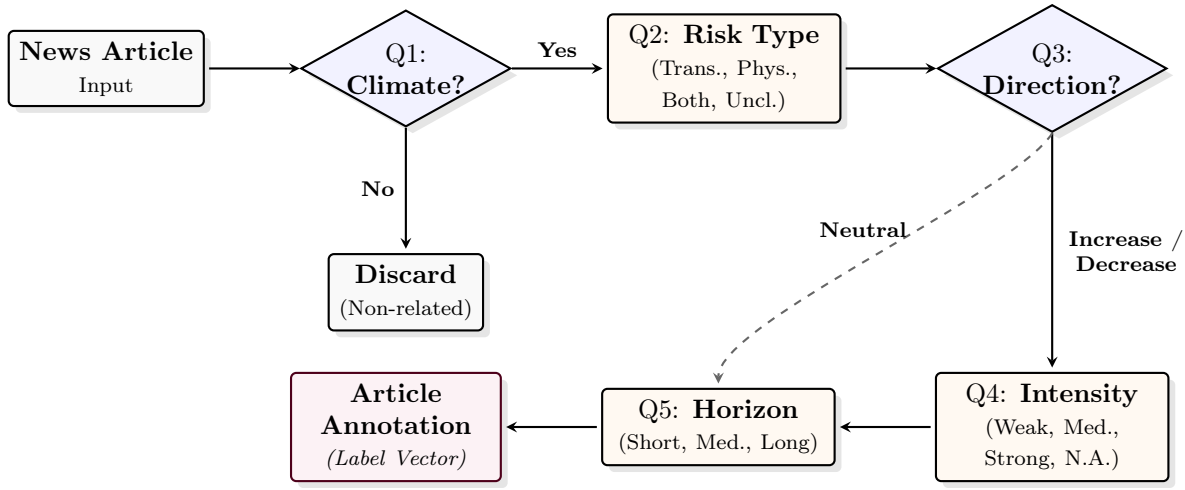


Figure 1. LLM Annotation Workflow

Each article is processed sequentially through five classification tasks. Articles unrelated to climate change (Q1 = No) are discarded. For climate-related articles, the model identifies the risk type (Q2), its direction (Q3), intensity (Q4), and time horizon (Q5). When the direction is neutral (Q3 = Neutral), intensity is not assessed (Q4 = N.A.) and the article proceeds directly to horizon classification. The output is a five-element label vector used to construct the daily risk indicators.

The choice of prompt formulation follows a systematic evaluation. I tested three specifications on a manually annotated sample of approximately 300 articles, randomly sampled across outlets and over the full sample period (2010–2024): a baseline prompt with direct questions only (P.1), a structured prompt with concise definitions (P.2), and an extended prompt with additional examples and edge cases (P.3). Classification performance is broadly stable across specifications; adding more detailed instructions did not improve—and in some

[1]The definition of risk *intensity* is provided in Appendix A. Briefly, the intensity depends on the geographical scope, the degree to which the discussed measure or event is binding, and its quantitative impact.

[2]The definition of *time horizon* is provided in Appendix A. Briefly, a *short* horizon means that the risk is expected to materialize within 0–3 years, a *medium* horizon within 3–10 years, and a *long* horizon beyond 10 years.

cases degraded—F1-scores (Appendix B). This finding motivates the relatively concise definitions used in the final prompt: overly specific instructions appear to constrain the model's ability to generalize across the heterogeneous language of news articles. I retain prompt P.2 for the full corpus classification.

Table 1: LLM Classification Performance on Manually Annotated Articles

Question	Support	Micro (%)			Macro (%)			MAE	nMAE
		F1	P	R	F1	P	R		
Q1	271	86.7	86.7	86.7	86.3	88.5	85.9		
Q2	141	87.9	87.9	87.9	62.7	63.9	62.3		
Q3	141	88.7	88.7	88.7	64.7	67.6	62.9	0.184	0.092
Q4	135	61.5	61.5	61.5	42.3	45.5	43.3	0.474	0.237
Q5	135	55.6	55.6	55.6	31.4	31.4	35.4	0.556	0.278

Notes: This table reports the performance of the LLM-based classification on the manually annotated newspaper sample. Q1–Q5 correspond to the five annotation tasks. For each task, the table reports F1-score, precision (P), and recall (R), computed using both micro- and macro-averaging; all values are expressed in percentages. Micro-averaged metrics weight observations by their frequency, while macro-averaged metrics give equal weight to each class. For the ordinal classification tasks (Q3–Q5), the table additionally reports the mean absolute error (MAE) and the normalized MAE (nMAE). Support denotes the number of manually annotated articles used for evaluation.

Table 1 reports the classification performance of the selected prompt (P.2) on the manually labelled subset. For the first three questions (Q1–Q3), micro-averaged F1-scores exceed 85%, and macro-averaged F1-scores range from 62% to 86%. These levels are consistent with recent applications of LLMs to financial and economic text classification, where F1-scores in the 60–85% range are typical for subjective annotation tasks involving sentiment, risk type, and temporal orientation (Cook et al., 2023).

For Questions 4 and 5, macro F1-scores fall below 50%, reflecting the inherent difficulty of distinguishing adjacent ordinal categories (e.g., *Medium* vs. *Strong* intensity). However, the normalized mean absolute error (nMAE) of 0.24 for Q4 and 0.28 for Q5 indicates that classification errors are concentrated in adjacent categories. To interpret these values: nMAE is bounded between 0 (perfect classification) and 1 (systematic misclassification at opposite ends of the scale); an nMAE below 0.5 implies that the average prediction error is less than

one ordinal category. The observed values thus confirm that the model rarely confuses *Short* with *Long* or *Weak* with *Strong*—errors that would materially distort the risk indicators. Because adjacent categories receive nearby weights in Equation 1 (e.g., $w_I = 1$ for Medium vs. $w_I = 2$ for Strong), confusing adjacent categories has a smaller effect on the aggregate score than confusing distant ones (e.g., Weak vs. Strong).

Given these performance levels, I retain all five questions to construct the daily textual risk indicators presented in the next subsection.

2.1.2 Robustness to Data Contamination

A potential concern with LLM-based classification is that the model may leverage information from its training data rather than relying solely on the article text. In an asset-pricing context, this could introduce forward-looking bias: if the model's classifications are influenced by knowledge of subsequent events—such as the passage of legislation, market reactions, or realized climate outcomes—the resulting indicators would embed information unavailable to investors at the time of publication. This would invalidate the interpretation of the indicators as real-time measures of climate risk perceptions and bias predictive regressions. I address this concern through three complementary checks.

Pre- vs. post-cutoff performance. To test whether the model relies on memorized content, I compare classification performance on articles published before and after the model's training cut-off date. For this test, I use *GPT-3.5-Turbo*, whose training cut-off (September 2021) is publicly documented, allowing a clean separation between articles the model may have seen during training and those it could not have seen. If memorization were driving the results, performance should deteriorate on post-cutoff articles. As reported in Table 12 (Appendix B), classification accuracy is similar across both groups, consistent with the model relying on article content rather than prior knowledge.

Prompt design. The prompt structure further limits forward-looking bias by construction. The model is asked to classify what the article *says*—e.g., "does the article suggest that risk is increasing or decreasing?"—not to assess the ex-post importance of the events described or their subsequent market impact. The questions are intentionally general and do not reference specific climate events, policy outcomes, or asset returns. This design reduces the scope for forward-looking bias: even if the model has knowledge of post-publication developments, it is not instructed to use that knowledge. Related evidence suggests that forward-looking biases remain modest even in LLM-based forecasting tasks, where models are explicitly asked to predict future outcomes (Chen et al., 2022; He et al., 2025); such concerns are structurally less prominent in text-classification settings like mine, where the model is asked only to characterize article content rather than forecast outcomes.

Q1–Q3 ablation. Among the five classification dimensions, intensity (Q4) and horizon (Q5) are most susceptible to hindsight bias. Judging whether a risk is "strong" or "weak," or whether it materializes in the "short" or "long" term, may be influenced by knowledge of how events ultimately unfolded. By contrast, classifying an article as climate-related (Q1), identifying the risk type (Q2), and coding the directional signal conveyed by the text (Q3) are more directly tied to article content. I therefore construct alternative indicators using only Q1–Q3, setting intensity and horizon weights to one in Equation 1. The main asset-pricing results—cross-sectional risk premia, return predictability, and ICAPM consistency—remain unchanged (Appendix F). This confirms that the findings do not depend on the dimensions most vulnerable to look-ahead concerns.

2.2 Transition and Physical Risk Indicators

To construct the transition and physical risk indicators, I first assign a numerical score to each article based on its LLM-labelled attributes, and then aggregate these scores at a daily frequency.

2.2.1 Article Scoring

Each article a is assigned a scalar score that (i) filters out non–climate-related content, (ii) encodes the *direction* of the risk (increase, neutral, decrease), and (iii) weights the signal by *intensity* and *time horizon*.

(i) Filtering. If the LLM labels as climate-related, the article is scored; otherwise its score is set to zero:

$$\mathbf{1}_a^{\text{clim}} = \begin{cases} 1 & \text{if article } a \text{ is labelled as climate-related,} \\ 0 & \text{otherwise.} \end{cases}$$

This filtering step removes articles that are not closely related to climate change from the construction of the transition and physical risk indices, thereby reducing noise in the final indicators.

(ii) Mappings. I then map each risk dimension to a numerical weight as follows:

$$D_a = \begin{cases} +1 & \text{if Increase,} \\ 0 & \text{if Neutral,} \\ -1 & \text{if Decrease,} \end{cases} \quad I_a = \begin{cases} 0.5 & \text{if Weak,} \\ 1.0 & \text{if Medium,} \\ 2.0 & \text{if Strong,} \\ 1.0 & \text{if NA (7\%),} \end{cases} \quad H_a = \begin{cases} 2.0 & \text{if Short,} \\ 1.0 & \text{if Medium,} \\ 0.5 & \text{if Long,} \\ 1.0 & \text{if NA (6\%).} \end{cases}$$

When the classifier returns NA for intensity or horizon, the neutral value of 1.0 is assigned, ensuring that such articles contribute to the indicator without disproportionately inflating or deflating the score. These cases are infrequent, representing only 7% of articles for intensity and 6% for horizon, and thus have limited influence on the aggregate indicators. Under these weights, an article conveying a *strong* climate risk receives the same weight as four *weak*

articles. Stronger intensities and shorter horizons are up-weighted, while weaker intensities and longer horizons are down-weighted. As a result, the indicators are more sensitive to climate-related information that is more likely to affect firms materially.

(iii) Article score. Finally, each article is assigned a score based on its risk-dimension weights and signed according to its risk direction. The signed contribution of article a is given by the direction multiplied by the sum of the intensity and horizon weights:

$$\text{Score}_a = \mathbf{1}_a^{\text{clim}} D_a (I_a + H_a). \tag{1}$$

This rule yields a positive (negative) score when the article signals an increase (decrease) in the relevant climate risk, with the magnitude scaled up for stronger intensities and shorter horizons, and scaled down for weaker intensities and longer horizons.

Transition vs Physical. Since each article is classified as either *transition* or *physical*, the score of article a is assigned to the corresponding risk category based on its risk-type label. Each article therefore contributes either to the transition-risk indicator or to the physical-risk indicator, depending on its risk-type label.

2.2.2 Daily Indicators

From the articles scores, I construct two daily newspaper-based indicators: a *Transition Risk Index* (TRI) and a *Physical Risk Index* (PRI). For the daily aggregation (step *(i)* to *(iii)*) I follow the methodology of Ardia et al. (2023), inspired by Baker et al. (2016).

Let $o \in \mathcal{O}$ index media outlets and t index days (from 2010-01-01 to 2024-12-31). For outlet o on day t, let $\mathcal{A}_{o,t}$ denote the set of climate-related articles identified by the LLM classifier. Each article $a \in \mathcal{A}_{o,t}$ is associated with an article-level risk score, corresponding to Equation 1: a Transition Risk Score for articles labeled as transition risk, and a Physical Risk Score for articles labeled as physical risk:

$$\text{Tr.RiskScore}_a \quad \text{Ph.RiskScore}_a, \qquad a \in \mathcal{A}_{o,t}. \tag{2}$$

(i) Outlet–day aggregation. For each outlet and day, I aggregate article scores by summation:

$$s_{o,t}^{\text{TRI}} = \sum_{a \in \mathcal{A}_{o,t}} \text{Tr.RiskScore}_a \quad \text{and} \quad s_{o,t}^{\text{PRI}} = \sum_{a \in \mathcal{A}_{o,t}} \text{Ph.RiskScore}_a. \tag{3}$$

(ii) Outlet–level normalization. I then normalize the aggregated score by each outlet's historical standard deviation:

$$\sigma_o^{\text{TRI}} = \text{sd}\big(\{s_{o,t}^{\text{TRI}}\}_t\big) \quad \text{and} \quad \sigma_o^{\text{PRI}} = \text{sd}\big(\{s_{o,t}^{\text{PRI}}\}_t\big), \tag{4}$$

and scale:

$$\tilde{s}_{o,t}^{\text{TRI}} = \frac{s_{o,t}^{\text{TRI}}}{\sigma_o^{\text{TRI}}} \quad \text{and} \quad \tilde{s}_{o,t}^{\text{PRI}} = \frac{s_{o,t}^{\text{PRI}}}{\sigma_o^{\text{PRI}}}. \tag{5}$$

This transformation accounts for systematic differences in scale and volatility across newspapers. Outlets vary substantially in their editorial style, coverage intensity, and propensity to report on climate-related issues. Without normalization, outlets with higher levels or greater variance in coverage would mechanically dominate the aggregate indicator. Scaling each series by its outlet-specific standard deviation places all outlets on a comparable footing and ensures that the aggregate transition and physical risk indicators reflect common information rather than outlet-specific reporting patterns.

(iii) Daily cross–outlet aggregation. I aggregate the outlet-level normalized scores by computing, for each day t, the cross-sectional average across all outlets that publish at least one article:

$$\mathrm{TRI}_t^{\mathrm{raw}} = \frac{1}{|\mathcal{O}_t|} \sum_{o \in \mathcal{O}_t} \tilde{s}_{o,t}^{\mathrm{TRI}} \quad \text{and} \quad \mathrm{PRI}_t^{\mathrm{raw}} = \frac{1}{|\mathcal{O}_t|} \sum_{o \in \mathcal{O}_t} \tilde{s}_{o,t}^{\mathrm{PRI}}, \tag{6}$$

where $\mathcal{O}_t \subseteq \mathcal{O}$ is the set of outlets active at date t.

(iv) Calendar completion. Finally, I complete the daily calendar between the minimum and maximum sample dates to construct the *transition risk indicator* (TRI) and the *physical risk indicator* (PRI). Days with no climate-related publications are interpolated using a Kalman smoother. Over the sample period 2014-01-01–2024-12-31 (used in the asset pricing analysis), only 62 out of 4,018 days (1.5%) require interpolation, ensuring that this step has no material impact on the empirical results.

2.3 Topic Modelling

To provide a descriptive overview of the content of climate-related news articles, I perform a topic modeling analysis using a standard embedding–clustering approach. Each article is first embedded using the `text-embedding-3-large` model, producing high-dimensional semantic representations. To reduce noise and improve tractability, the embeddings are projected onto a lower-dimensional space using principal component analysis. The reduced embeddings are then clustered using the k-means algorithm with $K = 40$ clusters. Cluster labels are generated by a large language model based on representative articles and indicative keywords. Finally, semantically related clusters are manually grouped into 10 broader thematic categories. This topic modeling exercise is used exclusively for descriptive purposes and does not enter the construction of the climate risk indicators or the asset pricing analysis. Full methodological details are provided in Appendix C.

3 Data

The analysis relies on two main data sources: textual data in the form of newspaper articles, which are used to construct the transition and physical climate risk indicators, and financial data, including stock returns and standard risk factors.

3.1 Textual Indicators

3.1.1 Newspaper Data

To construct textual indicators proxying transition and physical climate risk, I collect 161,962 newspaper articles published across seven major outlets between January 1, 2010 and December 31, 2024. Article URLs are obtained from *MediaCloud*, an open-source media analysis platform originally developed by the MIT Center for Civic Media and the Berkman Klein Center at Harvard University. The platform is currently maintained and developed through a consortium involving the University of Massachusetts Amherst, Northeastern University, and the Media Ecosystems Analysis Group. MediaCloud provides a systematically collected archive of online news content from thousands of outlets, together with consistent metadata such as publication dates, outlet identity, and URLs.

From MediaCloud, I retrieve article URLs whose body text contains at least one of the following climate-related expressions: *climate change, global warming, climate crisis, carbon emissions, Paris Agreement, renewable energy, climate policy, climate risk, sea level rise, carbon tax*. The keyword list is intentionally broad in order to maximize recall at the initial retrieval stage. Many articles containing these expressions are not primarily related to climate change and are subsequently filtered out using a large language model (LLM). This two-step procedure minimizes the risk of excluding relevant climate-related content.

The selected outlets are the *Daily Mail*, *Los Angeles Times*, *New York Post*, *National Public Radio* (NPR), *The Globe and Mail*, *The Guardian*, and *The Washington Post*. This selection yields a corpus that is substantially broader and more diverse than those used in

prior studies of climate risk in financial markets. For comparison, Engle et al. (2020) rely exclusively on the *Wall Street Journal*, and Faccini et al. (2023) use only Reuters newswires. The present corpus spans seven outlets across three countries, includes both broadsheet and tabloid formats, and covers the full political spectrum—from the *Daily Mail* and *New York Post* on the right to *The Guardian* and *NPR* on the left.

The outlets span the United States (*Los Angeles Times, New York Post, NPR, Washington Post*), the United Kingdom (*Daily Mail, The Guardian*), and Canada (*The Globe and Mail*). Although the empirical analysis focuses on U.S. equity markets, including international outlets is appropriate for two reasons. First, climate risk is inherently global: U.S. firms are exposed to regulatory developments abroad, international climate agreements, and extreme weather events worldwide. Second, English-language financial news circulates freely across borders, and major climate events—COP summits, IPCC reports, hurricanes—receive consistent coverage across all outlets regardless of geography. The outlet-level normalization described in subsubsection 2.2.2 ensures that differences in editorial style or publication volume across outlets do not mechanically influence the aggregate indicators.[3]

I then web-scrape the full text of each article, yielding a final corpus of 161,962 documents. The sample starts in 2010, reflecting limitations in the stability and accessibility of URLs for older articles, which prevents reliable large-scale text retrieval prior to this date and constrains the construction of a longer consistent time series.

Table 2 reports, for each outlet, the number of collected articles, the share classified as climate-related by the LLM, and the average transition and physical risk scores. Overall, 37.36% of the retrieved articles are excluded by the LLM filter, leaving 101,456 climate-related articles for constructing the indicators.

The share of articles classified as climate-related is relatively similar across outlets. Average physical risk scores also display comparable levels, suggesting limited heterogeneity in

[3]The *New York Times* and *Wall Street Journal* are excluded due to technical constraints on systematic full-text retrieval. Given the breadth of the existing corpus—which already includes four major U.S. outlets—this exclusion is unlikely to materially affect the indicators.

Table 2: Descriptive Statistics per Outlet

Outlet	Country	Articles			Indicators	
		Total	Green	% Green	Av. Transition Risk Score	Av. Physical Risk Score
Los Angeles Times	USA	13,170	8255	62.68	0.86	2.53
New York Post	USA	5,566	3,459	62.15	1.68	2.56
NPR	USA	6,463	4,477	69.27	1.02	2.66
Washington Post	USA	25,382	14,710	57.95	0.42	2.63
The Guardian	UK	51,130	36,965	72.30	1.11	2.67
Daily Mail	UK	52,186	28,175	53.99	1.51	2.56
The Globe and Mail	Canada	8,065	5,415	67.14	1.81	2.77
Total		161,962	101,456	62.64		

Notes: This table reports outlet-level descriptive statistics. *Total* is the number of collected articles; *Green* is the number of articles classified as climate-related by the LLM (Q1 = Yes); *% Green* is the corresponding share. *Av. Transition Risk Score* and *Av. Physical Risk Score* are the mean article-level scores computed over climate-related articles classified as transition risk (Q2 = Transition) and physical risk (Q2 = Physical), respectively.

the coverage of physical climate-related events.

By contrast, transition risk scores exhibit more variation across outlets. The *Los Angeles Times*, *NPR*, and the *Washington Post* have lower average transition risk scores, with means equal to or below 1. As seen in Figure 2, these outlets also display negative yearly average transition risk scores during 2016–2018, a period that coincides with major shifts in U.S. federal climate policy (including the election of Donald Trump). This heterogeneity underscores differences in how newspapers report transition-related developments and motivates the normalization and aggregation procedures described in subsection 2.2.

For physical risk, Figure 2 shows a gradual increase in yearly average physical risk scores across all outlets, consistent with the levels reported in Table 2 and with the growing frequency of physical climate events over time.

Figure 3a plots the monthly number of climate-related articles, while Figure 3b displays the monthly distribution of topics. Prior to 2014, the volume of articles is very low—typically fewer than 50 articles per month—and topic shares exhibit substantial month-to-month fluctuations. After 2014, the number of articles increases noticeably (exceeding 100

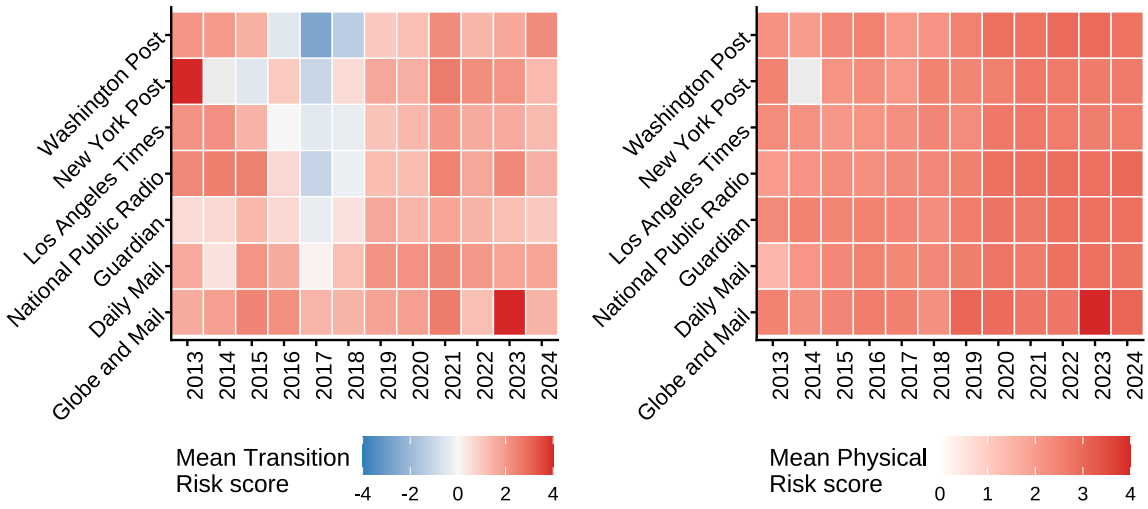


Figure 2. Average Transition and Physical Risk Scores by Outlet and Year

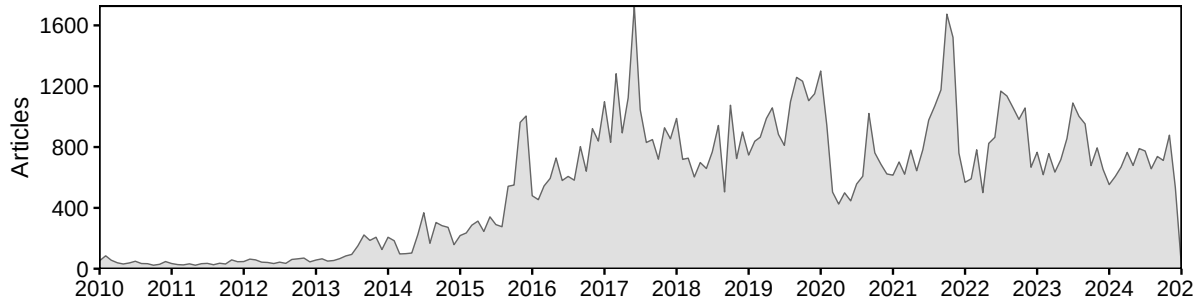
This figure displays the average article-level transition risk score (left panel) and physical risk score (right panel) by outlet and year. For the transition risk score, blue indicates negative values (risk-decreasing news) and red indicates positive values (risk-increasing news). For the physical risk score, darker red indicates higher values. Gray cells indicate missing data.

articles per month), and the distribution of topics becomes more stable. For this reason, the TRI and PRI time series are constructed over the period 2014-01-01 to 2024-12-31, thereby avoiding the noise associated with sparse media coverage in the earlier period.

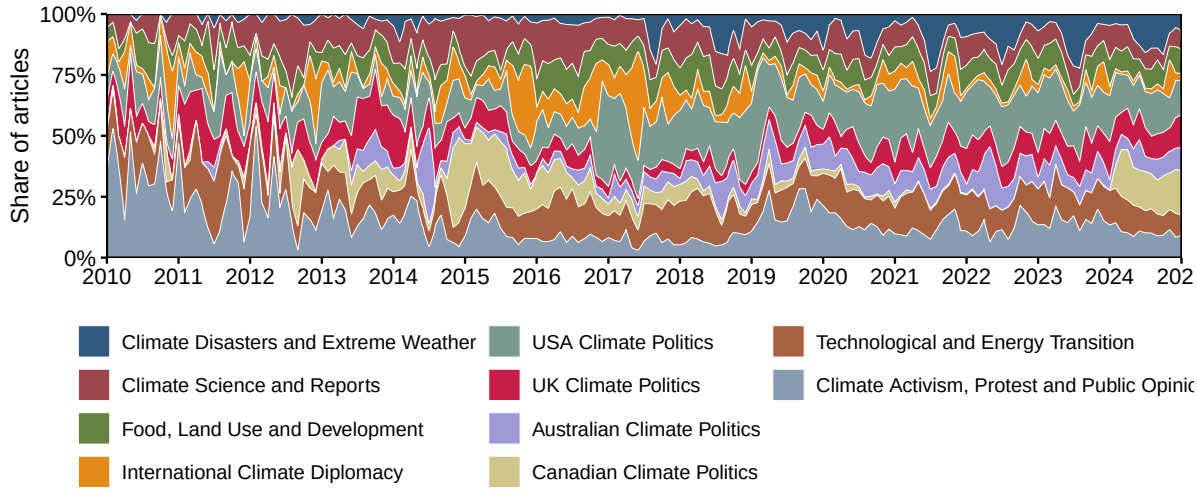
This sample choice is consistent with recent evidence that climate risk pricing in equity markets emerged only in the mid-2010s. Faccini et al. (2023) find that their climate risk indicator is priced only after 2012, and Bua et al. (2024) document significant asset pricing effects only from 2015 onward. These findings suggest that climate risk became financially material—and thus reflected in news coverage and asset prices—around this period, likely driven by the Paris Agreement (2015) and subsequent regulatory developments.

The number of climate-related articles exhibits sharp peaks around major climate-policy events. For example, article volume rises significantly in December 2015 (COP21) and in June 2017 (the U.S. announcement of its intended withdrawal from the Paris Agreement). At the time of these events, the share of articles classified under the "International Climate Diplomacy" topic also increases. Outside these event-driven spikes, the dominant topic in the corpus throughout the sample is "USA Climate Politics", which consistently represents

the largest share of monthly articles.



(a) Monthly Volume of Climate-Related Articles



(b) Thematic Composition of Climate-Related Articles

Figure 3. Volume and Thematic Composition of Climate-Related News Coverage

Panel (a) displays the monthly count of articles classified as climate-related by the LLM. Panel (b) shows the share of each thematic cluster within climate-related articles over time. Themes are identified using a topic modeling approach described in Appendix C.

3.1.2 Transition and Physical Risk Indicators

The TRI and PRI are constructed from the LLM outputs, which consist of answers to a sequence of five multiple-choice questions. Each article therefore receives five labels corresponding to: (i) climate relevance, (ii) risk type, (iii) risk direction, (iv) risk intensity, and (v) time horizon. Figure 4 illustrates the distribution of these labels through an alluvial diagram.

Among articles classified as transition-related, most are labelled as reflecting an *increasing* level of transition risk, with a non-trivial share labelled as *decreasing*. Within this decreasing group, the LLM frequently assigns high intensity and short-horizon labels, indicating that these articles tend to describe salient developments even when they point to a reduction in risk.

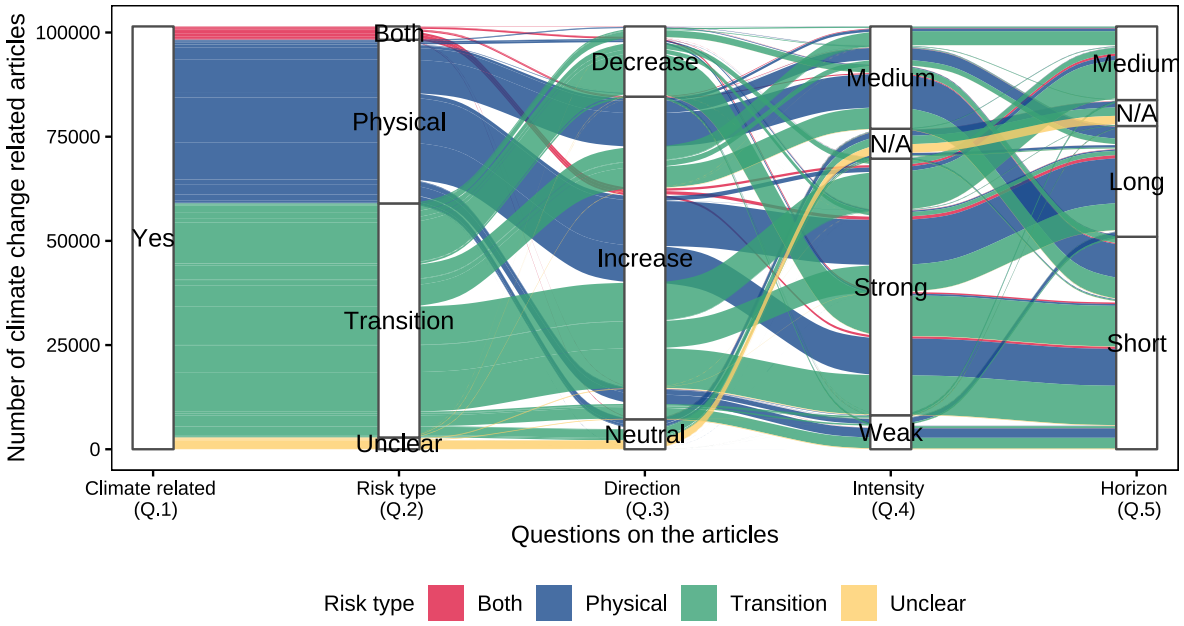


Figure 4. Alluvial Representation of the LLM-Based Classification of News Articles

This alluvial diagram illustrates the classification of articles across the five LLM questions, read from left to right. Each vertical bar represents a question: climate relevance (Q1), risk type (Q2), risk direction (Q3), risk intensity (Q4), and time horizon (Q5). The width of each flow is proportional to the number of articles. Colors indicate risk type: green for transition, blue for physical, red for both, and yellow for unclear. For example, the green flows show that most transition risk articles are classified as risk-increasing (Q3), with strong intensity (Q4) and short horizon (Q5).

Articles classified as physical-risk related are overwhelmingly labelled as *increasing*, consistent with the fact that media coverage of physical climate-related events predominantly focuses on escalating or ongoing risks.

Overall, more than three quarters of the climate-related articles are labelled as indicating an *increase* in climate risk, and more than half are labelled as *strong* in terms of intensity. The distribution across time-horizon categories is more balanced, although the *short* horizon

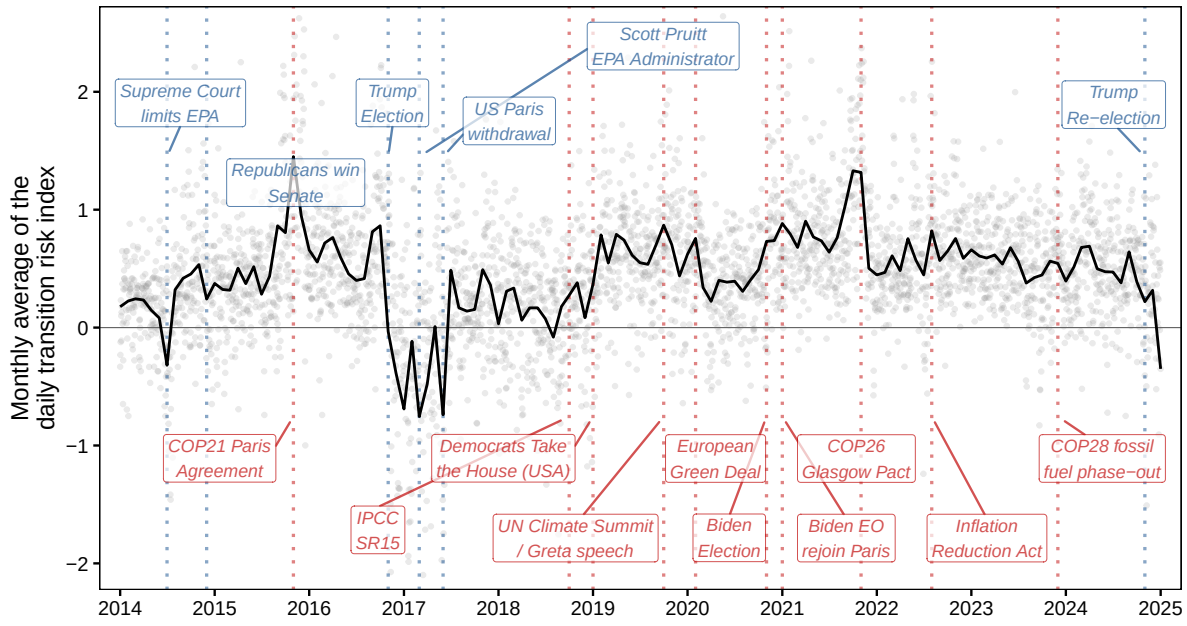
is the most prevalent category.

From these article-level classifications, I construct the transition (TRI) and physical (PRI) risk indicators, which proxy the levels of transition and physical climate risk, respectively. A positive (negative) value of the TRI indicates an increase (decrease) in transition risk, with a similar interpretation for the PRI.

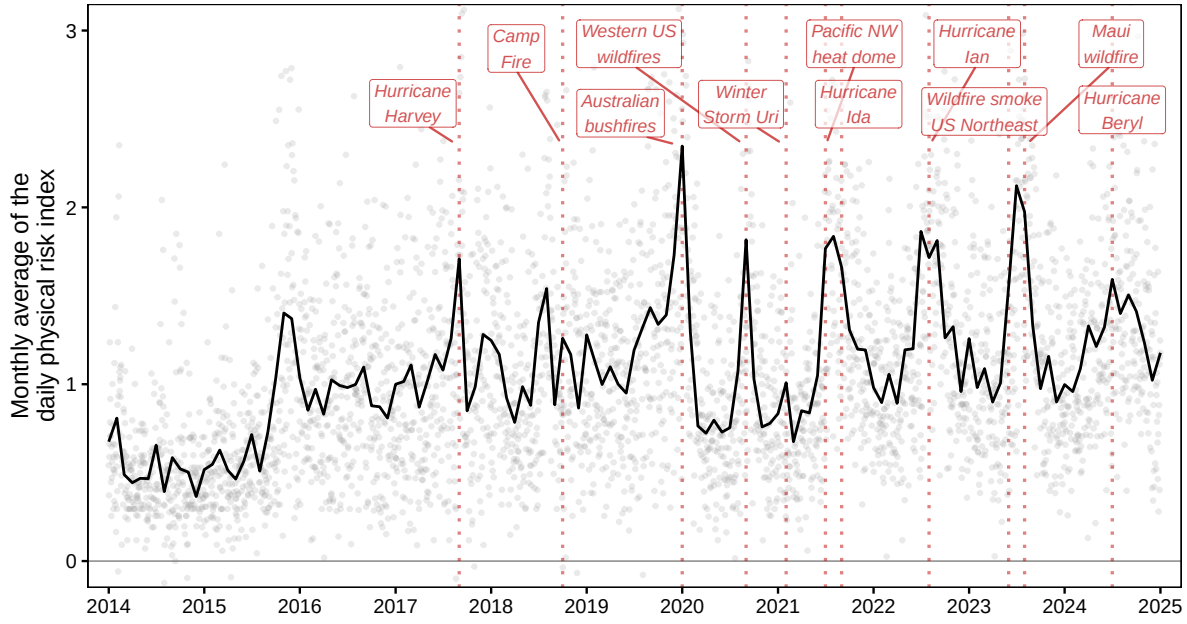
Figure 5a displays the monthly average of the transition risk indicator (TRI) from January 2014 to December 2024, together with major policy and regulatory events affecting climate transition risk. The TRI is on average positive over the sample, with the exception of the period between November 2016 and June 2017, during which the indicator turns negative. This interval coincides with several major U.S. political events: the election of Donald Trump, the announcement of the U.S. withdrawal from the Paris Agreement, and Executive Order 13783, which rolled back a number of federal environmental regulations. After mid-2017, the indicator stabilizes around zero until January 2019, following the 2018 midterm elections in which the Democratic Party gained control of the U.S. House of Representatives. From 2019 onward, the TRI monthly average fluctuates around 0.5, with pronounced spikes around events such as the election of Joe Biden, COP26, and the passage of the Inflation Reduction Act. Toward the end of the sample, the indicator turns negative again following the election of Donald Trump for a second term. Overall, the TRI dynamics align closely with major transition-related political events, suggesting that the indicator effectively proxies for transition risk.

Regarding the Physical Risk Indicator (PRI), Figure 5b displays its monthly average over the same period. The PRI exhibits a moderate upward trend, with noticeable spikes around major natural disasters such as hurricanes, wildfires, and severe storms. This pattern is consistent with the increasing frequency and intensity of physical climate-related events documented in recent years.

Thus, the PRI time series appears to proxy for physical climate risk, exhibiting higher



(a) Transition Risk Indicator (TRI)



(b) Physical Risk Indicator (PRI)

Figure 5. Daily and Monthly Climate Risk Indicators

Panel (a) displays the transition risk indicator (TRI) and Panel (b) displays the physical risk indicator (PRI) from January 2014 to December 2024. Gray dots represent daily values; the black line shows the monthly average. Vertical lines indicate major climate-related events, colored by their expected effect on risk: red for risk-increasing events, blue for risk-decreasing events.

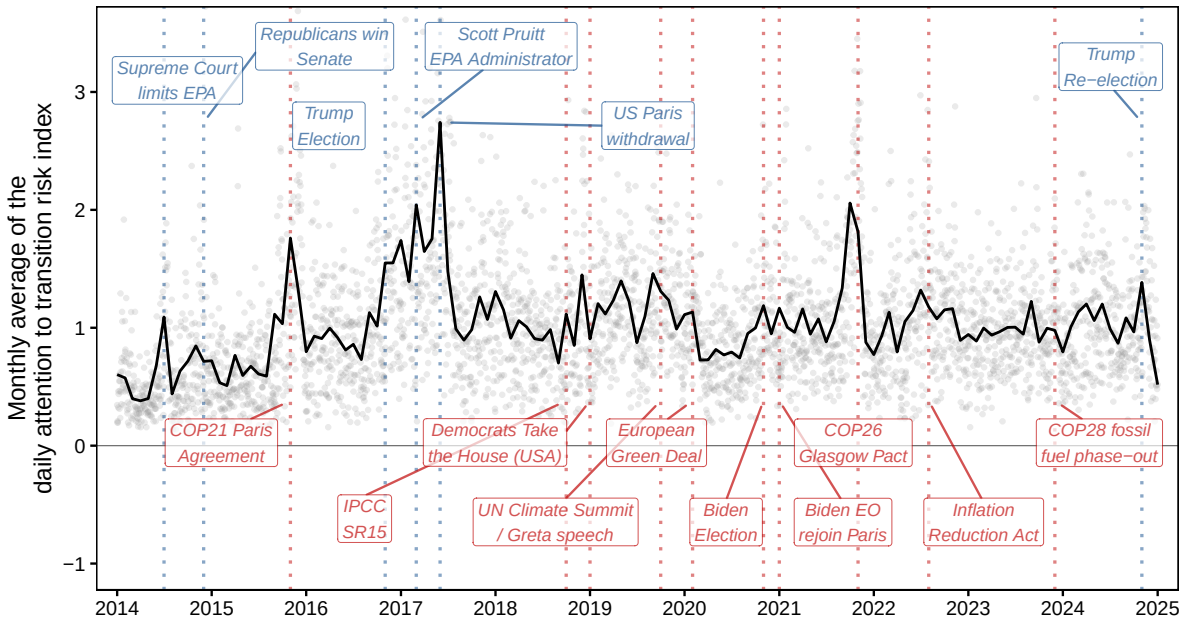
values during periods of major natural disasters and a gradual upward drift over time.

Finally, I also construct attention indicators for transition (AT) and physical (AP) risk to highlight the methodological contribution of the LLM, namely the inclusion of the *risk direction* dimension. These attention-based indicators weight each article by its intensity and horizon, but not by its direction.

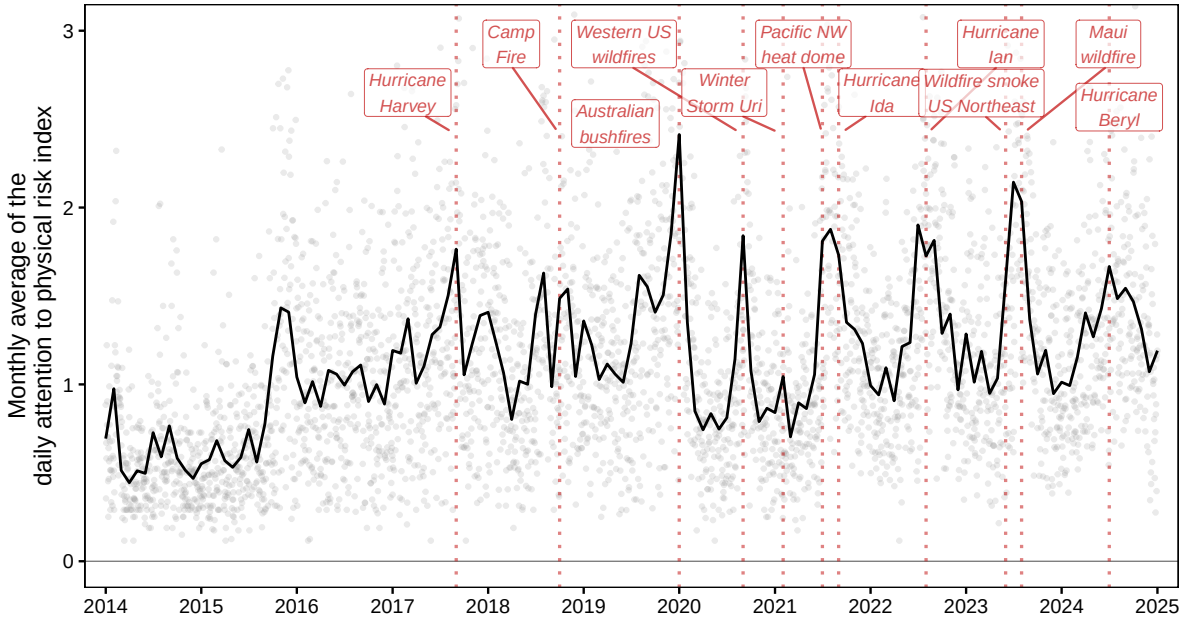
Figure 6a displays the monthly average of the AT indicator over the same period and with the same event markers as Figure 5a. During the period characterized by Democratic control of the U.S. House of Representatives (January 2019–January 2021) and the subsequent Democratic administration (January 2021–November 2024), the AT closely resembles the TRI. In contrast, during periods in which both the presidency and the House are controlled by Republicans (November 2016–January 2019 and from November 2024 onward), the two indicators diverge substantially. During these periods, the TRI declines, indicating a reduction in transition risk, whereas the AT primarily captures heightened attention to transition-related news without reflecting the direction of the underlying risk.

This comparison illustrates that attention is an imperfect proxy for transition risk: it responds strongly to salient events regardless of whether they increase or decrease the level of risk. For instance, the AT rises both around COP26 and around Donald Trump's election, even though these events have opposite implications for transition risk. As a consequence, attention-based measures approximate transition risk only in periods when most events tend to increase the risk, such as during Democratic administrations. The LLM-based TRI, by incorporating risk direction, avoids this limitation and provides a more informative proxy for transition risk.

By contrast, Figure 6b shows that the attention-based indicator for physical risk (AP) is nearly indistinguishable from the PRI itself. For physical risk, the direction label does not materially affect the indicator, which is expected given that physical climate risks rarely decrease in the short and medium run.



(a) Attention to Transition Risk (AT)



(b) Attention to Physical Risk (AP)

Figure 6. Daily and Monthly Climate Risk Attention Indicators

Panel (a) displays the attention to transition risk indicator (AT) and Panel (b) displays the attention to physical risk indicator (AP) from January 2014 to December 2024. These indicators are constructed by setting the direction weight to one in Equation 1, while keeping intensity and horizon weights unchanged. Unlike the directional indicators in Figure 5, these attention indicators capture the volume and intensity of climate risk coverage regardless of whether the news suggests risk is increasing or decreasing. Gray dots represent daily values; the black line shows the monthly average. Vertical lines indicate major climate-related events, colored by their expected effect on risk: red for risk-increasing events, blue for risk-decreasing events.

Additional descriptive statistics on the topics discussed in each outlet are reported in Appendix C.

3.2 Financial Data

For the asset-pricing analysis, I use firm-level stock returns, portfolio returns, and standard risk factors, complemented by additional control variables.

3.2.1 Stock sample

Firm-level data are obtained from the Sharadar Nasdaq Data Link platform, specifically from the Sharadar Equity Prices (SEP) and Sharadar Core US Fundamentals (SF1) datasets. These databases provide survivorship-bias-free historical prices, delisting returns, and accounting fundamentals for more than 16,000 listed and delisted U.S. firms. The stock sample is composed of the *domestic common stocks*, listed on the NYSE or NASDAQ.

Then, following standard practice in empirical asset pricing (e.g., Engle et al. (2020)), I exclude from the stock sample: (i) penny stocks, defined as firms with a closing price below \$5 at $t-1$, and (ii) microcaps, defined as firms whose market capitalization lies below the 20th percentile of the NYSE market-capitalization distribution at time $t-1$.

After applying these filters and retaining delisted securities with valid return histories, the final stock sample contains 5,125 unique firms over the 2010–2025 period. On average, 2,134 firms have available returns on a given trading day, and 2,141 firms on a given month.

3.2.2 Portfolio sample and control variables

Additional financial data are sourced from the Kenneth R. French Data Library, including the Fama–French three and five factors (Fama and French, 1993, 2015): the market excess return (MKT), the size (SMB), the value (HML), the profitability (RMW), the investment (CMA), and the momentum factor (Carhart, 1997) (MOM). I also use the 25 size–book-to-market portfolios and the 30 and 49 industry portfolios, which serve as test assets in the

Fama-MacBeth regressions.

Because fluctuations in the TRI often coincide with major political and policy events, I include the Economic Policy Uncertainty (EPU) Index of Baker et al. (2016), obtained from the authors' website, to control for broader policy-related uncertainty.

For the predictive regressions, I use standard predictors from the Goyal–Welch dataset (term spread, dividend yield, and default spread; Welch and Goyal, 2008), downloaded from Amit Goyal's website. The VIX and West Texas Intermediate (WTI) crude oil prices are sourced from FRED. Following Maio and Santa-Clara (2012), I proxy changes in the investment opportunity set using the aggregate value-weighted market return from the Kenneth R. French Data Library. All variables are used at a monthly frequency in the predictive regressions.

3.2.3 Correlation

Table 3 reports the daily pairwise correlations between the variables used in the empirical analysis. The TRI, PRI, and EPU enter the analysis in innovations rather than levels, consistent with their use in the asset-pricing tests. Innovations are obtained as residuals from autoregressive processes, with the lag order set to eight based on the Bayesian Information Criterion (BIC). Further details on this procedure are provided in the methodology section 4.1.

The innovations of the TRI and PRI exhibit very low correlations with both the control variables and the EPU innovations, indicating that transition and physical risk shocks are distinct from economic policy uncertainty shocks. The only notable correlations appear among the standard risk factors, as expected.

Taken together, Table 3 reveals no substantial correlation patterns that would compromise the results in the subsequent asset-pricing tests.

Table 3: Correlation Matrix of Daily TRI/PRI Innovations and Financial Variables

Variable	TRI*	PRI*	MKT	SMB	HML	RMW	CMA	MOM	EPU*	WTI	VIX
TRI*	1.00	0.04	-0.02	-0.03	0.01	-0.02	-0.00	0.01	-0.03	-0.01	0.02
PRI*		**1.00**	-0.00	-0.01	0.00	-0.01	0.00	-0.00	0.02	-0.01	-0.01
MKT			**1.00**	**0.22**	-0.07	**-0.20**	**-0.30**	**-0.13**	0.01	**0.12**	**-0.16**
SMB				**1.00**	-0.01	**-0.41**	**-0.14**	**-0.20**	-0.00	0.00	-0.00
HML					1.00	**0.32**	**0.59**	**-0.33**	-0.07	0.08	-0.03
RMW						1.00	**0.30**	-0.03	-0.04	0.03	0.03
CMA							1.00	0.02	-0.04	0.05	0.06
MOM								1.00	0.05	-0.08	-0.00
EPU*									1.00	-0.03	0.08
WTI										**1.00**	-0.09
VIX											**1.00**

Notes: This table reports pairwise correlations among the daily time-series variables from January 3, 2014 to December 31, 2024. Variables marked with * denote innovations, obtained as residuals from a fitted AR(8) process (with the lag order selected using the Bayesian Information Criterion). MKT, SMB, and HML are the Fama–French three factors (Fama and French, 1993), corresponding to the excess market return, the size factor, and the value factor, respectively. RMW and CMA are the profitability and investment factors (Fama and French, 2015). MOM is the momentum factor (Carhart, 1997). EPU denotes the Economic Policy Uncertainty index (Baker et al., 2016). TRI and PRI are the transition and physical risk indicators. WTI corresponds to West Texas Intermediate (WTI) crude oil price returns (Fred St Louis, DCOILWTICO). Coefficients with an absolute value greater than 0.1 are shown in bold.

4 Asset Pricing Empirical Results

4.1 Methodology

The methodology follows four steps to test whether transition / physical risk shocks are priced in the cross-section of stock returns, and whether the TRI and PRI can be interpreted as state variables within the Intertemporal Capital Asset Pricing Model (ICAPM) (Merton, 1973) framework.

4.1.1 Transition and Physical Risk Innovations

Following the ICAPM (Merton, 1973), candidate state variables are used in innovations in the subsequent asset-pricing analysis. Daily innovations of the TRI and PRI are extracted over the sample period from January 1, 2014 to December 31, 2024. For each indicator, an autoregressive model of order eight is estimated, with the lag length selected using the

Bayesian Information Criterion (BIC). Let F_t denote the TRI or PRI:

$$F_t = \alpha + \sum_{p=1}^{8} \phi_p F_{t-p} + \varepsilon_t. \tag{7}$$

The estimated residuals $\hat{\varepsilon}_t$ are interpreted as innovations and used in the asset-pricing analysis. The relatively high autoregressive order reflects the strong persistence of climate-related news coverage. Climate events typically generate sustained media attention over several consecutive days, implying that shocks to textual indicators unfold gradually rather than instantaneously. Using an $AR(8)$ specification therefore allows the innovations to capture unexpected changes in climate risk that are not mechanically driven by short-term media persistence.

In the literature, Engle et al. (2020) estimate innovations using an $AR(1)$ specification. However, their analysis relies on monthly data, for which news shocks exhibit limited persistence and an $AR(1)$ process adequately captures unexpected variation in the indicator. In contrast, I use daily data, for which climate-related news displays substantially higher short-run persistence. Accordingly, a higher-order autoregressive specification is required to properly isolate climate-risk shocks.

4.1.2 Portfolio Construction: TRI and PRI-Sorted Portfolios

The second step of the asset pricing analysis consists in forming portfolios sorted on exposure to shocks in the textual risk indicator. Specifically, firms are sorted according to their estimated sensitivity to innovations in the transition (or physical) risk factor. This approach assesses whether assets with different exposures to unexpected changes in climate risk earn systematically different returns, indicating that the factor captures economically relevant priced risk.

Let $\tilde{F}_t = \hat{\varepsilon}_t$ denote the innovation of the textual factor F (either TRI or PRI), obtained from the autoregressive specification in Equation 7. For each asset i, I estimate the following regression using daily data over a rolling window of 12 months, updated monthly:

$$r_{i,t} - r_{f,t} = c_i + \beta_{i,t}\tilde{F}_t + \boldsymbol{\gamma}_i' \boldsymbol{X}_t + \varepsilon_{i,t}, \tag{8}$$

where $\boldsymbol{X}_t$ denotes a vector of control factors. I consider three specifications: the Fama–French three-factor model (Fama and French, 1993), the Fama–French three-factor model augmented with momentum (Carhart, 1997), and the Fama–French five-factor model (Fama and French, 2015). This procedure yields a time-varying estimate of each asset's exposure $\hat{\beta}_{i,t}$ to innovations in the textual risk factor.

At the end of each month, assets are sorted into quintiles based on their estimated exposure $\hat{\beta}_{i,t-1}$ to the climate risk shock, and value-weighted portfolios are formed accordingly. Quintile 1 (Q1) contains assets with the lowest exposure to the factor, while Quintile 5 (Q5) contains assets with the highest exposure. Assets in Q1 experience negative return responses to climate risk shocks, whereas assets in Q5 tend to respond positively. Accordingly, Q1 can be interpreted as a "brown portfolio" and Q5 as a "green" portfolio. The long–short portfolio Q5–Q1 therefore captures the return differential between assets that benefit from climate risk shocks and those that are adversely affected.

The returns of the quintile portfolios and of the long–short portfolio are then used to estimate abnormal returns (alphas) under several asset pricing models. Specifically, I consider (i) the Fama–French three-factor model (Fama and French, 1993), (ii) the three-factor model augmented with momentum (Carhart, 1997), and (iii) the Fama–French five-factor model (Fama and French, 2015). In all cases, factor loadings are estimated consistently with the first-stage specification.

4.1.3 Fama–MacBeth Estimation Procedure

The third step employs the two-stage Fama and MacBeth (1973) methodology to assess whether innovations in transition and physical climate risk are priced in the cross section of stock returns and to estimate the sign and magnitude of the associated risk premia.

Unlike the beta-sorted portfolio approach, the Fama–MacBeth analysis relies on a set of

diversified test assets rather than individual stocks. Specifically, I use the 25 size–book-to-market portfolios together with the 30 and 49 industry portfolios from the Kenneth R. French Data Library, following Faccini et al. (2023). As shown by Lewellen et al. (2010), augmenting size–value portfolios with industry portfolios improves the reliability of cross-sectional asset pricing tests by increasing the dispersion in factor loadings.

To mitigate errors-in-variables in the estimation of factor exposures, I estimate first-stage betas using a 12-month rolling window of daily data, updated monthly. This choice yields approximately 250 observations per estimation window and substantially reduces measurement error relative to shorter windows, in line with Sautner et al. (2023).

In the second stage, I estimate the following cross-sectional regression:

$$E(r_i) - r_f = \bar{c} + \boldsymbol{\lambda}'_{\tilde{F}} \boldsymbol{\beta}_i^{\tilde{F}} + \boldsymbol{\lambda}'_X \boldsymbol{\beta}_i^X + \bar{\varepsilon}_i, \tag{9}$$

$\boldsymbol{\beta}_i^{\tilde{F}}$ denotes the vector of factor loadings capturing portfolio i's exposure to innovations in the textual climate risk indicators, namely the Transition Risk Index (TRI) and the Physical Risk Index (PRI).

$\boldsymbol{\beta}_i^X$ denotes the vector of factor loadings associated with the control variables for portfolio i. These controls include the Fama–French three factors (Fama and French, 1993)—the excess market return (MKT), the size factor (SMB), and the value factor (HML)—the momentum factor (MOM) (Carhart, 1997), as well as innovations in the Economic Policy Uncertainty (EPU) index (Baker et al., 2016). The EPU innovations are obtained as residuals from an autoregressive model of order eight, selected to minimize the Bayesian Information Criterion (BIC).

4.1.4 ICAPM Testable Implications

Lastly, to assess whether the transition and physical risk indicators can be interpreted as state variables in the ICAPM (Merton, 1973), several empirical implications of the model must be satisfied. First, the climate risk indicators should forecast future investment op-

portunities (Cochrane, 2009; Maio and Santa-Clara, 2012). Second, innovations in these indicators should command a non-zero risk premium in the cross section of asset returns, consistent with the ICAPM (i.e., if a candidate state variable predicts a deterioration in the investment opportunity set, its innovations should be associated with a negative risk premium). Finally, the sign and magnitude of the estimated market price of risk should be economically consistent with its interpretation as a coefficient of risk aversion (Maio and Santa-Clara, 2012).

The latter two criteria can be assessed using the two-stage Fama–MacBeth methodology described above. To evaluate the first implication of the ICAPM—namely, whether the climate risk indicators forecast future investment opportunities—I estimate predictive regressions of aggregate stock market returns, which serve as a proxy for changes in the investment opportunity set. Specifically, I estimate the following regression:

$$r_{t,t+q} = a_q + \beta_q z_t + \varepsilon_{t,t+q}, \tag{10}$$

where $r_{t,t+q} \equiv r_{t+1} + \ldots + r_{t+q}$ the continuously compounded market returns over q periods, $\varepsilon_{t,t+q}$ is the residuals, z_t is the candidate state variable in levels (either the TRI or PRI) at date t. $q \in [1, 3, 12, 24, 36, 48]$ months ahead. Following Maio and Santa-Clara (2012).

To further assess whether the predictive content of the TRI and PRI is distinct from standard macroeconomic predictors, I also estimate the following multiple predictive regression:

$$r_{t,t+q} = a_q + \beta_q^{TRI} TRI_t + \beta_q^{PRI} PRI_t + \boldsymbol{\gamma}_q' \boldsymbol{X_t} + \varepsilon_{t,t+q}, \tag{11}$$

where $\boldsymbol{X_t}$ denotes a vector of control variables commonly used in the return predictability literature (Petkova, 2006; Hahn and Lee, 2006; Welch and Goyal, 2008), including the term spread, the default spread, the market dividend yield (Fama and French, 1988), and the

one-month Treasury bill rate. In addition, I include West Texas Intermediate (WTI) crude oil returns and the VIX index. Including these controls allows me to assess whether the predictive power of the climate-risk indicators is subsumed by well-known macroeconomic or uncertainty-related predictors.

In both regressions, TRI and PRI are included in levels, rather than as innovations.

4.2 Empirical Results

This section reports three main findings. First, portfolios sorted on exposure to transition-risk shocks earn significant abnormal returns, whereas no such pattern emerges for physical risk. Second, Fama–MacBeth regressions show that exposure to transition-risk innovations commands a significantly negative risk premium in the cross section of stock returns, while the physical-risk premium is not statistically different from zero. Third, transition risk predicts future aggregate market returns with a sign consistent with the estimated risk premium, supporting an ICAPM interpretation of the results.

4.2.1 Alphas of TRI and PRI Sorted Portfolios

Table 4 reports the main characteristics and factor exposures of the portfolios sorted on TRI innovation betas. Panel A confirms that the sorting procedure effectively differentiates assets according to their exposure to transition-risk shocks. The estimated TRI innovation betas increase monotonically from the lowest quintile (Q1), which exhibits strongly negative exposure, to the highest quintile (Q5), which displays strongly positive exposure.

Panel A also shows that the number of firms is balanced across quintiles and that there is no monotonic pattern in market capitalization, suggesting that the sorting is not mechanically driven by firm size or sample composition.

Table 4: Properties of Portfolios Sorted on TRI Innovation Betas

	1	2	3	4	5	5-1
	Panel A: Portfolio Characteristics					
Nb. Firms	396.558	396.408	396.083	395.925	395.750	
Return (VW, %)	0.671	1.099	0.840	1.014	1.119	
Market Cap (B$)	13.654	19.965	21.452	17.903	10.866	
TRI Innov. $\hat{\beta}$ (EW)	-0.004	-0.001	0.000	0.001	0.004	
	Panel B: Factor Loadings					
$\hat{\alpha}$	-0.004*	0.000	-0.000	-0.000	0.002	0.006**
	(-1.68)	(0.01)	(-0.37)	(-0.43)	(1.52)	(2.10)
MKT	1.112***	1.051***	0.936***	0.984***	1.078***	-0.034
	(12.58)	(24.16)	(33.56)	(31.71)	(21.46)	(-0.35)
SMB	0.021	-0.091	-0.093**	0.045	0.211***	0.190
	(0.20)	(-1.48)	(-2.09)	(0.87)	(3.66)	(1.41)
HML	-0.335***	0.020	0.042	0.092	-0.089	0.246*
	(-2.74)	(0.52)	(1.54)	(1.56)	(-1.40)	(1.72)
RMW	-0.149	0.014	0.009	0.144**	-0.130	0.019
	(-1.33)	(0.23)	(0.14)	(2.12)	(-1.39)	(0.11)
CMA	0.273	0.009	-0.104**	-0.006	0.057	-0.216
	(1.39)	(0.14)	(-2.01)	(-0.06)	(0.60)	(-0.93)
MOM	-0.143*	0.047	-0.007	0.063*	0.011	0.154
	(-1.85)	(0.83)	(-0.27)	(1.72)	(0.19)	(1.33)

Notes: Each month, stocks are sorted into five value-weighted portfolios (quintiles) based on their estimated exposure to innovations in the TRI. Quintile 1 (1) contains the lowest TRI innovation betas, and Quintile 5 (5) the highest. TRI innovation betas are estimated using daily excess returns and a 12-month rolling window. Panel A reports time-series averages (January 2015 – December 2024) of portfolio characteristics, including the average number of firms (Nb. Firms), the average value-weighted monthly return (in percent), the average market capitalization (in billions of USD), and the average equal-weighted TRI innovation beta. Panel B reports factor loadings from time-series regressions of portfolio excess returns on the Fama–French five factors (Fama and French, 2015) augmented with momentum (Carhart, 1997), for each quintile and for the long–short portfolio (5–1). *t*-statistics (in parentheses) are computed using Newey–West (Newey and West, 1987) standard errors with 6 lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Panel B reports the factor loadings of the TRI-sorted portfolios with respect to standard asset-pricing factors. The long–short portfolio (Q5–Q1) exhibits a market beta close to zero and generally limited exposure to the Fama–French size, value, profitability, investment (Fama and French, 1993, 2015) and Carhart momentum factors (Carhart, 1997). While some

loadings are statistically different from zero, their magnitudes remain small. Overall, this suggests that the TRI-sorted strategy is not primarily driven by conventional risk factors. Consequently, the abnormal performance documented in Table 5 is unlikely to be fully explained by standard factor exposures, and is instead consistent with compensation related to transition risks shocks.

Table 5: Alphas of Value-Weighted Portfolios Sorted on Transition-Risk Innovation Betas

	1	2	3	4	5	5-1
	January 2015 – December 2024 ***(120 months)***					
Fama-French 3-factor alpha	-0.45*	0.04	-0.06	0.06	0.10	0.55**
	(-1.83)	(0.43)	(-0.79)	(0.63)	(0.83)	(2.04)
Carhart 4-factor alpha	-0.39**	0.09	-0.13	0.07	0.07	0.47*
	(-1.96)	(1.07)	(-1.52)	(0.78)	(0.56)	(1.92)
Fama-French 5-factor alpha	-0.47**	-0.02	0.09	-0.10	0.23*	0.70***
	(-1.98)	(-0.24)	(1.38)	(-1.05)	(1.80)	(2.69)
	November 2016 – December 2018 ***(26 months)***					
Fama-French 3-factor alpha	0.20	0.25	-0.06	-0.00	-0.30**	-0.51
	(0.58)	(1.21)	(-0.62)	(-0.03)	(-2.25)	(-1.34)
Carhart 4-factor alpha	0.01	0.27**	-0.07	0.01	-0.32**	-0.33
	(0.07)	(2.52)	(-0.63)	(0.17)	(-2.07)	(-1.37)
Fama-French 5-factor alpha	0.01	0.17	0.05	0.05	-0.46***	-0.47*
	(0.06)	(0.85)	(0.40)	(0.62)	(-3.67)	(-1.90)

Notes: Each month, stocks are sorted into five value-weighted portfolios (quintiles) based on their transition-risk (TRI) innovation betas (1 = lowest-beta quintile, 5 = highest). TRI innovation betas are estimated for each stock using a 12-month rolling window of daily returns, together with the appropriate control variables for each specification. Portfolio alphas are computed from monthly returns using several asset-pricing models: the Fama–French three-factor model (Fama and French, 1993), the Carhart four-factor model (Carhart, 1997), and the Fama–French five-factor model (Fama and French, 2015). Alphas are reported in monthly percentage terms. t-statistics (in parentheses) are based on Newey–West heteroskedasticity- and autocorrelation-consistent standard errors (Newey and West, 1987) with a lag length of 6. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 5 reports the estimated alphas of the quintile portfolios and the corresponding long–short portfolio formed on exposure to transition-risk (TRI) innovations. Alphas are estimated using the Fama–French three-factor model (Fama and French, 1993), the Fama–French five-factor model (Fama and French, 2015), and the momentum-augmented Fama–French

three-factor model (Carhart, 1997).

The results reveal a monotonically increasing pattern in estimated alphas across TRI innovations-beta sorted portfolios. The strategy that longs the high-exposure quintile and shorts the low-exposure quintile yields a spread alpha that is statistically significant across all three specifications, ranging from 0.47% ($t = 1.92$) under the Carhart four-factor model to 0.70% ($t = 2.69$) under the Fama-French five-factor model. Interestingly, the alpha is predominantly driven by the short leg (Q1), which exhibits significant underperformance. This suggests that stocks with negative exposure to transition risk shocks have been subject to significant downward price adjustments over the 2015-2024 period.

This period is characterized by frequent increases in transition risk shocks associated with major climate policy developments, such as COP21 agreements, the Inflation Reduction Act. Over this sample, the transition risk indicator (TRI) therefore predominantly reflects a *bad state* of the world for firms subject to environmental regulation. Consistent with this interpretation, assets that are negatively exposed to TRI innovations (Q1 portfolios) exhibit negative and significant abnormal returns, indicating that they provide poor hedging properties against increases in transition risk and therefore underperform during periods in which transition risk shocks materialize frequently. These findings are consistent with the equilibrium mechanism in Pástor et al. (2021), in which brown assets underperform green assets during periods of heightened climate risk shocks.

By contrast, assets with positive exposure to TRI innovations (Q5 portfolios) tend to perform better during such episodes, reflecting their role as partial hedges against transition risk shocks. The fact that Q5 alphas are statistically significant only under the Fama–French five-factor specification suggests that this hedging channel is present but not dominant, and that part of the performance is absorbed by standard risk factors.

Moreover, the period from November 2016 to December 2018 (corresponding to the Trump administration and a Republican majority in the U.S. House of Representatives) is characterized by a temporary easing of U.S. climate policy. During this period, the tran-

sition risk indicator (TRI) predominantly reflects a *good state* of the world, with declining transition risk. Consistent with the ICAPM, the sign of abnormal returns tends to reverse during this period: the high-exposure portfolio (Q5) exhibits negative abnormal returns. This pattern reflects the state-contingent realization of returns for assets that hedge against increases in transition risk, rather than a reversal of the underlying risk premium. While these patterns align with the intertemporal hedging mechanism, this sub-period analysis relies on a limited number of observations (26 months) and should therefore be interpreted as illustrative rather than conclusive.

Taken together, the full-sample and sub-period evidence supports a state-dependent interpretation of transition risk, whereby the sign of abnormal returns depends on whether transition risk represents a *bad* or a *good* state of the world, for polluting firms (i.e. deterioration or an improvement in investment opportunities). Over the full sample, however, the TRI predominantly reflects a bad state of the world.

Importantly, these abnormal returns are consistent with the negative risk premium associated with exposure to transition-risk innovations in the Fama–MacBeth regressions (subsubsection 4.2.2). In an intertemporal hedging framework, assets that covary positively with a bad state variable (i.e., one that signals deteriorating investment opportunities) command lower expected returns (Cochrane, 2009), even though they may deliver positive abnormal returns during periods in which adverse states materialize more frequently.

Regarding the physical risk shocks, Table 6 reports the estimated alphas of the different PRI innovations sorted portfolios and the long-short portfolio. PRI shocks do not generate significant abnormal return for the quintiles and the long-short portfolio. These results are consistent with the existing literature (Faccini et al., 2023).

Table 6: Alphas of Value-Weighted Portfolios Sorted on Physical-Risk Innovation Betas

	1	2	3	4	5	5-1
	January 2015 – December 2024 ***(120 months)***					
Fama-French 3-factor alpha	-0.57**	0.02	-0.03	0.24**	-0.09	0.48
	(-2.13)	(0.25)	(-0.30)	(2.49)	(-0.64)	(1.40)
Carhart 4-factor alpha	-0.59**	-0.06	-0.00	0.24***	-0.02	0.57
	(-2.15)	(-0.53)	(-0.04)	(3.04)	(-0.10)	(1.64)
Fama-French 5-factor alpha	-0.59**	0.03	0.07	0.14	-0.07	0.52
	(-2.14)	(0.27)	(0.76)	(1.56)	(-0.49)	(1.50)

Notes: Each month, stocks are sorted into five value-weighted portfolios (quintiles) based on their physical-risk (PRI) innovation betas (1 = lowest-beta quintile, 5 = highest). PRI innovation betas are estimated for each stock using a 12-month rolling window of daily returns, together with the appropriate control variables for each specification. Portfolio alphas are computed from monthly returns using several asset-pricing models: the Fama–French three-factor model (Fama and French, 1993), the Carhart four-factor model (Carhart, 1997), and the Fama–French five-factor model (Fama and French, 2015). Alphas are reported in monthly percentage terms. t-statistics (in parentheses) are based on Newey–West heteroskedasticity- and autocorrelation-consistent standard errors (Newey and West, 1987) with a lag length of 6. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

4.2.2 Pricing of Transition and Physical Risk: Fama–MacBeth Evidence

This section reports Fama–MacBeth estimates of the risk premia associated with exposure to transition and physical risk innovations. Table 7 presents the results along with standard asset-pricing controls.

I find a negative and statistically significant risk premium for exposure to TRI innovations: portfolios with higher TRI betas earn significantly lower expected excess returns. This result holds across both asset universes, with estimated premia ranging from -1.61 (t = -1.94) to -1.89 (t = -2.25) depending on the specification. The risk premium is estimated more precisely when additionally controlling for PRI innovations. The weaker significance in the 74-portfolio universe may reflect noisier first-stage TRI beta estimates at a 12-month horizon, as this set includes 49 industry portfolios with more idiosyncratic return variation. Consistent with an errors-in-variables interpretation, Appendix D shows that using a 24-month rolling window–thereby improving the precision of first-stage beta estimates–yields second-pass premia that are statistically significant at the 5% level across all specifications;

results also remain robust across alternative sets of control variables.

Table 7: Fama–MacBeth Risk-Premium Estimates for Transition and Physical Risk Indicator Innovations

	74 Portfolios			55 Portfolios		
	(*i*)	(*ii*)	(*iii*)	(*i*)	(*ii*)	(*iii*)
TRI Innov.	-1.339		-1.612*	-1.823**		-1.893**
	(-1.59)		(-1.94)	(-2.00)		(-2.25)
PRI Innov.		-0.624	-0.500		-0.848	-0.757
		(-0.63)	(-0.51)		(-0.74)	(-0.70)
MKT	0.010	0.009	0.008	0.014**	0.010	0.012*
	(1.37)	(1.18)	(1.13)	(2.33)	(1.57)	(1.77)
SMB	0.000	0.000	0.000	-0.000	-0.000	-0.000
	(0.03)	(0.09)	(0.03)	(-0.11)	(-0.07)	(-0.13)
HML	-0.001	-0.001	-0.001	-0.001	-0.002	-0.001
	(-0.20)	(-0.15)	(-0.16)	(-0.34)	(-0.35)	(-0.35)
MOM	0.007	0.007	0.006	0.008	0.007	0.006
	(1.20)	(1.13)	(1.11)	(1.28)	(1.04)	(1.02)
EPU Innov.	-74.830	-87.240	-84.020	-144.407	-105.725	-137.216
	(-0.79)	(-0.96)	(-0.91)	(-1.38)	(-1.08)	(-1.41)
Sample Period	December 2014 – November 2024					
Observations	8,880	8,880	8,880	6,600	6,600	6,600

Notes: This table reports risk premium estimates from a Fama–MacBeth regression (Fama and MacBeth, 1973) over the period December–2014 to November–2024. The first-stage regressions are estimated on the period January–2014 to December–2024 using daily data and a 12-month rolling window that advances by one month. The same control variables as in the second stage regressions are included. I consider two asset universes. The *74 portfolios* correspond to the 49 industry portfolios and the 25 Size–Book-to-Market portfolios (5×5). The *55 portfolios* correspond to the same 25 Size–BM portfolios and a subset of 30 industry portfolios. Control variables include the Fama–French three factors (*MKT*, *SMB*, *HML*) (Fama and French, 1993), the momentum factor (*MOM*) (Carhart, 1997), and the innovations of the Economic Policy Uncertainty index (Baker et al., 2016), extracted as residuals from an AR(8) model selected by the BIC criterion. Specifications *(i)* and *(ii)* include one climate beta at a time—either the transition-risk innovation beta or the physical-risk innovation beta—whereas specification *(iii)* includes both innovation-based climate betas jointly. t-statistics are computed using Newey–West standard errors (Newey and West, 1987) with a lag of 6. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

In contrast, exposure to PRI innovations is not significantly priced in any specification, a result that remains unchanged across all robustness checks reported in Appendix D.

These findings support an intertemporal hedging interpretation. Portfolios whose re-

turns covary positively with TRI innovations provide a hedge against transition-risk shocks and therefore command lower expected excess returns, as investors are willing to pay to hold assets that perform relatively well when transition risk materializes. This evidence is consistent with a negative price of transition risk, as implied by the equilibrium model of Barnett (2023), and with the predictions of Pástor et al. (2021), in which assets that hedge climate-related risks earn lower expected returns in equilibrium.

In addition, the estimated annual risk premium associated with market exposure is positive and equals approximately 10.8% and 12% for the 74- and 55-portfolio test assets, respectively. Following Maio and Santa-Clara (2012), I assess the economic plausibility of this estimate by scaling it by the variance of the market excess return. The implied coefficient of relative risk aversion ranges from about 3.8 to 6.7, which is economically reasonable and consistent with the third ICAPM restriction discussed in Maio and Santa-Clara (2012).

Taken together, the economically plausible implied risk aversion coefficient and the significantly negative price of risk for exposure to TRI innovations support the internal consistency of an ICAPM-based interpretation of the results.

Next, I examine whether TRI and PRI innovations predict aggregate market returns, and whether the sign of this predictability is consistent with the sign of the estimated risk premia, thereby addressing the remaining ICAPM restrictions discussed in Maio and Santa-Clara (2012).

4.2.3 ICAPM Restrictions and Market Return Predictability

Lastly, I find that the Transition and Physical Risk Indicators in levels predict negative aggregate market returns at different horizons, even after controlling for standard predictors used in the return predictability literature.

Table 8 reports predictive regressions using the transition and physical risk indicators without additional control variables. The TRI significantly predicts negative aggregate market returns at the 48-month horizon, while the PRI exhibits predictive power at the

one-month horizon.

When standard control variables are included, the predictive coefficients become more precisely estimated and retain the same negative sign, as reported in Table 9. In this specification, the TRI predicts negative aggregate market returns at the 12 and 48 month horizons. For the PRI, predictive power is concentrated at the one-month horizon and, in some specifications, at the 48-month horizons.

Table 8: Single Predictive Regressions of Aggregate Market Returns on Transition and Physical Risk Indices

$q =$	1	3	12	24	36	48
TRI	-1.085	-1.798	-6.526	2.235	-3.996	-8.145***
t-stat$_{\text{NW}}$	(-1.47)	(-1.07)	(-0.88)	(0.15)	(-0.69)	(-2.92)
t-stat$_{\text{HH}}$	(-1.39)	(-1.20)	(-1.06)	(0.27)	(-0.73)	(-1.98)
Adj. R^2 (%)	-0.03	0.14	2.30	-0.62	1.33	7.78
PRI	-2.532**	-3.594	1.669	9.972	-1.768	5.397
t-stat$_{\text{NW}}$	(-2.54)	(-1.48)	(0.26)	(1.42)	(-0.43)	(0.73)
t-stat$_{\text{HH}}$	(-2.53)	(-1.48)	(0.29)	(1.89)	(-0.40)	(0.80)
Adj. R^2 (%)	3.84	3.47	-0.61	5.12	-0.66	2.28

Notes: This table reports single-predictive regressions of the value-weighted U.S. market return compounded over horizons of 1, 3, 12, 24, 36, and 48-month on the current level of the Transition Risk Indicator (TRI) or the Physical Risk Indicator (PRI). The sample runs from January 2014 to December 2024. For each horizon q, the first q months of data are lost by construction. Newey and West (1987) and Hansen and Hodrick (1980) q-lag t-statistics are reported below the coefficient estimates. The estimates are in percentage. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

The difference in predictive horizons between TRI and PRI is consistent with the distinct nature of these risks. Physical risk shocks—such as hurricanes, wildfires, or extreme weather events—materialize suddenly and affect asset prices almost immediately, explaining why PRI predicts returns at the one-month horizon. In contrast, transition risk operates through slower-moving channels: regulatory announcements, policy proposals, and legislative processes unfold over multiple years before fully affecting firm cash flows and discount rates. For instance, a carbon tax proposal must be debated, passed, and implemented before its economic impact is fully realized. This delay explains why TRI exhibits predictive power

at longer horizons (12 and 48 months).

Taken together, the evidence shows that the transition risk indicator (TRI) predicts negative aggregate market returns and that its innovations carry a negative and statistically significant risk premium, alongside an economically sensible market risk price. These findings are consistent with the ICAPM restrictions discussed in Maio and Santa-Clara (2012).

Because the TRI forecasts lower future market returns, investors have an incentive to hedge against increases in transition risk by holding assets whose returns covary positively with TRI innovations. As a result, such assets command lower expected excess returns, implying a negative price of transition risk (Table 7).

When transition risk shocks materialize, these same assets deliver positive abnormal returns (Table 5), consistent with their role as hedging assets. Overall, this evidence supports the interpretation of transition risk as a relevant ICAPM state variable.

Moreover, Appendix E shows that removing the directional component from the TRI and PRI (yielding the attention to transition risk (AT) and attention to physical risk (AP) indicators) breaks the ICAPM sign-consistency between cross-sectional pricing and aggregate market return predictability. In particular, once direction is removed, the TRI in levels no longer forecasts aggregate market returns with a sign that is consistent with the estimated cross-sectional price of risk. This evidence underscores that accounting for risk direction is essential for identifying an intertemporal coherent transition-risk premium.

Table 9: Multiple Predictive Regressions of Aggregate Market Returns

	TRI	PRI	TERM	DEF	DY	RF	WTI	VIX	Adj. R^2 (%)
Panel A: $q = 1$									
(i)	-0.005	-0.015***	-0.016*	0.003	-0.000	-0.005			6.49
	(-1.21)	(-3.75)	(-1.79)	(0.53)	(-0.03)	(-0.63)			
	[-1.15]	[-3.80]	[-1.86]	[0.51]	[-0.03]	[-0.66]			
(ii)	-0.003	-0.012***	-0.010	-0.009	0.011	0.006		0.015**	10.08
	(-0.64)	(-2.71)	(-1.08)	(-1.16)	(1.34)	(0.70)		(2.47)	
	[-0.61]	[-2.76]	[-1.12]	[-1.13]	[1.31]	[0.72]		[2.55]	
(iii)	-0.005	-0.015***	-0.018*	0.003	-0.001	-0.007	-0.005		6.91
	(-1.24)	(-3.81)	(-1.89)	(0.52)	(-0.10)	(-0.88)	(-1.22)		
	[-1.17]	[-3.83]	[-1.98]	[0.50]	[-0.09]	[-0.93]	[-1.33]		
(iv)	-0.003	-0.012***	-0.011	-0.008	0.010	0.004	-0.002	0.014**	9.56
	(-0.69)	(-2.78)	(-1.15)	(-1.06)	(1.26)	(0.47)	(-0.65)	(2.39)	
	[-0.66]	[-2.81]	[-1.19]	[-1.01]	[1.21]	[0.48]	[-0.65]	[2.40]	
Panel B: $q = 12$									
(i)	-0.036***	-0.011	-0.091**	0.038***	0.050**	-0.021			50.54
	(-3.63)	(-0.90)	(-2.10)	(3.46)	(2.44)	(-0.44)			
	[-3.59]	[-0.84]	[-1.90]	[3.46]	[2.37]	[-0.40]			
(ii)	-0.028***	0.002	-0.064*	-0.009	0.093***	0.024		0.058***	57.57
	(-3.35)	(0.15)	(-1.85)	(-0.59)	(4.00)	(0.60)		(3.52)	
	[-3.45]	[0.14]	[-1.72]	[-0.57]	[3.82]	[0.58]		[3.27]	
(iii)	-0.036***	-0.012	-0.097**	0.038***	0.049**	-0.028	-0.014*		51.22
	(-3.52)	(-0.98)	(-2.15)	(3.28)	(2.37)	(-0.57)	(-1.87)		
	[-3.40]	[-0.91]	[-1.94]	[3.17]	[2.31]	[-0.51]	[-1.98]		
(iv)	-0.028***	0.001	-0.066*	-0.008	0.092***	0.022	-0.003	0.056***	57.23
	(-3.32)	(0.10)	(-1.85)	(-0.47)	(3.80)	(0.52)	(-0.37)	(3.12)	
	[-3.37]	[0.09]	[-1.72]	[-0.44]	[3.56]	[0.50]	[-0.35]	[2.85]	
Panel C: $q = 48$									
(i)	-0.034***	0.031	-0.078***	-0.021	0.065***	-0.139***			25.78
	(-5.77)	(1.11)	(-3.84)	(-1.16)	(3.02)	(-6.31)			
	[-2.99]	[1.47]	[-2.47]	[-0.85]	[1.55]	[-3.88]			
(ii)	-0.029***	0.045*	-0.057***	-0.055***	0.106***	-0.117***		0.042***	32.02
	(-5.69)	(1.80)	(-2.98)	(-2.60)	(3.96)	(-4.71)		(2.87)	
	[-2.52]	[2.43]	[-1.70]	[-1.97]	[2.33]	[-3.03]		[1.78]	
(iii)	-0.034***	0.029	-0.086***	-0.023	0.069***	-0.151***	-0.012		26.64
	(-5.78)	(1.01)	(-3.61)	(-1.27)	(3.23)	(-5.92)	(-1.63)		
	[-3.11]	[1.37]	[-2.61]	[-0.97]	[1.73]	[-4.12]	[-1.05]		
(iv)	-0.029***	0.044*	-0.061***	-0.054***	0.105***	-0.121***	-0.004	0.040***	31.26
	(-5.86)	(1.65)	(-2.76)	(-2.80)	(4.18)	(-4.41)	(-0.69)	(2.96)	
	[-2.55]	[2.21]	[-1.83]	[-1.89]	[2.30]	[-3.12]	[-0.43]	[1.69]	

Notes: This table reports multiple-predictive regressions of the value-weighted U.S. market return compounded over horizons $q \in \{1, 12, 48\}$ months on the TRI and PRI in levels, with control variables: term spread (TERM), default spread (DEF), dividend yield (DY), risk-free rate (RF), VIX, and WTI crude oil returns. All regressors are standardized. Sample: January 2014 – November 2024. The first q months are lost by construction. Newey and West (1987) and Hansen and Hodrick (1980) q-lag t-statistics reported in parentheses and brackets, respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

5 Conclusion

This paper provides new evidence on the pricing of climate transition risk in equity markets. Using novel directional transition and physical risk indicators constructed with large language models, I show that exposure to transition-risk innovations commands a negative and statistically significant risk premium in the cross section of U.S. stock returns, whereas exposure to physical-risk innovations is not priced. By capturing both increasing and decreasing risk information, my measure resolves the identification ambiguity in prior unidirectional approaches.

The empirical evidence—from portfolio sorts, Fama–MacBeth regressions, and aggregate market return predictability—is jointly consistent with an ICAPM interpretation. Transition risk forecasts deteriorations in future investment opportunities, leading investors to hedge against increases in transition risk by holding assets whose returns covary positively with transition-risk innovations. As a result, these assets have lower expected excess returns, yet earn positive abnormal returns when transition-risk shocks materialize. Thus, transition risk can be interpreted as a state variable within an ICAPM framework.

These findings align with the equilibrium predictions of climate asset pricing models such as Pástor et al. (2021) and Barnett (2023), and satisfy the ICAPM restrictions discussed in Maio and Santa-Clara (2012). More broadly, the results highlight the importance of accounting for the direction, horizon, and intensity of climate-related risks when studying their asset pricing implications. Importantly, removing the directional component breaks the ICAPM sign-consistency between cross-sectional pricing and aggregate market return predictability.

The absence of a significant physical risk premium should be interpreted with caution. Unlike transition risk, physical climate risk has a strong spatial and localized dimension, which may not be fully captured by aggregate textual indicators constructed from national newspaper coverage. As a result, the physical risk indicator used in this paper may attenuate cross-sectional variation in firm-level exposure to climate hazards.

Consistent with this interpretation, prior studies that rely on granular weather and spatial data—rather than media-based measures—document significant effects of physical climate risk on asset prices (e.g., Acharya et al. (2022); Braun et al. (2025)). Therefore, the insignificant physical risk premium found in this paper may reflect limitations in measurement rather than the absence of pricing per se. Exploring alternative proxies for physical risk that exploit high-frequency and geographically disaggregated data (as in Tankov and Tantet (2019)) represents a promising avenue for future research.

Data Availability Statement

The climate risk indicators constructed in this paper are available from the author upon request. Firm-level data are from Sharadar Nasdaq Data Link. Fama-French factors are from Kenneth French's Data Library. News articles are retrieved from MediaCloud; full text is subject to copyright restrictions.

References

Acharya, Viral V., Timothy Johnson, Suresh Sundaresan, and Tuomas Tomunen, 2022, Is physical climate risk priced? evidence from regional variation in exposure to heat stress, NBER Working Paper No. 30445.

Ardia, David, Keven Bluteau, Kris Boudt, and Koen Inghelbrecht, 2023, Climate Change Concerns and the Performance of Green vs. Brown Stocks, *Management Science* 69, 7607–7632.

Bail, Christopher A, 2024, Can generative ai improve social science?, *Proceedings of the National Academy of Sciences* 121, e2314021121.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *The Quarterly Journal of Economics* 131, 1593–1636.

Baldauf, Markus, Lorenzo Garlappi, and Constantine Yannelis, 2020, Does climate change affect real estate prices? only if you believe in it, *The Review of Financial Studies* 33, 1256–1295.

Barnett, Michael, 2023, Climate change and uncertainty: An asset pricing perspective, *Management Science* 69, 7562–7584.

Bessec, Marie, and Julien Fouquau, 2024, A Green Wave in Media: A Change of Tack in Stock Markets, *Oxford Bulletin of Economics and Statistics* 86, 1026–1057.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan, 2003, Latent dirichlet allocation, *Journal of Machine Learning Research* 3, 993–1022.

Bolton, Patrick, and Marcin Kacperczyk, 2021, Do investors care about carbon risk?, *Journal of Financial Economics* 142, 517–549.

Bolton, Patrick, and Marcin Kacperczyk, 2023, Global pricing of carbon-transition risk, *The Journal of Finance* 78, 3677–3754.

Braun, Alexander, Julia Braun, and Florian Weigert, 2025, Extreme-weather risk and the cross-section of stock returns, *Journal of Risk and Insurance* .

Bua, Giovanna, Daniel Kapp, Federico Ramella, and Lavinia Rognone, 2024, Transition versus physical climate risk pricing in European financial markets: a text-based approach, *The European Journal of Finance* 30, 2076–2110.

Bybee, Leland, Bryan Kelly, and Yinan Su, 2023, Narrative asset pricing: Interpretable systematic risk factors from news text, *The Review of Financial Studies* 36, 4759–4787.

Campbell, John Y., 1996, Understanding risk and return, *Journal of Political Economy* .

Carhart, Mark M., 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Chen, Yifei, Bryan T. Kelly, and Dacheng Xiu, 2022, Expected returns and large language models, Working Paper 4416687, SSRN.

Cochrane, John H., 2009, *Asset Pricing* (Princeton University Press).

Cook, Thomas R., Sophia Kazinnik, Anne Lundgaard Hansen, and Peter McAdam, 2023, Evaluating local language models: An application to financial earnings calls, Research Working Paper RWP 23-12, Federal Reserve Bank of Kansas City.

De Nard, Gianluca, Robert F. Engle, and Bryan Kelly, 2024, Factor-Mimicking Portfolios for Climate Risk, *Financial Analysts Journal* 80, 37–58.

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2020, Hedging Climate Change News, *The Review of Financial Studies* 33, 1184–1216.

Faccini, Renato, Rastin Matin, and George Skiadopoulos, 2023, Dissecting climate risks: Are they reflected in stock prices?, *Journal of Banking & Finance* 155, 106948.

Fama, Eugene F., 1991, Efficient capital markets: II, *The Journal of Finance* 46, 1575–1617.

Fama, Eugene F, and Kenneth R French, 1988, Dividend yields and expected stock returns, *Journal of financial economics* 22, 3–25.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and Kenneth R. French, 2015, A five-factor asset pricing model, *Journal of Financial Economics* 116, 1–22.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Gavriilidis, Konstantinos, 2021, Measuring climate policy uncertainty, SSRN Working Paper.

Giglio, Stefano, Bryan Kelly, and Johannes Stroebel, 2021, Climate Finance, *Annual Review of Financial Economics* 13, 15–36.

Gilardi, Fabrizio, Meysam Alizadeh, and Maël Kubli, 2023, ChatGPT outperforms crowd workers for text-annotation tasks, *Proceedings of the National Academy of Sciences* 120, e2305016120.

Grootendorst, Maarten, 2022, BERTopic: Neural topic modeling with a class-based TF-IDF procedure, arXiv:2203.05794.

Hahn, Jaehoon, and Hangyong Lee, 2006, Yield spreads as alternative risk factors for size and book-to-market, *Journal of Financial and Quantitative Analysis* 41, 245–269.

Hansen, Lars Peter, and Robert J. Hodrick, 1980, Forward exchange rates as optimal predictors of future spot rates: An econometric analysis, *Journal of Political Economy* 88, 829–853.

He, Songrun, Linying Lv, Asaf Manela, and Jimmy Wu, 2025, Chronologically consistent large language models, Preprint 2502.21206, arXiv.

Jha, Manish, Jialin Qian, Michael Weber, and Baozhong Yang, 2024, Chatgpt and corporate policies, NBER Working Paper No. 32161.

Krueger, Philipp, Zacharias Sautner, and Laura T. Starks, 2020, The importance of climate risks for institutional investors, *The Review of Financial Studies* 33, 1067–1111.

Lewellen, Jonathan, Stefan Nagel, and Jay Shanken, 2010, A skeptical appraisal of asset pricing tests, *Journal of Financial Economics* 96, 175–194.

Lloyd, S., 1982, Least squares quantization in PCM, *IEEE Transactions on Information Theory* 28, 129–137.

Maio, Paulo, and Pedro Santa-Clara, 2012, Multifactor models and their consistency with the ICAPM, *Journal of Financial Economics* 106, 586–613.

Mellon, Jonathan, Jack Bailey, Ralph Scott, James Breckwoldt, Marta Miori, and Phillip Schmedeman, 2024, Do AIs know what the most important issue is? using language models to code open-text social survey responses at scale, *Research & Politics* 11, 20531680241231468.

Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.

Newey, Whitney K., and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Petkova, Ralitsa, 2006, Do the Fama–French factors proxy for innovations in predictive variables?, *The Journal of Finance* 61, 581–612.

Pástor, L, Robert F. Stambaugh, and Lucian A. Taylor, 2021, Sustainable investing in equilibrium, *Journal of Financial Economics* 142, 550–571.

Pástor, L, Robert F. Stambaugh, and Lucian A. Taylor, 2022, Dissecting green returns, *Journal of Financial Economics* 146, 403–424.

Sautner, Zacharias, Laurence van Lent, Grigory Vilkov, and Ruishen Zhang, 2023, Pricing climate change exposure, *Management Science* 69, 7540–7561.

Tankov, Peter, and Alexis Tantet, 2019, Climate data for physical risk assessment in finance, SSRN Working Paper.

Treepongkaruna, Sirimon, Kam Fong Chan, and Ihtisham Malik, 2023, Climate policy uncertainty and the cross-section of stock returns, *Finance Research Letters* 55, 103837.

Welch, Ivo, and Amit Goyal, 2008, A comprehensive look at the empirical performance of equity premium prediction, *The Review of Financial Studies* 21, 1455–1508.

Zhang, Shaojun, 2025, Carbon Returns across the Globe, *The Journal of Finance* 80, 615–645.

Zhang, SY, 2022, Are investors sensitive to climate-related transition and physical risks? evidence from global stock markets. res. int. bus. finance, 62, 101710.

Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang, 2024, Can large language models transform computational social science?, arXiv:2305.03514.

A Prompt

The full prompt used in the construction of the Transition and Physical Risk Indicators is provided below.

You will be asked a sequence of questions about the following article. Always output a complete JSON object with all the fields. If you answered "No" to the first question, then fill the other fields with "N/A" or [].

Article: `{article_text}`

Question 1: Is the following article related to climate change?

Definitions:

- Answer "Yes" if climate change, global warming, greenhouse gas emissions, environmental or energy transition policies, extreme weather events, adaptation or mitigation strategies, or their social/economic consequences are a substantial part of the article. This includes cases where climate change is not the only topic, but is discussed in detail or plays a significant role in the narrative.
- Answer "No" only if climate change is absent, or if it is mentioned only briefly.

Rules:

- Use "No" only when climate change is minor, anecdotal, or not part of the article's real focus.

```
{{ "is_climate_related":  "Yes" | "No" }}
```

Question 2: What type of climate risk does it mainly refer to?

Definitions:

- "Transition": risks linked to the transition toward a low-carbon economy (e.g., new regulations, carbon taxes).
- "Physical": risks linked to the physical impacts of climate change (e.g., the impact of extreme weather events caused by climate change). For instance, an article discussing the economic consequences of a natural disaster caused by climate change will be classified as "Physical" risk.
- "Both": the article clearly includes both transition and physical risks equally.
- "Unclear": climate change is discussed, but the risk type cannot be determined.

Rules:

- If one risk type is clearly dominant in the narrative, choose that category (e.g., if transition risks are mentioned much more often than physical risks, choose Transition).
- If both risk types are discussed with comparable importance, choose "Both".
- If the text does not allow classification, choose "Unclear".

```
{{ "risk_type":  "Transition" | "Physical" | "Both" | "Unclear" }}
```

Question 3: From the general point of view of a polluting firm reasoning in the short run and at its own level (not macro), does the article suggest that the risk is increasing, decreasing, or remaining neutral?

Examples for an increase:

- A carbon tax is implemented, the risk for a firm increases as it has to pay a tax on its carbon emissions.

- Rising global temperatures are expected to increase the frequency of hurricanes, raising physical risks for coastal infrastructure.

Examples for a decrease:

- Lawmakers voted down stricter emission standards, easing transition risks for the automotive sector.
- The EPA's power is weakened, easing transition risks for firms.
- The EU delayed the implementation of stricter climate regulations, reducing immediate transition risks for manufacturers.
- The development of drought-resistant crops lowers physical risks for farmers facing hotter, drier summers.

```
{{ "risk_direction":  "Increase" | "Decrease" | "Neutral" }}
```

Question 4: If you have not answered "Neutral" previously, what is the intensity of the risk described in the article?

Definitions:

- "Weak": The measure or event has a limited geographic scope (local, e.g., a city or county) and the quantified impacts are minor. It may also be exploratory, preliminary, or non-binding (e.g., an announcement, consultation, or draft policy).
- "Medium": The measure or event has a regional scope (e.g., at the state level in the United States) or is sector-specific, but not national or global. Quantified impacts are moderate. It may be serious in intent but not fully binding (e.g., partial commitments, directives).
- "Strong": The measure or event has a national or international scope, or involves major actors (e.g., the EU, US, China, G7). Quantified impacts are significant. It is binding, or categorical (e.g., a "ban," "mandatory phase-out," or "emergency declared").

- "N/A": Only use N/A if the previous answer was Neutral. When in doubt, choose the closest among Weak, Medium, or Strong.

```
{{ "intensity":  "Weak" | "Medium" | "Strong" | "N/A" }}
```

Question 5: What is the time horizon of the identified climate risk in the article?

Definitions:

- "Short": The risk is expected to materialize within 0–3 years, the language often refers to current events, near-term shocks, or upcoming policies.
- "Medium": The risk is expected to occur in the 3–10 year range, impacts are not immediate but foreseeable and already under discussion.
- "Long": The risk unfolds in a 10+ year timeframe, often linked to structural changes (technology shifts, sea-level rise, global warming trajectories).
- "N/A": Use if the article does not specify any clear timeframe.

```
{{ "horizon":  "Short" | "Medium" | "Long" | "N/A" }}
```

B LLM Robustness Checks

To assess the robustness of the paper's LLM-based approach, I proceed in two steps. First, I evaluate three alternative prompt designs on a manually annotated subset of the newspaper corpus, and I compare two LLMs (*GPT-5-mini* and *GPT-5-nano*) to select the configuration with the best classification performance. Second, to probe potential data-contamination concerns—i.e., the possibility that the model leverages memorized information from its training data rather than the provided input text—I evaluate performance on articles published after the model's training cut-off date.

Overall, performance varies meaningfully across prompt designs, which is less consistent with a pure memorization explanation and suggests an important role for how the model processes the input under different instructions. Table 12 confirms the absence of data contamination.

The prompt selected for the analysis is P.2 (reported in Appendix A).

B.1 Evaluation of Prompts and Models on the Human-Annotated Sample

Three prompts are tested:

P.1 Baseline prompt with direct questions, without additional instructions or definitions,

P.2 Structured prompt with direct questions supplemented by instructions and definitions,

P.3 Same as the P.2 but with more details in the instructions and definitions.

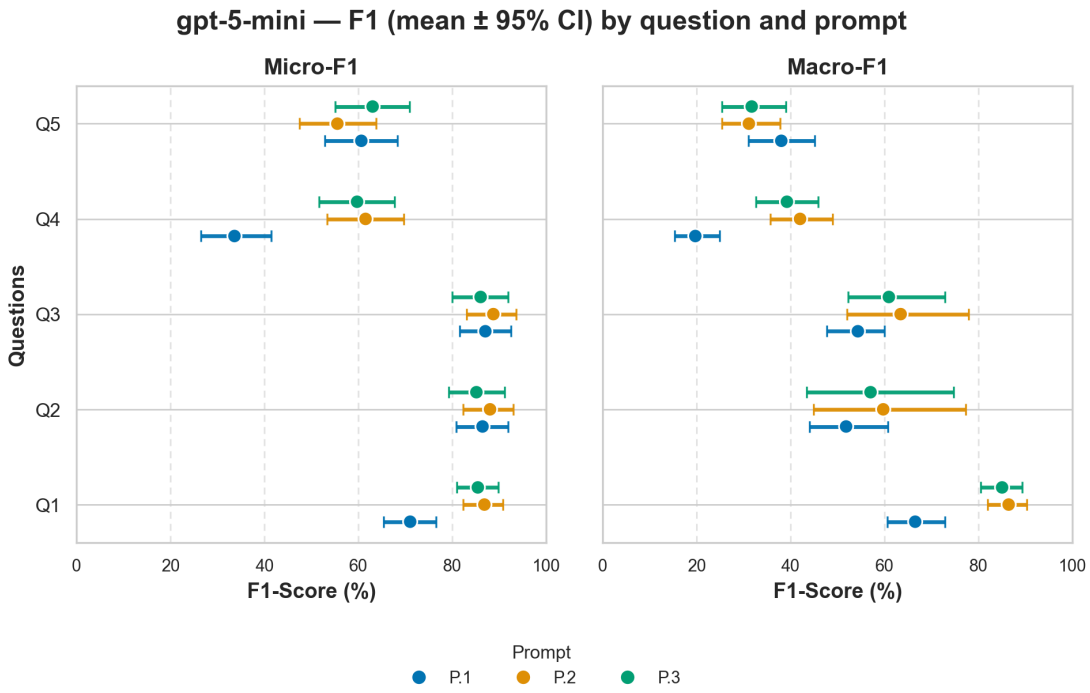


Figure 7. Sensitivity of LLM Classification Performance to Prompt Design

This figure reports micro and macro averaged F1-scores for the five annotation tasks (Q1–Q5) obtained with *GPT-5-mini* under three alternative prompt designs (P.1–P.3). Points denote mean F1-scores, and horizontal bars indicate 95% confidence intervals computed from 1,000 bootstrap replications. Differences across prompts highlight the sensitivity of classification performance to instruction design.

Table 10: LLM Classification Performance on the Manually Annotated Newspaper Sample

Model	Prompt	Question	Support	Micro (%)			Macro (%)			MAE	nMAE
				F1	P	R	F1	P	R		
gpt-5-mini	P.1	Q1	272	71.0	71.0	71.0	66.5	82.4	68.7		
		Q2	146	86.3	86.3	86.3	52.1	51.8	52.4		
		Q3	146	87.0	87.0	87.0	54.5	57.2	53.0	0.212	0.106
		Q4	140	33.6	33.6	33.6	19.8	29.3	24.2	0.771	0.386
		Q5	142	60.6	60.6	60.6	38.2	37.8	41.5	0.5	0.25
	P.2	Q1	271	86.7	86.7	86.7	86.3	88.5	85.9		
		Q2	141	87.9	87.9	87.9	62.7	63.9	62.3		
		Q3	141	88.7	88.7	88.7	64.7	67.6	62.9	0.184	0.092
		Q4	135	61.5	61.5	61.5	42.3	45.5	43.3	0.474	0.237
		Q5	135	55.6	55.6	55.6	31.4	31.4	35.4	0.556	0.278
	P.3	Q1	272	85.3	85.3	85.3	85.0	86.1	84.7		
		Q2	135	85.2	85.2	85.2	58.8	57.8	60.9		
		Q3	135	85.9	85.9	85.9	61.2	61.7	61.5	0.2	0.1
		Q4	124	59.7	59.7	59.7	39.4	43.9	40.3	0.524	0.262
		Q5	127	63.0	63.0	63.0	32.0	32.8	32.1	0.457	0.228
gpt-5-nano	P.1	Q1	272	68.4	68.4	68.4	63.0	80.3	65.9		
		Q2	145	85.5	85.5	85.5	52.7	53.5	52.0		
		Q3	145	84.8	84.8	84.8	55.3	65.0	51.9	0.234	0.117
		Q4	133	36.8	36.8	36.8	21.2	18.9	24.7	0.737	0.368
		Q5	137	53.3	53.3	53.3	35.0	36.1	40.6	0.65	0.325
	P.2	Q1	272	80.1	80.1	80.1	79.1	83.3	79.0		
		Q2	139	84.2	84.2	84.2	61.3	58.4	75.6		
		Q3	139	82.7	82.7	82.7	54.6	59.5	52.8	0.273	0.137
		Q4	129	62.8	62.8	62.8	42.0	52.2	43.3	0.465	0.233
		Q5	121	59.5	59.5	59.5	29.8	30.7	30.2	0.504	0.252
	P.3	Q1	269	81.8	81.8	81.8	81.4	82.5	81.1		
		Q2	136	83.8	83.8	83.8	54.5	54.6	60.1		
		Q3	136	75.0	75.0	75.0	52.1	60.1	55.2	0.316	0.158
		Q4	110	56.4	56.4	56.4	37.0	38.9	38.2	0.545	0.273
		Q5	121	58.7	58.7	58.7	28.8	29.1	29.4	0.496	0.248

Notes: This table reports classification performance on a manually annotated subset of the newspaper corpus for five annotation tasks (Q1–Q5). Results are shown for two models (*GPT-5-mini* and *GPT-5-nano*) and three prompt designs (P.1–P.3). For each task, the table reports micro and macro averaged F1-score, precision (P), and recall (R), expressed in percent. Support denotes the number of manually labeled articles used for evaluation. For the ordinal tasks (Q3–Q5), the table additionally reports mean absolute error (MAE) and normalized MAE (nMAE), where lower values indicate better performance.

Table 11: Bootstrapped F1-Scores for LLM Classification on the Manually Annotated Newspaper Sample

Model	Prompt	Question	Support	Micro (%)		Macro (%)	
				F1 (mean)	95% CI	F1 (mean)	95% CI
gpt-5-mini	P.1	Q1	272	71.0	[65.4, 76.5]	66.5	[60.6, 72.9]
		Q2	146	86.4	[80.8, 91.8]	51.8	[44.0, 60.7]
		Q3	146	87.0	[81.5, 92.5]	54.3	[47.7, 59.9]
		Q4	140	33.6	[26.4, 41.4]	19.7	[15.3, 24.9]
		Q5	142	60.6	[52.8, 68.3]	38.0	[31.0, 45.1]
	P.2	Q1	271	86.8	[82.3, 90.8]	86.4	[81.9, 90.3]
		Q2	141	88.0	[82.3, 92.9]	59.7	[44.9, 77.3]
		Q3	141	88.7	[83.0, 93.6]	63.4	[52.0, 77.9]
		Q4	135	61.5	[53.3, 69.6]	42.0	[35.7, 48.9]
		Q5	135	55.5	[47.4, 63.7]	31.1	[25.4, 37.7]
	P.3	Q1	272	85.4	[80.9, 89.7]	85.0	[80.4, 89.3]
		Q2	135	85.1	[79.2, 91.1]	57.0	[43.4, 74.7]
		Q3	135	86.0	[80.0, 91.9]	60.9	[52.2, 72.8]
		Q4	124	59.7	[51.6, 67.7]	39.2	[32.6, 45.8]
		Q5	127	63.0	[55.1, 70.9]	31.7	[25.4, 39.0]
gpt-5-nano	P.1	Q1	272	68.5	[62.9, 74.3]	63.0	[56.7, 69.3]
		Q2	145	85.6	[80.0, 91.0]	52.5	[44.2, 61.0]
		Q3	145	84.9	[78.6, 90.3]	54.5	[44.9, 66.6]
		Q4	133	36.8	[28.6, 45.1]	21.0	[16.8, 25.3]
		Q5	137	53.1	[44.5, 61.3]	34.7	[27.8, 41.6]
	P.2	Q1	272	80.2	[75.4, 84.9]	79.2	[74.0, 84.1]
		Q2	139	84.0	[77.7, 89.9]	60.0	[47.4, 73.4]
		Q3	139	82.7	[76.3, 89.2]	53.9	[42.8, 67.4]
		Q4	129	62.8	[54.3, 70.5]	41.8	[35.4, 48.8]
		Q5	121	59.5	[50.4, 68.6]	29.6	[23.3, 36.4]
	P.3	Q1	269	81.9	[77.3, 86.2]	81.4	[76.7, 86.0]
		Q2	136	83.8	[77.9, 89.7]	53.4	[43.1, 67.2]
		Q3	136	75.0	[67.6, 82.4]	51.6	[41.9, 61.2]
		Q4	110	56.3	[46.4, 65.5]	36.7	[29.4, 44.3]
		Q5	121	58.9	[50.4, 67.8]	28.4	[22.2, 35.6]

Notes: This table reports micro and macro averaged F1-scores for the five annotation tasks (Q1–Q5) on a manually annotated subset of the newspaper corpus. Results are shown for two models (*GPT-5-mini* and *GPT-5-nano*) and three prompt designs (P.1–P.3). For each task, the table reports the mean F1-score and the associated 95% confidence interval, computed from 1,000 bootstrap replications (resampling at the article level). All values are expressed in percent; higher values indicate better classification performance.

B.2 Robustness to Training-Data Contamination: Post–Cut-off Articles

Table 12: Pre- vs Post–Training Cut-off Performance of the LLM Classification

Group	Prompt	Question	Support	Micro (%)			Macro (%)			MAE	nMAE
				F1	P	R	F1	P	R		
		Q1	174	77.0	77.0	77.0	74.9	82.3	75.0		
		Q2	92	80.4	80.4	80.4	49.5	49.1	57.6		
G.1	P.2	Q3	91	58.2	58.2	58.2	31.4	39.4	35.4	0.538	0.269
		Q4	68	41.2	41.2	41.2	24.0	21.8	26.7	0.632	0.316
		Q5	72	43.1	43.1	43.1	29.7	31.2	32.1	0.764	0.382
		Q1	98	75.5	75.5	75.5	73.7	82.2	74.6		
		Q2	50	76.0	76.0	76.0	46.2	45.2	47.8		
G.2	P.2	Q3	49	71.4	71.4	71.4	36.2	36.2	58.3	0.327	0.163
		Q4	37	40.5	40.5	40.5	21.1	20.4	24.5	0.649	0.324
		Q5	40	62.5	62.5	62.5	29.0	29.5	31.7	0.525	0.263

Notes: This table reports classification performance on the manually annotated newspaper sample, split by publication date relative to the model's training cut-off (September 1, 2021). Group 1 (Pre-cutoff) includes articles published before the cut-off date, and Group 2 (Post-cutoff) includes articles published on or after the cut-off date. Results are shown for prompt P.2 and the indicated LLM model. For each annotation task (Q1–Q5), the table reports micro- and macro-averaged F1-score, precision (P), and recall (R), expressed in percent; Support denotes the number of manually labeled articles. For ordinal tasks, the table additionally reports MAE and normalized MAE (lower values indicate better performance). GPT-3.5-Turbo is used for this exercise because its earlier training cut-off allows a larger post-cutoff evaluation sample, strengthening the power of the contamination test.

Table 13: Bootstrapped F1-Scores Before and After the Model's Training Cut-off

Group	Prompt	Question	Support	Micro (%)		Macro (%)	
				F1 (mean)	95% CI	F1 (mean)	95% CI
G.1	P.2	Q1	174	77.0	[70.1, 82.8]	74.9	[68.1, 81.2]
		Q2	92	80.4	[72.8, 88.0]	49.0	[41.3, 58.1]
		Q3	91	58.4	[48.4, 68.1]	31.3	[24.9, 40.1]
		Q4	68	40.9	[29.4, 53.0]	23.6	[17.5, 29.9]
		Q5	72	42.9	[31.9, 54.2]	29.2	[20.9, 37.7]
G.2	P.2	Q1	98	75.5	[67.3, 83.7]	73.5	[63.8, 82.5]
		Q2	50	76.0	[64.0, 88.0]	45.6	[36.7, 57.4]
		Q3	49	71.1	[57.1, 83.7]	35.2	[26.2, 46.5]
		Q4	37	40.6	[24.3, 56.8]	20.8	[12.5, 29.3]
		Q5	40	62.9	[47.5, 77.5]	28.7	[18.8, 37.6]

Notes: This table reports bootstrapped micro- and macro-averaged F1-scores on the manually annotated newspaper sample, split by publication date relative to the model's training cut-off (September 1, 2021). Group 1 (Pre-cutoff) includes articles published before the cut-off date, and Group 2 (Post-cutoff) includes articles published on or after the cut-off date. Results are shown for prompt P.2 and the indicated LLM model. For each task (Q1–Q5), the table reports the mean F1-score and a 95% confidence interval computed from 1,000 bootstrap replications (resampling at the article level). All values are expressed in percent. GPT-3.5-Turbo is used for this exercise because its earlier training cut-off allows a larger post-cutoff evaluation sample, strengthening the power of the contamination test.

C Topic Modelling

I use the `text-embedding-3-large` model to generate embeddings for the content of each climate-related article. These embeddings were then clustered using the K-means algorithm (Lloyd, 1982) with $K = 40$ clusters. Finally, I group semantically similar clusters into 10 broader thematic categories.

I choose $k = 40$ as it corresponds to the highest silhouette score (Figure 8), and according to the elbow method, the inertia starts to decrease at a slower rate beyond this point.

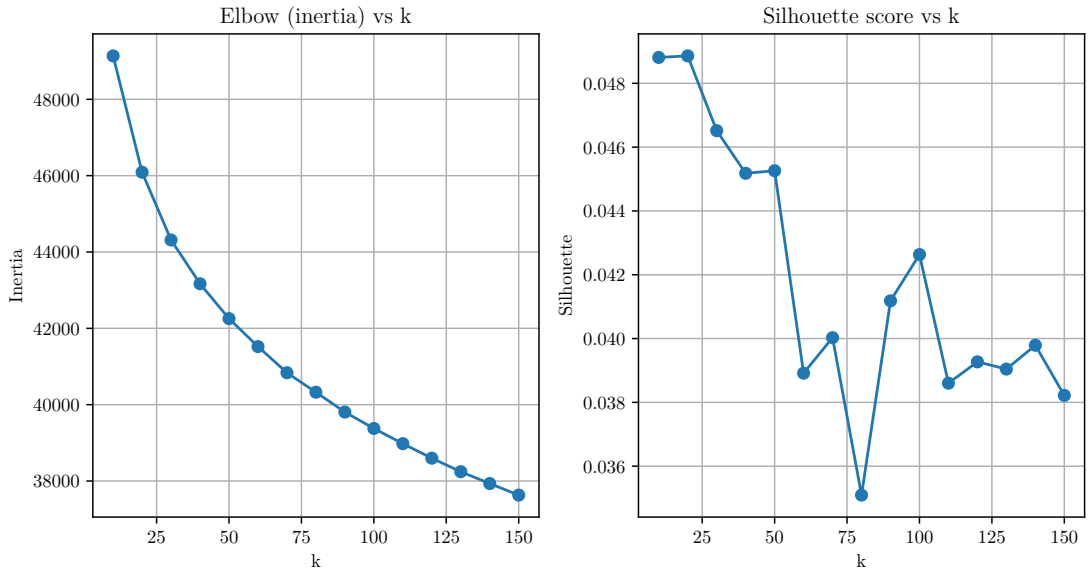


Figure 8. K-means Cluster Selection Diagnostics: Inertia (Elbow) and Silhouette Scores

The computations are based on the full sample of articles. The embeddings were obtained using the `text-embedding-3-large` model and reduced via a PCA with 500 components.

C.1 The clustering

To identify the topics discussed in the sample of news articles, I use BERTopic (Grootendorst, 2022) with the k-means clustering algorithm (Lloyd, 1982). BERTopic operates through several stages, which are described in this subsections.

Each document i is embedded using `text-embedding-3-large`, and thus transformed

into an embedding vector $v_i^\top \in \mathbb{R}^d$ with $d = 3072$. For the N documents, we obtain the following matrix $X \in \mathbb{R}^{N \times d}$ that contains the embeddings of all articles:

$$X = \begin{bmatrix} v_1^\top \\ v_2^\top \\ \vdots \\ v_N^\top \end{bmatrix} \tag{12}$$

To reduce the noise in the embeddings and improve computational efficiency, I apply a Principal Component Analysis (PCA) to X and retain 500 components:

$$Z = XW \Leftrightarrow Z = \begin{bmatrix} z_1^\top \\ z_2^\top \\ \vdots \\ z_N^\top \end{bmatrix} \in \mathbb{R}^{N \times 500} \tag{13}$$

Where $W \in \mathbb{R}^{d \times 500}$ is the matrix containing the eigenvalues of X. Now, each document is represented by $z_i = v_i^\top W \in \mathbb{R}^{500}$

Then, I perform the k-means clustering algorithm with $k = 40$. The objective of the algorithm is to partition the documents into k clusters by minimizing the within-cluster Euclidean distance:

$$\min_{C_1,\ldots,C_k} \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2 \tag{14}$$

Where μ_i is the centroid of cluster i (i.e., the average reduced embedding of all documents in that cluster), C_i denotes the set of documents assigned to cluster i, and x represents the embedding vector of a document (z_i).

Finally, for each cluster, 4 representative documents are selected based on their Euclidean distance from the cluster centroid μ_i: the closest document to the centroid, a central one, an averagely distant one, and one near the cluster's border. These selected documents are then

provided to the GPT-5 large language model (LLM) to generate a descriptive cluster name.

GPT-5 Prompt *You are labeling thematic clusters of articles in a research corpus where every document already relates to climate change and environmental issues. Your goal is to assign each cluster a clear, descriptive label that summarizes its main topic. Be specific enough to reflect the focus of the cluster, but general enough to apply to all documents within it. Use 5–20 words and avoid generic phrases like "climate change" or "environmental issues" unless essential. Return only the label, with no punctuation at the end.*

Indicative keywords: `{keywords}`
Representative documents: `{docs}`
Return ONLY the label

C.2 Thematic indicators construction

I then construct monthly, theme-level transition and physical climate risk indicators from the article-level GPT scores, using the same aggregation procedure as for the daily TRI and PRI (subsubsection 2.2.2), but applied at the monthly frequency and within themes.

C.3 Results

Table 14: Average Transition and Physical Risk Scores by Theme

Theme	Physical Risk		Transition Risk	
	Nb. Articles	Av. Scores	Nb. Articles	Av. Scores
Australian Climate Politics	1447	2.86	5923	1.00
Canadian Climate Politics	939	2.69	3876	1.87
Climate Activism, Protest and Public Opinion	4421	2.51	6196	1.59
International Climate Diplomacy	1128	2.25	6292	0.51
Technological and Energy Transition	2238	2.42	10474	1.99
UK Climate Politics	705	2.69	7724	1.36
USA Climate Politics	4609	2.32	12939	0.33
Climate Disasters and Extreme Weather	7262	3.13	13	2.27
Climate Science and Reports	9750	2.53	337	1.43
Food, Land Use and Development	6724	2.53	2236	1.77

Notes: This table reports the number of articles and average risk scores by thematic cluster. Themes above the dashed line are primarily associated with transition risk (regulatory, political, and technological developments); themes below are primarily associated with physical risk (extreme weather, scientific reports, environmental impacts). The number of articles reflects how often each theme appears within transition-classified (Q2 = Transition) versus physical-classified (Q2 = Physical) articles. For instance, "USA Climate Politics" contains 12,939 transition risk articles but only 4,609 physical risk articles, consistent with its policy-oriented content. The average scores are computed only over articles classified in each category and are not directly comparable across columns due to differences in sample composition.

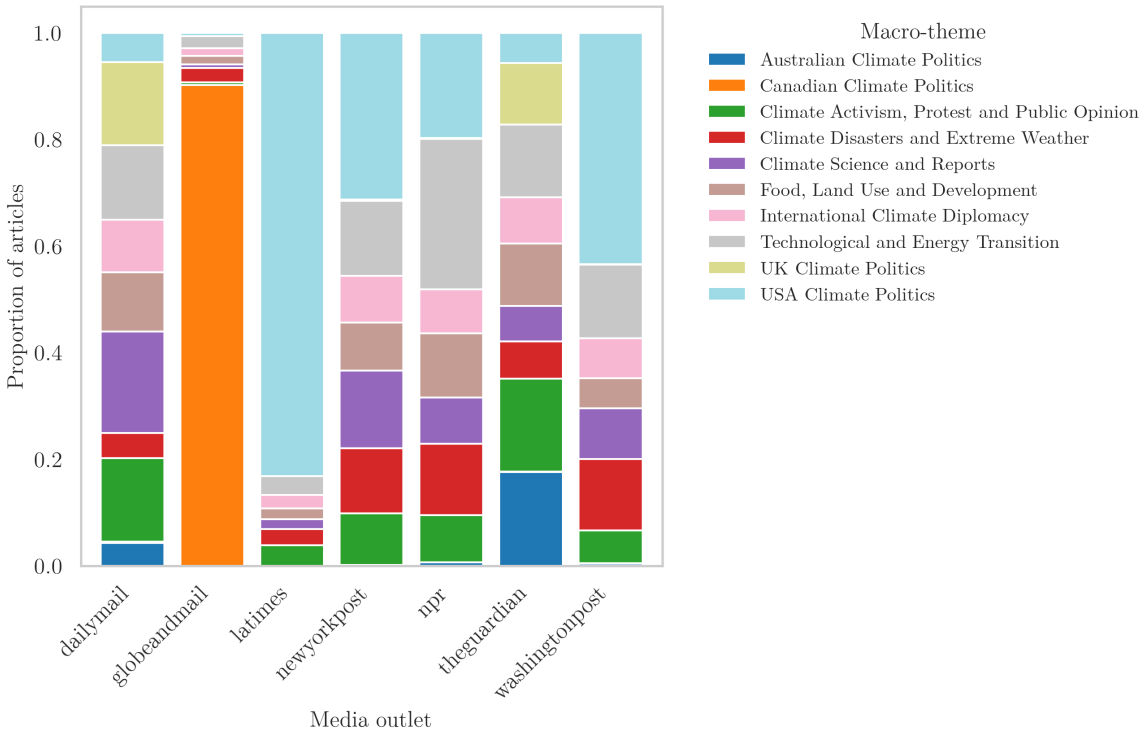


Figure 9. Thematic Composition of Climate-Related Articles by Outlet

This figure displays the distribution of thematic clusters within climate-related articles for each outlet. Each bar represents the share of articles belonging to each theme. Outlets exhibit distinct thematic profiles: The Globe and Mail focuses predominantly on Canadian Climate Politics, while U.S. outlets (Los Angeles Times, New York Post, NPR, Washington Post) emphasize USA Climate Politics. UK outlets (Daily Mail, The Guardian) display more balanced coverage across themes.

D Asset Pricing Robustness Checks

Table 15: Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls (12-Month Rolling Window)

	74 Portfolios						55 Portfolios					
	FF5		FFC		FF3		FF5		FFC		FF3	
Term	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
$\widetilde{TRI}$	-1.026	-0.822	-1.612*	-1.561*	-1.553*	-1.362*	-1.195	-0.607	-1.893**	-1.577*	-1.839**	-1.376
	(-1.29)	(-1.07)	(-1.94)	(-1.88)	(-1.93)	(-1.68)	(-1.38)	(-0.76)	(-2.25)	(-1.80)	(-2.00)	(-1.50)
$\widetilde{PRI}$	-0.381	-0.036	-0.500	-0.136	-0.148	0.139	-1.150	-0.840	-0.757	-0.505	-0.472	-0.315
	(-0.37)	(-0.03)	(-0.51)	(-0.13)	(-0.15)	(0.12)	(-0.90)	(-0.59)	(-0.70)	(-0.42)	(-0.42)	(-0.24)
MKT	0.005	0.003	0.008	0.007	0.008	0.006	0.006	0.004	0.012*	0.009	0.010*	0.006
	(0.93)	(0.57)	(1.13)	(1.01)	(1.32)	(0.87)	(1.27)	(0.81)	(1.77)	(1.48)	(1.68)	(0.90)
SMB	-0.001	-0.001	0.000	-0.000	0.000	0.000	-0.001	-0.001	-0.000	-0.000	0.000	-0.000
	(-0.35)	(-0.35)	(0.03)	(-0.00)	(0.12)	(0.03)	(-0.53)	(-0.43)	(-0.13)	(-0.07)	(0.01)	(-0.01)
HML	0.000	0.000	-0.001	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.00)	(0.00)	(-0.16)	(-0.01)	(-0.16)	(-0.12)	(-0.24)	(-0.27)	(-0.35)	(-0.16)	(-0.31)	(-0.24)
$\widetilde{EPU}$	-74.000		-84.020		-98.653		-128.431		-137.216		-165.788*	
	(-0.98)		(-0.91)		(-1.08)		(-1.24)		(-1.41)		(-1.81)	
MOM			0.006	0.007					0.006	0.008		
			(1.11)	(1.17)					(1.02)	(1.32)		
RMW	0.001	0.001					0.000	0.001				
	(0.39)	(0.63)					(0.22)	(0.70)				
CMA	-0.001	-0.000					-0.000	-0.000				
	(-0.21)	(-0.13)					(-0.02)	(-0.05)				
Intercept	0.004	0.006	0.000	0.002	0.000	0.003	0.002	0.004	-0.003	-0.001	-0.002	0.003
	(0.91)	(1.26)	(0.08)	(0.28)	(0.02)	(0.56)	(0.59)	(0.90)	(-0.63)	(-0.15)	(-0.43)	(0.55)
Sample Period	December 2014 – November 2024											
Observations	8880	8880	8880	8880	8880	8880	6600	6600	6600	6600	6600	6600

Notes: This table reports second-pass Fama–MacBeth risk premium estimates over the period 2014–2024. Betas are estimated using daily data and a 12-month rolling window. Results are shown for two asset universes: 74 portfolios (25 Size–Book-to-Market and 49 industry portfolios) and 55 portfolios (25 Size–BM and 30 industry portfolios), all obtained from the Kenneth R. French Data Library. I consider alternative factor control sets (FF3, FFC, and FF5), with and without innovations in the Economic Policy Uncertainty index. Transition (TRI) and physical (PRI) climate risk factors enter in innovations. Newey–West (Newey and West, 1987) (6) t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 16: Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls (24-Month Rolling Window)

	74 Portfolios						55 Portfolios					
	FF5		FFC		FF3		FF5		FFC		FF3	
Term	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
$\widetilde{TRI}$	-2.159**	-1.891**	-2.271**	-2.388**	-2.483**	-2.367**	-3.227***	-2.904**	-3.877***	-3.890***	-3.981***	-3.703***
	(-2.30)	(-2.16)	(-2.36)	(-2.51)	(-2.57)	(-2.45)	(-2.81)	(-2.53)	(-3.26)	(-2.92)	(-3.33)	(-2.71)
$\widetilde{PRI}$	0.538	0.346	0.471	0.214	0.551	0.194	-0.351	-0.469	-0.275	-0.458	-0.257	-0.500
	(0.31)	(0.21)	(0.28)	(0.13)	(0.32)	(0.12)	(-0.19)	(-0.26)	(-0.14)	(-0.25)	(-0.14)	(-0.28)
MKT	0.007	0.007	0.009	0.008	0.007	0.007	0.010**	0.010*	0.010	0.009	0.008	0.007
	(1.21)	(1.15)	(1.28)	(1.07)	(1.17)	(1.03)	(1.96)	(1.94)	(1.42)	(1.32)	(1.29)	(1.14)
SMB	-0.001	-0.001	-0.000	0.000	-0.000	-0.000	-0.001	-0.000	0.000	0.000	-0.000	0.000
	(-0.26)	(-0.20)	(-0.04)	(0.00)	(-0.08)	(-0.03)	(-0.22)	(-0.18)	(0.04)	(0.06)	(-0.02)	(0.02)
HML	0.000	0.001	-0.001	0.000	-0.000	0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.10)	(0.21)	(-0.11)	(0.03)	(-0.01)	(0.10)	(-0.21)	(-0.19)	(-0.29)	(-0.19)	(-0.26)	(-0.24)
$\widetilde{EPU}$	62.064		62.683		56.408		17.065		-4.739		-18.875	
	(0.64)		(0.67)		(0.54)		(0.19)		(-0.05)		(-0.19)	
MOM			-0.008	-0.009					-0.006	-0.004		
			(-1.04)	(-1.15)					(-0.72)	(-0.55)		
RMW	0.001	0.001					0.001	0.001				
	(0.26)	(0.38)					(0.33)	(0.60)				
CMA	-0.003	-0.002					-0.003	-0.003				
	(-0.80)	(-0.57)					(-0.84)	(-0.83)				
Intercept	0.002	0.002	0.000	0.002	0.002	0.003	0.000	0.000	-0.000	0.000	0.001	0.002
	(0.44)	(0.43)	(0.05)	(0.27)	(0.40)	(0.53)	(0.04)	(0.01)	(-0.07)	(0.08)	(0.23)	(0.47)
Sample Period	December 2015 – November 2024											
Observations	7992	7992	7992	7992	7992	7992	5940	5940	5940	5940	5940	5940

Notes: This table reports second-pass Fama–MacBeth risk premium estimates over the period 2014–2024. Betas are estimated using daily data and a 12-month rolling window. Results are shown for two asset universes: 74 portfolios (25 Size–Book-to-Market and 49 industry portfolios) and 55 portfolios (25 Size–BM and 30 industry portfolios), all obtained from the Kenneth R. French Data Library. I consider alternative factor control sets (FF3, FFC, and FF5), with and without innovations in the Economic Policy Uncertainty index. Transition (TRI) and physical (PRI) climate risk factors enter in innovations. Newey–West (Newey and West, 1987) (6) t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

E Additional Results for the Attention Indicators

Table 17: Single Predictive Regressions of Aggregate Market Returns on the Transition and Physical Attention Indices

$q =$	1	3	12	24	36	48
AT	-0.089	-0.657	-1.096	3.652	-2.313	11.199***
t-stat$_{\text{NW}}$	(-0.10)	(-0.37)	(-0.18)	(0.51)	(-0.80)	(3.04)
t-stat$_{\text{HH}}$	(-0.11)	(-0.38)	(-0.17)	(0.55)	(-0.77)	(2.99)
Adj. R^2 (%)	-0.76	-0.66	-0.76	-0.13	-0.28	16.86
AP	-2.541***	-3.428	1.829	9.897	-0.730	6.073
t-stat$_{\text{NW}}$	(-2.59)	(-1.36)	(0.32)	(1.64)	(-0.20)	(0.84)
t-stat$_{\text{HH}}$	(-2.59)	(-1.40)	(0.35)	(2.11)	(-0.16)	(0.89)
Adj. R^2 (%)	4.05	3.23	-0.56	5.37	-0.98	3.76

Notes: This table reports single-predictive regressions of the value-weighted U.S. market return compounded over horizons of 1, 3, 12, 24, 36, and 48-month on the current level of the Transition Attention Indicator (AT) or the Physical Attention Indicator (AP). The sample runs from January 2014 to December 2024. For each horizon q, the first q months of data are lost by construction. Newey–West q-lag t-statistics (Newey and West, 1987) and Hansen–Hodrick (Hansen and Hodrick, 1980) q-lag t-statistics are reported below the coefficient estimates. The estimates are in percentage. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 18: Multiple Predictive Regression of Aggregate Market Returns with Attention Indicators

	AT	AP	TERM	DEF	DY	RF	WTI	VIX	Adj. R^2 (%)
Panel A: $q = 1$									
(i)	0.005*	-0.017***	-0.014	0.001	0.003	-0.002			6.83
	(1.74)	(-3.95)	(-1.48)	(0.11)	(0.66)	(-0.22)			
	[1.70]	[-3.96]	[-1.49]	[0.10]	[0.65]	[-0.22]			
(ii)	0.007**	-0.014***	-0.007	-0.012*	0.015**	0.010		0.017***	11.95
	(2.34)	(-3.15)	(-0.76)	(-1.94)	(2.27)	(1.13)		(2.90)	
	[2.26]	[-3.17]	[-0.77]	[-1.93]	[2.24]	[1.13]		[2.98]	
(iii)	0.005*	-0.018***	-0.016	0.000	0.003	-0.004	-0.005		7.33
	(1.69)	(-3.99)	(-1.59)	(0.08)	(0.57)	(-0.48)	(-1.20)		
	[1.64]	[-3.99]	[-1.65]	[0.07]	[0.55]	[-0.51]	[-1.32]		
(iv)	0.007**	-0.015***	-0.008	-0.011*	0.014**	0.008	-0.002	0.016***	11.38
	(2.27)	(-3.24)	(-0.81)	(-1.86)	(2.26)	(0.89)	(-0.55)	(2.86)	
	[2.18]	[-3.26]	[-0.82]	[-1.81]	[2.19]	[0.89]	[-0.56]	[2.83]	
Panel B: $q = 12$									
(i)	-0.003	-0.010	-0.087*	0.018	0.073***	-0.009			45.33
	(-0.24)	(-0.93)	(-1.78)	(1.48)	(3.52)	(-0.18)			
	[-0.22]	[-0.90]	[-1.61]	[1.45]	[3.42]	[-0.17]			
(ii)	0.002	0.001	-0.058	-0.028*	0.114***	0.039		0.065***	54.50
	(0.17)	(0.07)	(-1.50)	(-1.83)	(4.85)	(0.90)		(3.88)	
	[0.16]	[0.07]	[-1.40]	[-1.83]	[4.48]	[0.86]		[3.64]	
(iii)	-0.004	-0.012	-0.094*	0.017	0.072***	-0.017	-0.014*		45.99
	(-0.27)	(-1.02)	(-1.87)	(1.41)	(3.49)	(-0.32)	(-1.92)		
	[-0.25]	[-0.98]	[-1.68]	[1.37]	[3.33]	[-0.29]	[-2.18]		
(iv)	0.002	0.001	-0.058	-0.028*	0.113***	0.038	-0.001	0.064***	54.10
	(0.16)	(0.05)	(-1.50)	(-1.67)	(4.69)	(0.85)	(-0.13)	(3.45)	
	[0.15]	[0.05]	[-1.40]	[-1.63]	[4.31]	[0.82]	[-0.12]	[3.20]	
Panel C: $q = 48$									
(i)	0.032***	0.018	-0.066***	-0.031**	0.063***	-0.138***			28.56
	(10.44)	(1.00)	(-5.23)	(-2.41)	(4.01)	(-8.29)			
	[4.85]	[0.83]	[-2.43]	[-1.84]	[1.86]	[-4.88]			
(ii)	0.037***	0.033**	-0.038***	-0.071***	0.111***	-0.107***		0.058***	41.62
	(11.16)	(2.28)	(-3.11)	(-5.96)	(4.45)	(-5.77)		(3.73)	
	[5.27]	[2.22]	[-1.46]	[-4.16]	[3.27]	[-3.41]		[2.50]	
(iii)	0.032***	0.016	-0.074***	-0.032***	0.068***	-0.149***	-0.011		29.28
	(11.20)	(0.83)	(-4.37)	(-2.63)	(4.11)	(-6.47)	(-1.37)		
	[4.60]	[0.74]	[-2.47]	[-2.08]	[2.05]	[-5.21]	[-0.91]		
(iv)	0.037***	0.033**	-0.036***	-0.071***	0.112***	-0.105***	0.002	0.059***	40.87
	(10.82)	(2.18)	(-2.66)	(-6.30)	(4.44)	(-5.02)	(0.33)	(3.99)	
	[5.29]	[2.12]	[-1.44]	[-4.10]	[3.25]	[-3.50]	[0.22]	[2.59]	

Notes: This table reports multiple-predictive regressions of the value-weighted U.S. market return compounded over horizons $q \in \{1, 12, 48\}$ months on the AT and AP in levels, with control variables: term spread (TERM), default spread (DEF), dividend yield (DY), risk-free rate (RF), VIX, and WTI crude oil returns. All regressors are standardized. Sample: January 2014 – November 2024. The first q months are lost by construction. Newey and West (1987) and Hansen and Hodrick (1980) q-lag t-statistics reported in parentheses and brackets, respectively. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 19: Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls (12-Month Rolling Window) for Attention Indicators

	74 Portfolios						55 Portfolios					
	FF5		FFC		FF3		FF5		FFC		FF3	
Term	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
$\widetilde{AT}$	0.133	0.193	-0.492	-0.697	-0.548	-0.488	-1.012	-0.766	-1.631*	-1.731**	-1.605*	-1.519
	(0.19)	(0.29)	(-0.79)	(-1.20)	(-0.76)	(-0.69)	(-1.21)	(-0.83)	(-1.92)	(-2.04)	(-1.66)	(-1.56)
$\widetilde{AP}$	-0.487	-0.205	-0.650	-0.327	-0.189	-0.019	-1.354	-1.291	-1.214	-1.078	-0.885	-0.905
	(-0.51)	(-0.21)	(-0.75)	(-0.37)	(-0.21)	(-0.02)	(-1.25)	(-1.07)	(-1.37)	(-1.08)	(-0.98)	(-0.82)
MKT	0.006	0.005	0.010	0.008	0.009	0.007	0.004	0.004	0.011	0.010	0.008	0.006
	(1.23)	(0.85)	(1.34)	(1.21)	(1.33)	(0.96)	(0.84)	(0.79)	(1.54)	(1.41)	(1.24)	(0.83)
SMB	-0.001	-0.001	-0.000	-0.000	0.000	-0.000	-0.001	-0.001	-0.000	-0.000	-0.000	0.000
	(-0.38)	(-0.41)	(-0.00)	(-0.03)	(0.02)	(-0.06)	(-0.53)	(-0.42)	(-0.15)	(-0.05)	(-0.02)	(0.02)
HML	0.000	0.000	-0.000	0.000	0.000	0.000	-0.001	-0.001	-0.001	-0.000	-0.001	-0.001
	(0.04)	(0.06)	(-0.08)	(0.05)	(0.01)	(0.04)	(-0.16)	(-0.19)	(-0.25)	(-0.07)	(-0.18)	(-0.13)
$\widetilde{EPU}$	-95.492		-99.876		-101.547		-138.419		-130.351		-157.102*	
	(-1.27)		(-1.04)		(-1.14)		(-1.59)		(-1.38)		(-1.87)	
MOM			0.007	0.007					0.008	0.009		
			(1.20)	(1.27)					(1.36)	(1.58)		
RMW	0.001	0.002					0.001	0.002				
	(0.45)	(0.66)					(0.24)	(0.69)				
CMA	-0.000	-0.000					-0.000	-0.000				
	(-0.13)	(-0.12)					(-0.03)	(-0.06)				
Intercept	0.002	0.004	-0.001	-0.000	-0.001	0.002	0.005	0.004	-0.003	-0.001	-0.000	0.002
	(0.53)	(0.92)	(-0.25)	(-0.07)	(-0.15)	(0.31)	(1.15)	(1.01)	(-0.48)	(-0.28)	(-0.07)	(0.45)
Sample Period	December 2014 – November 2024											
Observations	8,880	8,880	8,880	8,880	8,880	8,880	6,600	6,600	6,600	6,600	6,600	6,600

Notes: This table reports second-pass Fama–MacBeth risk premium estimates over the period 2014–2024. Betas are estimated using daily data and a 12-month rolling window. Results are shown for two asset universes: 74 portfolios (25 Size–Book-to-Market and 49 industry portfolios) and 55 portfolios (25 Size–BM and 30 industry portfolios), all obtained from the Kenneth R. French Data Library. I consider alternative factor control sets (FF3, FFC, and FF5), with and without innovations in the Economic Policy Uncertainty index. Transition (AT) and Physical (AP) Attention indicator factors enter in innovations. Newey–West (Newey and West, 1987) (6) t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

Table 20: Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls (24-Month Rolling Window) for Attention Indicators

	74 Portfolios						55 Portfolios					
	FF5		FFC		FF3		FF5		FFC		FF3	
Term	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
$\widetilde{AT}$	-1.226	-0.977	-1.098	-0.999	-1.103	-0.843	-1.528	-1.012	-1.586	-1.156	-1.291	-0.632
	(-1.08)	(-0.75)	(-0.94)	(-0.80)	(-0.97)	(-0.66)	(-1.07)	(-0.62)	(-1.17)	(-0.77)	(-0.85)	(-0.37)
$\widetilde{AP}$	0.949	0.637	0.769	0.396	1.006	0.608	0.279	-0.138	0.420	-0.010	0.539	0.157
	(0.64)	(0.48)	(0.53)	(0.31)	(0.68)	(0.47)	(0.18)	(-0.11)	(0.27)	(-0.01)	(0.34)	(0.12)
MKT	0.010	0.009	0.010	0.009	0.010	0.008	0.011*	0.012*	0.010	0.010	0.010	0.009
	(1.48)	(1.41)	(1.44)	(1.20)	(1.46)	(1.21)	(1.81)	(1.93)	(1.58)	(1.49)	(1.63)	(1.47)
SMB	-0.001	-0.001	-0.001	-0.000	-0.001	-0.000	-0.001	-0.001	-0.001	-0.000	-0.001	-0.000
	(-0.39)	(-0.31)	(-0.19)	(-0.10)	(-0.23)	(-0.13)	(-0.51)	(-0.41)	(-0.24)	(-0.10)	(-0.28)	(-0.14)
HML	0.000	0.001	-0.000	0.000	0.000	0.000	-0.001	-0.001	-0.002	-0.001	-0.001	-0.001
	(0.09)	(0.16)	(-0.10)	(0.04)	(0.00)	(0.10)	(-0.17)	(-0.11)	(-0.32)	(-0.11)	(-0.28)	(-0.12)
$\widetilde{EPU}$	111.886		111.778		107.510		116.270		107.584		102.701	
	(1.18)		(1.11)		(0.96)		(1.11)		(0.85)		(0.78)	
MOM			-0.007	-0.008					-0.005	-0.005		
			(-0.98)	(-1.11)					(-0.72)	(-0.71)		
RMW	-0.001	-0.001					-0.001	-0.001				
	(-0.24)	(-0.26)					(-0.49)	(-0.31)				
CMA	-0.002	-0.001					-0.002	-0.002				
	(-0.45)	(-0.25)					(-0.52)	(-0.38)				
Intercept	-0.000	0.000	-0.001	0.000	-0.000	0.001	-0.001	-0.002	-0.001	-0.001	-0.001	-0.000
	(-0.02)	(0.00)	(-0.19)	(0.02)	(-0.10)	(0.14)	(-0.19)	(-0.34)	(-0.29)	(-0.22)	(-0.23)	(-0.07)
Sample Period	December 2015 – November 2024											
Observations	7,992	7,992	7,992	7,992	7,992	7,992	5,940	5,940	5,940	5,940	5,940	5,940

Notes: This table reports second-pass Fama–MacBeth risk premium estimates over the period 2015–2024. Betas are estimated using daily data and a 24-month rolling window. Results are shown for two asset universes: 74 portfolios (25 Size–Book-to-Market and 49 industry portfolios) and 55 portfolios (25 Size–BM and 30 industry portfolios), all obtained from the Kenneth R. French Data Library. I consider alternative factor control sets (FF3, FFC, and FF5), with and without innovations in the Economic Policy Uncertainty index. Transition (AT) and Physical (AP) Attention factors enter in innovations. Newey–West (Newey and West, 1987) (6) t-statistics are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

F Robustness to Excluding Intensity and Horizon Labels (Q4–Q5 Ablations)

Table 21: Single Predictive Regressions of Aggregate Market Returns: Q1-Q3 Only Indicators (Robustness)

$q =$	1	3	12	24	36	48
$\text{TRI}_{\text{Q1–Q3}}$	-1.119	-1.791	-6.135	3.733	-4.669	-6.666*
$\text{t-stat}_{\text{NW}}$	(-1.34)	(-1.01)	(-0.82)	(0.25)	(-0.82)	(-1.95)
$\text{t-stat}_{\text{HH}}$	(-1.28)	(-1.16)	(-0.96)	(0.41)	(-0.84)	(-1.43)
Adj. R^2 (%)	-0.03	0.09	1.80	-0.10	2.04	4.32
$\text{PRI}_{\text{Q1–Q3}}$	-2.667**	-3.718	0.826	11.577	0.208	5.730
$\text{t-stat}_{\text{NW}}$	(-2.38)	(-1.43)	(0.12)	(1.53)	(0.05)	(0.71)
$\text{t-stat}_{\text{HH}}$	(-2.38)	(-1.43)	(0.13)	(2.12)	(0.04)	(0.76)
Adj. R^2 (%)	3.59	3.10	-0.79	6.26	-1.05	2.49

Notes: This table reports single-predictive regressions of the value-weighted U.S. market return compounded over horizons of 1, 3, 12, 24, 36, and 48-month on the current level of the transition and physical risk indicators constructed using Q1–Q3 only ($\text{TRI}_{\text{Q1–Q3}}$ and $\text{PRI}_{\text{Q1–Q3}}$). The sample runs from January 2014 to December 2024. For each horizon q, the first q months of data are lost by construction. Newey–West q-lag t-statistics (Newey and West, 1987) and Hansen–Hodrick (Hansen and Hodrick, 1980) q-lag t-statistics are reported below the coefficient estimates. The estimates are in percentage. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 22: Fama–MacBeth Risk Premia with Alternative Factor Controls: Q1–Q3-Only Indicator (Excluding Q4–Q5), 12-Month Rolling Window

	74 Portfolios						55 Portfolios					
	FF5		FFC		FF3		FF5		FFC		FF3	
Term	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)	(i)	(ii)
$\widetilde{TRI}_{\text{Q1-Q3}}$	-0.976	-0.794	-1.793**	-1.864**	-1.709**	-1.638*	-1.241	-0.662	-2.146**	-2.010**	-2.135**	-1.763*
	(-1.25)	(-1.05)	(-2.21)	(-2.28)	(-2.06)	(-1.95)	(-1.31)	(-0.78)	(-2.40)	(-2.15)	(-2.10)	(-1.78)
$\widetilde{PRI}_{\text{Q1-Q3}}$	-0.354	-0.085	-0.394	-0.041	-0.177	-0.001	-1.058	-1.072	-0.627	-0.501	-0.423	-0.498
	(-0.37)	(-0.09)	(-0.44)	(-0.05)	(-0.19)	(-0.00)	(-0.85)	(-0.76)	(-0.60)	(-0.44)	(-0.38)	(-0.38)
MKT	0.005	0.003	0.008	0.007	0.008	0.006	0.006	0.003	0.012*	0.009	0.010*	0.005
	(1.03)	(0.67)	(1.14)	(1.08)	(1.31)	(0.89)	(1.25)	(0.63)	(1.85)	(1.62)	(1.65)	(0.79)
SMB	-0.001	-0.001	0.000	0.000	0.000	0.000	-0.001	-0.001	-0.000	-0.000	0.000	0.000
	(-0.28)	(-0.27)	(0.03)	(0.01)	(0.12)	(0.03)	(-0.44)	(-0.34)	(-0.11)	(-0.05)	(0.05)	(0.03)
HML	-0.000	-0.000	-0.001	0.000	-0.001	-0.000	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(-0.01)	(-0.01)	(-0.14)	(0.02)	(-0.12)	(-0.10)	(-0.23)	(-0.23)	(-0.32)	(-0.14)	(-0.26)	(-0.20)
$\widetilde{EPU}$	-94.789		-91.324		-104.132		-140.688		-141.999		-163.935*	
	(-1.22)		(-1.00)		(-1.15)		(-1.30)		(-1.41)		(-1.70)	
MOM			0.007	0.007					0.007	0.008		
			(1.17)	(1.24)					(1.12)	(1.44)		
RMW	0.001	0.002					0.001	0.002				
	(0.51)	(0.71)					(0.26)	(0.72)				
CMA	-0.001	-0.001					-0.000	-0.000				
	(-0.20)	(-0.16)					(-0.01)	(-0.10)				
Intercept	0.004	0.005	0.001	0.001	0.000	0.003	0.003	0.005	-0.003	-0.001	-0.001	0.004
	(0.83)	(1.18)	(0.10)	(0.24)	(0.05)	(0.57)	(0.65)	(1.11)	(-0.65)	(-0.19)	(-0.30)	(0.75)
Sample Period	December 2014 – November 2024											
Observations	8880	8880	8880	8880	8880	8880	6600	6600	6600	6600	6600	6600

Notes: This table reports second-pass Fama–MacBeth risk premium estimates over the period 2014–2024. Betas are estimated using daily data and a 12-month rolling window. Results are shown for two asset universes: 74 portfolios (25 Size–Book-to-Market and 49 industry portfolios) and 55 portfolios (25 Size–BM and 30 industry portfolios), all obtained from the Kenneth R. French Data Library. I consider alternative factor control sets (FF3, FFC, and FF5), with and without innovations in the Economic Policy Uncertainty index. The transition and physical risk factors constructed using Q1–Q3 only—$\widetilde{TRI}_{\text{Q1–Q3}}$ and $\widetilde{PRI}_{\text{Q1–Q3}}$—enter as innovations. Newey–West (Newey and West, 1987) t-statistics with lag length 6 are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.