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

Clément Toussaint †

First Version: March 17, 2026; This Version: August 25, 2026

**Abstract**

Climate news can signal rising or falling risk, yet existing news-based climate risk measures do not distinguish between the two. Using a large language model, I classify over 160,000 newspaper articles and build daily transition and physical climate risk indicators that capture the direction, intensity, and horizon of the risk each article conveys. Transition-risk shocks carry a significantly negative price of risk in the cross-section of U.S. stock returns, and higher transition risk predicts lower future market returns—the joint pattern implied by intertemporal hedging. Physical risk forecasts returns but is not priced, consistent with its localized nature. In horse-race tests, only the directional indicator is priced, and removing direction reverses the sign of the predictive relation, breaking the consistency between pricing and predictability. Direction is thus what turns climate news into a measure of transition risk: without it, media coverage captures attention, not risk.

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

*JEL Classification:* G12, Q54.

*I thank Jérôme Glachant and Thomas Renault for helpful comments and suggestions. I also thank Sylvain Marsat and Skander Maraoui for their thoughtful discussions, as well as participants at the 2026 AFFI Annual Conference and the 42nd GdRE Annual Symposium on Money, Banking and Finance for valuable feedback. All remaining errors are my own.

†Université Paris 1 Panthéon-Sorbonne, CNRS, Centre d'Économie de la Sorbonne, 106-112 boulevard de l'Hôpital, 75642 Paris Cedex 13, France. Email: 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; Chu et al., 2025). 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 signals 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 reflects 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). A natural concern with LLM-based measurement is look-ahead bias: a model trained on data extending beyond an article's publication date could draw on knowledge of subsequent events rather than on the article itself. I address this concern through three complementary checks: a comparison of classification performance on articles published before and after the model's training cut-off date, a prompt design that asks the model to characterize article content rather than to assess outcomes, and an ablation restricted to the dimensions least vulnerable to look-ahead bias. None of these checks alters the main results (Section 2.1.2). I apply this framework to a corpus of more than 160,000 newspaper articles, of which around 100,000 are classified as climate-related, producing directional, horizon-specific, and intensity-weighted transition (TRI) and physical (PRI) 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 January 2014–July 2025 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, with average realized abnormal returns rising from negative values in the low-beta quintiles to positive values in the highest. Fama and MacBeth (1973) regressions show a significantly negative risk premium associated with exposure to TRI innovations; through the lens of the ICAPM, this reflects investors' willingness to accept lower expected returns on assets that hedge against transition-risk shocks. The two findings are complementary rather than contradictory: hedging assets command lower expected excess returns, yet earn positive realized abnormal returns in periods when transition-risk shocks materialize frequently, as in my sample. 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). Moreover, the estimated premium is not subsumed by existing daily measures: in horse-race Fama–MacBeth regressions that include the climate news indicators of Ardia et al. (2023) and Faccini et al. (2023), the TRI premium remains negative and significant while the competing measures are not priced. This is consistent with their unidirectional design. Over my sample, climate policy moves in both directions—the Paris Agreement and the Inflation Reduction Act on one side, the U.S. withdrawal from the Paris Agreement and the subsequent regulatory rollbacks on the other—so that a given volume of climate coverage can signal rising or falling transition risk depending on the episode. A measure that does not sign the news aggregates the two, and the resulting series carries no consistent risk content over such a sample. The directional indicator is not exposed to this problem by construction.[1]

Second, the results are consistent with an intertemporal hedging interpretation in the

[1]Faccini et al. (2023) report significant pricing of their indicator over 2012–2018. Their result does not contradict this explanation: they describe that period as one of climate policy rollback in the United States (p. 2), a setting in which distinguishing risk-increasing from risk-decreasing news does not play a material role.

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, the long–short portfolio sorted on exposure to PRI innovations earns no significant abnormal return in any specification. 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 climate finance literature and empirical asset pricing. First, I provide cross-sectional evidence that transition risk innovations are priced in U.S. stock returns, with pricing content that is not captured by existing daily news-based climate risk measures. 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 an LLM-based text classification methodology. Importantly, incorporating direction is essential for identifying transition risk as an ICAPM state variable. When direction is removed and the indicator is reduced to a unidirectional attention measure, both requirements fail: exposure to attention shocks commands no significant risk premium, and the attention measure no longer forecasts aggregate market returns with a sign consistent with the cross-sectional estimates. In other words, direction is what turns climate news coverage into a proxy for transition risk: attention measures conflate risk-increasing and risk-decreasing news, muting the very signal that carries the premium.

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 two 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.

## 2.1 LLM-Based Climate Risk Annotation

### 2.1.1 Annotation Framework and Classification Validation

I use an LLM, *GPT-5-mini*, to analyze the content of each news article. A growing literature shows that LLMs can support social-science measurement through structured text coding (Bail, 2024), including applications to economically meaningful policy content (Cook et al., 2023; Jha et al., 2024; Lopez-Lira and Tang, 2026).

For each article, I prompt the LLM 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, and the full prompt is provided in the Internet Appendix Section IA2 (prompt P.2):

[Figure 1 about here.]

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*[2] *of the risk described in the article?* Weak / Medium / Strong / N.A.,

Q.5 *What is the time horizon*[3] *of the identified climate risk in the article?* Short / Medium / Long / N.A.

I selected the prompt through a systematic evaluation, testing three specifications on a manually annotated sample of approximately 300 articles, randomly sampled across outlets and over the full sample period (2010–2025): 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 cases degraded—F1-scores (see Section IA2 of the Internet Appendix). 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 to classify the full corpus.

Table 1 reports the classification performance of the selected prompt (P.2) on the manually labeled 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).

[Table 1 about here.]

[2] Intensity depends on the geographical scope of the measure or event, the degree to which it is binding, and its quantitative impact (see Internet Appendix Section IA2 for the full definition).

[3] A *short* horizon corresponds to risks expected to materialize within 0–3 years, a *medium* horizon within 3–10 years, and a *long* horizon beyond 10 years (see Internet Appendix Section IA2 for the full definition).

For Questions 4 and 5, macro F1-scores fall below 50%, reflecting the difficulty of distinguishing adjacent ordinal categories (e.g., *Medium* vs. *Strong* intensity). The normalized mean absolute error (nMAE) puts this weakness in perspective: bounded between 0 (perfect classification) and 1 (systematic misclassification at opposite ends of the scale), an nMAE below 0.5 implies an average error of less than one ordinal category. At 0.24 for Q4 and 0.28 for Q5, the observed values confirm that the model rarely confuses *Short* with *Long* or *Weak* with *Strong*—errors that would materially distort the risk indicators. Moreover, because adjacent categories receive nearby weights in equation (2) (e.g. $I_a = 1$ for Medium vs. $I_a = 2$ for Strong), such confusions have only a limited effect on the aggregate score.

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 A), 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 look-ahead 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 (3). The main asset-pricing results—cross-sectional risk premia, return predictability, and ICAPM consistency—remain unchanged (see Section IA5 of the Internet Appendix). 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-labeled 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 the article as climate-related, the article is scored; otherwise its score is set to zero:

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

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

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

$$D_a = \begin{cases} +1 & \text{if Increase,} \\ 0 & \text{if Neutral,} \\ -1 & \text{if Decrease,} \end{cases} \quad I_a = \begin{cases} 0.5 & \text{if Weak,} \\ 1.0 & \text{if Medium,} \\ 2.0 & \text{if Strong,} \\ 1.0 & \text{if NA (7\%),} \end{cases} \quad H_a = \begin{cases} 2.0 & \text{if Short,} \\ 1.0 & \text{if Medium,} \\ 0.5 & \text{if Long,} \\ 1.0 & \text{if NA (6\%).} \end{cases} \tag{2}$$

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, the intensity weight varies by a factor of four between weak and strong articles, and the horizon weight by a factor of four between long and short horizons; combined with the additive structure of the score (Eq. (3)), a strong article carries roughly twice the weight of a weak one at a given horizon. 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:

$$\text{Score}_a = \mathbf{1}_a^{\text{clim}} D_a(I_a + H_a). \tag{3}$$

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.** The risk-type label (Q2) determines which indicator each article contributes to: articles labeled *Transition* enter the transition-risk indicator, and articles labeled *Physical* enter the physical-risk indicator. Articles labeled *Both* (3% of the corpus) are excluded, as assigning them to both indicators would mechanically induce correlation between the two series; articles labeled *Unclear* (3% of the corpus) are excluded because their risk type cannot be reliably assigned.

### 2.2.2 Daily Indicators

From the article scores, I construct two daily newspaper-based indicators: a *Transition Risk Indicator* (TRI) and a *Physical Risk Indicator* (PRI). For the daily aggregation (steps (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 January 1, 2010 to June 30, 2025). 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 Eq. (3): a Transition Risk Score for articles labeled as transition risk, and a Physical Risk Score for articles labeled as physical risk: $\text{Tr.RiskScore}_a$ and $\text{Ph.RiskScore}_a$.

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

$$s_{o,t}^{\text{TRI}} = \sum_{a \in \mathcal{A}_{o,t}} \text{Tr.RiskScore}_a \quad \text{and} \quad s_{o,t}^{\text{PRI}} = \sum_{a \in \mathcal{A}_{o,t}} \text{Ph.RiskScore}_a. \tag{4}$$

**(ii) Outlet–level normalization.** I then normalize the aggregated score by each outlet's historical standard deviation, estimated over the pre-sample period January 1, 2010 to December 31, 2013[4]:

$$\sigma_o^{\text{TRI}} = \text{sd}\big(\{s_{o,t}^{\text{TRI}}\}_{t\leq 2013\text{-}12\text{-}31}\big) \quad \text{and} \quad \sigma_o^{\text{PRI}} = \text{sd}\big(\{s_{o,t}^{\text{PRI}}\}_{t\leq 2013\text{-}12\text{-}31}\big)\,, \tag{5}$$

and scale:

$$\tilde{s}_{o,t}^{\text{TRI}} = \frac{s_{o,t}^{\text{TRI}}}{\sigma_o^{\text{TRI}}} \quad \text{and} \quad \tilde{s}_{o,t}^{\text{PRI}} = \frac{s_{o,t}^{\text{PRI}}}{\sigma_o^{\text{PRI}}}. \tag{6}$$

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. Computing the standard deviations over a fixed pre-sample window ensures that the normalization incorporates no information from the asset pricing estimation period (January 1, 2014 to July 1, 2025): up to the outlet exceptions discussed in footnote 4, the indicators can therefore be interpreted as real-time measures, free of look-ahead bias from this step.

**(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:

$$\text{TRI}_t^{\text{raw}} = \frac{1}{|\mathcal{O}_t|}\sum_{o\in\mathcal{O}_t} \tilde{s}_{o,t}^{\text{TRI}} \quad \text{and} \quad \text{PRI}_t^{\text{raw}} = \frac{1}{|\mathcal{O}_t|}\sum_{o\in\mathcal{O}_t} \tilde{s}_{o,t}^{\text{PRI}}, \tag{7}$$

where $\mathcal{O}_t \subseteq \mathcal{O}$ is the set of outlets active at date $t$. This equal-weighted aggregation ensures that the indicators are not dominated by high-volume outlets: each outlet contributes equally, irrespective of how many climate-related articles it publishes.

[4]The *New York Post* (transition risk series) and the *New York Post*, the *Los Angeles Times*, and *Politico* (physical risk series) have fewer than 90 days with relevant coverage over 2010–2013; for these outlet series, the standard deviation is computed over the full sample period (2010–2025) instead. Since this scaling constant is a single time-invariant scalar per outlet, it only sets the outlet's relative weight in the cross-outlet average; computing all constants over the full sample, or strictly pre-2014 where feasible, leaves all results virtually unchanged.

**(iv) Calendar completion.** Then, I complete the daily calendar between the minimum and maximum sample dates, yielding the completed series $\mathrm{TRI}_t^{\mathrm{f}}$ and $\mathrm{PRI}_t^{\mathrm{f}}$. Days with no climate-related publications are interpolated using a Kalman smoother. Over the period January 1, 2014 to June 30, 2025, only 58 out of 4,199 days (1.4%) require interpolation, ensuring that this step has no material impact on the empirical results.

**(v) Damping of extreme coverage.** Finally, following the rationale of Ardia et al. (2023), I dampen days with extreme coverage, reflecting the idea that additional media coverage raises the perceived signal at a decreasing rate. Ardia et al. (2023) implement this damping with the square-root function, which is defined only for non-negative values. Because the directional indicators take negative values on days dominated by risk-decreasing news, I instead apply the inverse hyperbolic sine to the filtered series scaled by its pre-sample standard deviation:

$$\mathrm{TRI}_t = \operatorname{asinh}\left(\frac{\mathrm{TRI}_t^{\mathrm{f}}}{c_{\mathrm{TRI}}}\right), \quad \mathrm{PRI}_t = \operatorname{asinh}\left(\frac{\mathrm{PRI}_t^{\mathrm{f}}}{c_{\mathrm{PRI}}}\right), \tag{8}$$

with $\operatorname{asinh}(x) = \ln\bigl(x + \sqrt{x^2+1}\bigr)$, where $\mathrm{TRI}_t^{\mathrm{f}}$ and $\mathrm{PRI}_t^{\mathrm{f}}$ denote the completed series from step (iv) and the scaling constants $c_{\mathrm{TRI}}$ and $c_{\mathrm{PRI}}$ are the standard deviations of the corresponding series, estimated over the same pre-sample window as in step (ii) (January 1, 2010 to December 31, 2013). The inverse hyperbolic sine is defined on the whole real line, strictly increasing, and odd: it preserves the sign and ordering of the indicator, leaves typical days approximately unchanged ($\operatorname{asinh}(x) \approx x$ for small $|x|$), and compresses large positive and large negative spikes logarithmically. Estimating $c_{\mathrm{TRI}}$ and $c_{\mathrm{PRI}}$ over the pre-sample window keeps this final step, like the outlet-level normalization, free of look-ahead bias.

# 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 160,914 newspaper articles published across six major outlets between January 1, 2010 and June 30, 2025. 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, and currently maintained by 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; articles containing these expressions without being primarily related to climate change are subsequently filtered out by the LLM (see Section 2).

Outlet selection follows an explicit, ex-ante rule, documented in Table IA1.1 of the Internet Appendix. Starting from the 50 most-visited news websites in the United States (Similarweb, as compiled by Press Gazette), I retain outlets that are (i) newspapers, and (ii) national in audience, thereby excluding aggregators, newswires, online publications, magazines, and platforms. Among eligible outlets, I collect all those whose full text is systematically retrievable through MediaCloud; eligible outlets that fail this technical requirement are documented in the Internet Appendix (Section IA1).[5] The two U.K. outlets in the final corpus, the *Daily Mail* and *The Guardian*, enter through the same audience-based ranking: both figure among the 50 most-visited news websites in the United States, indicating a substantial American readership.[6]

This procedure yields six outlets: the *Daily Mail*, *Los Angeles Times*, *New York Post*,

[5]For instance, the *Wall Street Journal* meets all eligibility criteria but is not indexed by MediaCloud, which prevents systematic article retrieval. *USA Today* also meets the eligibility criteria and is indexed, but its URLs are RSS redirect links whose underlying articles are no longer accessible, including through the Internet Archive's Wayback Machine.

[6]The main results of the paper hold when the indicators are constructed from U.S.-only newspapers (see Internet Appendix Section IA6).

*Politico*, *The Guardian*, and *The Washington Post*. The resulting corpus is substantially broader than those used in prior studies of climate risk in financial markets: Engle et al. (2020) rely exclusively on the *Wall Street Journal*, and Faccini et al. (2023) use only Reuters newswires. It includes both broadsheet and tabloid formats and outlets with both Republican- and Democratic-leaning audiences, as measured by the Pew Research Center audience composition data reported in Table IA1.1.

I then web-scrape the full text of each article, yielding a final corpus of 160,914 documents. The sample starts in 2010, reflecting limitations in the stability and accessibility of URLs for older articles, which prevents reliable large-scale retrieval prior to this date.

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, 38.55% of the retrieved articles are excluded by the LLM filter, leaving 98,877 climate-related articles for constructing the indicators.

[Table 2 about here.]

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 the coverage of physical climate-related events.

By contrast, transition risk scores exhibit more variation across outlets. The *Washington Post*, *Los Angeles Times*, and *Politico* display the lowest average transition risk scores, with means below 1. As seen in Figure 2, most outlets record negative yearly average transition-risk scores during 2016–2018, a period of major U.S. federal climate-policy rollbacks following the election of Donald Trump. Importantly, this pattern is not confined to outlets with left-leaning audiences: the most negative scores in 2017 appear in the *Washington Post* and the *New York Post* alike, consistent with the direction dimension capturing the content of regulatory news—rollbacks reduce transition risk for polluting firms—rather than the editorial stance of the outlet reporting them. Appendix B complements this evidence with the monthly TRI and PRI series of each outlet, offering a finer view than the yearly averages of Figure 2: the 2016–2018 decline is visible across all six outlets, with varying magnitudes. This heterogeneity in levels motivates the normalization and equal-weighted aggregation described in Subsection 2.2. It also implies that publication volume does not drive the aggregate indicators: although the two U.K. outlets account for roughly 65% of collected articles, each outlet enters the daily aggregation with equal weight after normalization by its own pre-sample standard deviation, so that no outlet's volume mechanically dominates the series.

[Figure 2 about here.]

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 3 plots the monthly number of climate-related articles. Prior to 2014, coverage is sparse, with 63 articles per month on average; volume rises markedly thereafter, averaging 694 articles per month over 2014–2025. The TRI and PRI time series are therefore constructed over the period January 1, 2014 to June 30, 2025, avoiding the noise associated with sparse media coverage in the earlier period. The pre-2014 data are not discarded: they serve as the pre-sample window over which the outlet-level standard deviations are computed (Subsection 2.2), so that the normalization of the indicators relies exclusively on information available before the estimation sample begins.[7]

[Figure 3 about here.]

This sample choice also aligns with the timing of climate risk materiality documented in the literature: Faccini et al. (2023) find climate-related asset pricing effects only after 2012, and Bua et al. (2024) only from 2015 onward, following the Paris Agreement. The start of the estimation sample therefore reflects both the density of media coverage and the period over which climate risk plausibly became financially material.

Article volume also exhibits sharp peaks around major climate-policy events, most notably the June 2017 announcement of the intended U.S. withdrawal from the Paris Agreement and COP26 in late 2021 (Figure 3), providing a first indication that the corpus tracks the climate-policy news cycle.

#### 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,

[7] This choice is innocuous: computing the standard deviations over the full sample instead yields virtually identical indicators and leaves all empirical results unchanged.

and (v) time horizon. Figure 4 illustrates the distribution of these labels through an alluvial diagram.

Among articles classified as transition-related, most are labeled as reflecting an *increasing* level of transition risk, with a non-trivial share labeled 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.

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

[Figure 4 about here.]

Overall, three quarters of the climate-related articles are labeled as indicating an *increase* in climate risk, and more than half are labeled 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 5 about here.]

Figure 5a displays the monthly average of the transition risk indicator (TRI) from January 2014 to June 2025, 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 a sequence of major U.S. political events: the election of Donald Trump, the appointment of Scott Pruitt as EPA Administrator, Executive Order 13783, which rolled back a number of federal environmental regulations, and the announcement of the U.S. withdrawal from the Paris Agreement. After mid-2017, the indicator stabilizes around zero and remains subdued until the Democratic Party regains control of the U.S. House of Representatives in the November 2018 midterm elections. From 2019 onward, the TRI monthly average fluctuates around 1, 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 declines sharply 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 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-based approach, 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 6 about here.]

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 Democratic administration that followed (January 2021 until the November 2024 election), the AT closely co-moves with the TRI. In contrast, during the periods opened by Republican electoral victories (from the November 2016 elections to the seating of the new Democratic House in January 2019, and from the November 2024 elections 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. The asset-pricing consequences of this distinction are examined in Section IA4 of the Internet Appendix.

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.

[Table 3 about here.]

Table 3 closes this section by relating the four indicators to one another and to existing measures. Consistent with the visual evidence above, the TRI is essentially uncorrelated with its attention counterpart in levels (0.03), whereas the PRI and AP are highly correlated (0.85): direction is the key dimension for transition risk and a negligible one for physical risk. The table also addresses a natural concern, namely that the TRI and PRI may merely repackage information already contained in existing news-based climate indices.[8] The data suggest otherwise. The existing indices are substantially correlated with one another: in levels, the pairwise correlations between the MCCC and the topic-specific indices of Faccini et al. (2023) range from 0.15 to 0.53, consistent with these measures capturing a common climate-news component. By contrast, the TRI and PRI correlate only weakly with all of them, at most 0.21 and 0.27, indicating that the directional indicators embed information absent from existing measures, plausibly the direction, intensity, and horizon dimensions introduced in Section 2. The same reading holds in innovations: correlations among existing indices reach 0.41, whereas none involving the TRI or PRI exceeds 0.12. Revealingly, the AT correlates with the policy-oriented indices (0.2) about as much as they correlate with one another, even though it already embeds the intensity and horizon dimensions: what these measures share is attention, and what sets the TRI apart is direction.

[8] Monthly indicators, such as those of Engle et al. (2020), Gavriilidis (2021), and Bessec and Fouquau (2024) (weekly), cannot be compared at the daily frequency used throughout the paper.

## 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 domestic common stocks listed on the NYSE, NASDAQ, or NYSE American. 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 4,539 unique firms over the January 2014–July 2025 period. On average, 2,218 firms have available returns on a given trading day, and 2,222 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), and the investment (CMA) factors, as well as the momentum factor (MOM) of Carhart (1997). 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, default spread, and the one-month Treasury bill rate; 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 4 reports the daily pairwise correlations between the variables used in the empirical analysis. The TRI, PRI, and EPU enter the analysis in innovations, measured as daily first differences, consistent with their use in the asset-pricing tests (Subsection 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 4 reveals no substantial correlation patterns that would compromise the results in the subsequent asset-pricing tests.

[Table 4 about here.]

# 4 Asset Pricing Empirical Results

## 4.1 Methodology

The methodology follows four steps to test whether transition and 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) of Merton (1973). First, I extract daily innovations from the two indicators. Second, I form portfolios sorted on stocks' exposure to these innovations. Third, I estimate the associated risk premia using the two-pass Fama–MacBeth methodology. Fourth, I confront the evidence with the testable ICAPM restrictions of Maio and Santa-Clara (2012).

### 4.1.1 Transition and Physical Risk Innovations

Following the ICAPM (Merton, 1973), candidate state variables enter the asset-pricing analysis in innovations rather than in levels. I measure daily innovations as first differences of the indicators over the sample period from January 1, 2014 to June 30, 2025. Let $F_t$ denote the TRI or PRI:

$$\Delta F_t = F_t - F_{t-1}. \tag{9}$$

Throughout the asset-pricing analysis, $\Delta$ denotes daily first differences, and $\Delta$TRI and $\Delta$PRI are the transition- and physical-risk shocks entering the tests.

First differencing has two advantages in the present setting. First, it involves no parameter estimation: the innovation at date $t$ is computed from information available at dates $t-1$ and $t$ only, so the series is free of look-ahead bias by construction and can be interpreted as a real-time measure of climate-risk shocks. Second, the two-pass methodology below already relies on estimated quantities, namely the rolling first-stage betas; extracting innovations without estimation avoids compounding estimation errors across stages. Because the indicators are highly persistent at the daily frequency, first differences also closely approximate their unexpected component: for a near-unit-root process, the first difference coincides with the innovation up to a term proportional to the deviation of the autoregressive root from one.

The main alternative in the literature extracts innovations as residuals from an autoregressive process. Engle et al. (2020) fit an AR(1) to their monthly index over the full sample, and Bua et al. (2024) follow a similar approach. In my daily setting, full-sample estimation raises two concerns. The autoregressive coefficients depend on observations through the end of the sample, so the residual at date $t$ embeds future information, introducing look-ahead bias into the innovation series. In addition, the estimation error in the coefficients propagates into the innovations, and from there into the betas estimated on them, compounding the errors-in-variables problem inherent in the Fama and MacBeth (1973) procedure. A rolling implementation, as in Ardia et al. (2023), addresses the first concern: with coefficients estimated on past observations only, the innovations are free of look-ahead bias. It does not, however, address the second, and it consumes data: the estimation window is lost as burn-in, delaying the start of the innovation series, and of every downstream estimation window, by its length. This cost is material in a sample that is already short by asset-pricing standards. First differencing is thus the only approach that is simultaneously free of look-ahead bias, immune to parameter estimation error, and costless in observations.[9]

[9]The main results are unchanged when innovations are instead extracted as residuals from an AR(8)

### 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 indicators. Specifically, firms are sorted according to their estimated sensitivity to innovations in the transition (or physical) risk indicator. 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 $\Delta F_t$ denote the innovation of the textual indicator $F$ (either TRI or PRI), as defined in Eq. (9). For each asset $i$, I estimate the following regression using daily data over a rolling window of 12 months, updated monthly:

$$r_{i,t} - r_{f,t} = c_i + \beta_i \Delta F_t + \boldsymbol{\gamma}_i' \boldsymbol{X}_t + \varepsilon_{i,t}, \tag{10}$$

where $\boldsymbol{X}_t$ denotes a vector of control factors. I consider three specifications: the Fama and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Fama and French (2015) five-factor model. This procedure yields, at the end of each month, an estimate $\hat{\beta}_i$ of each asset's exposure to innovations in the textual risk indicator.

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 and French (1993) three-factor model, (ii) the Carhart (1997) four-factor model, and (iii) the Fama and French (2015) five-factor model. In all cases, factor loadings are estimated consistently with the first-stage specification.

process, with the lag order selected by minimizing the Bayesian Information Criterion.

#### 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 diversified test assets rather than individual stocks. Following the prescription of Lewellen et al. (2010) to expand the set of test portfolios beyond size and book-to-market sorts, I estimate the model on four universes: the 30 and 49 industry portfolios from the Kenneth R. French Data Library ($N = 30$ and $N = 49$), and each industry set augmented with the 25 size–book-to-market portfolios ($N = 55$ and $N = 74$), following Faccini et al. (2023). Industry portfolios are a natural choice of test assets: transition risk exposure is largely determined at the industry level, so they should generate substantial dispersion in climate betas. Moreover, Lewellen et al. (2010) show that the tight factor structure of the size–book-to-market portfolios can make cross-sectional fits look deceptively strong; combining them with industry portfolios breaks this structure, while varying the cross-sectional dimension provides a direct diagnostic against the small-sample distortions discussed below.

First-stage betas are estimated from 24-month rolling windows of daily data, updated monthly, yielding roughly 500 observations per window. The window is longer than in the portfolio sorts for two reasons. First, test assets are diversified portfolios whose factor exposures evolve more slowly than those of individual firms, so lengthening the window sacrifices little in terms of tracking time variation. Second, measurement error in first-stage betas directly contaminates second-stage risk prices, and estimation precision rises with window length: Bali et al. (2016, Ch. 8) document that the cross-sectional distribution of betas estimated from daily data tightens markedly as the window lengthens, with far fewer extreme estimates at the 24-month horizon.

In the second stage, I estimate a cross-sectional regression each month:

$$r_{i,t+1} - r_{f,t+1} = c_t + \boldsymbol{\lambda}'_{\Delta F,t}\boldsymbol{\beta}^{\Delta F}_{i,t} + \boldsymbol{\lambda}'_{X,t}\boldsymbol{\beta}^{X}_{i,t} + \varepsilon_{i,t+1}, \tag{11}$$

where $\boldsymbol{\beta}^{\Delta F}_{i,t}$ collects portfolio $i$'s loadings on innovations in the transition and physical risk indicators, estimated over the 24-month window ending in month $t$, and $\boldsymbol{\beta}^{X}_{i,t}$ collects the loadings on the control factors. The timing convention is thus strictly predictive: betas estimated with information through month $t$ are paired with the excess return of month $t+1$.[10] The reported risk prices are the time-series averages of the monthly estimates $\hat{\boldsymbol{\lambda}}_{t+1}$.

The controls include the Fama and French (1993) three factors—the excess market return (MKT), the size factor (SMB), and the value factor (HML)—the momentum factor (MOM) (Carhart, 1997), and first-differenced innovations in the Economic Policy Uncertainty (EPU) index (Baker et al., 2016).

Two features of this design deserve comment. First, betas are estimated on rolling windows rather than once over the full sample. This choice reflects the economic reality that exposures to transition risk vary over time, as firms alter their energy mix, face shifting regulation, and reposition strategically; constant loadings estimated over the full sample would be misspecified. The cost is that the analytical errors-in-variables corrections available in the classic two-pass setting do not carry over: the Shanken (1992) adjustment and the bootstrap procedures of Gospodinov and Robotti (2021) are derived under betas that are constant and estimated once over the full sample, so applying them mechanically to loadings re-estimated each month from rolling windows would not be justified by their underlying theory. Fama–MacBeth standard errors remain informative in this design. Because betas are re-estimated each month, first-stage estimation error varies over time and contributes to the month-to-month variation in the estimated risk prices. It is therefore reflected, at least partially, in the Fama–MacBeth standard errors. Moreover, longer estimation windows reduce the variance of first-stage estimation errors, mitigating attenuation bias arising from errors-in-variables. To the extent that the remaining estimation error behaves as classical measurement error, it is expected to attenuate estimated risk prices toward zero rather than mechanically inflate them.

Second, the number of monthly cross-sections ($T = 115$) is small relative to the largest cross-section ($N = 74$), precisely the conditions under which Kleibergen and Zhan (2020) document overrejection of Fama–MacBeth $t$-tests. The four test-asset universes span $T/N$ ratios from roughly 1.6 ($N = 74$) to 3.8 ($N = 30$): if significance were an artifact of the small-$T$ problem, it should be confined to the largest cross-sections and vanish in

[10]The second-stage sample runs from January 2016 (betas from the first 24-month window, January 2014–December 2015, paired with the January 2016 return) to July 2025 (last window, January 2023–June 2025), for a total of 115 monthly cross-sections.

the smaller universes where these distortions are attenuated. Section 4.2 shows that it is not.

#### 4.1.4 ICAPM Testable Implications

The fourth step assesses whether the transition and physical risk indicators can be interpreted as state variables in the ICAPM (Merton, 1973). Three testable implications of the model must be satisfied. First, the climate risk indicators should forecast future investment opportunities (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, with a sign consistent with the direction of predictability: if a candidate state variable predicts a deterioration in the investment opportunity set, its innovations should carry a negative price of risk. Third, the estimated price of risk on the market factor should be economically plausible when interpreted as the relative risk aversion coefficient of the representative investor (Maio and Santa-Clara, 2012).

The latter two criteria can be assessed with the two-stage Fama–MacBeth methodology described above. To evaluate the first implication, I estimate predictive regressions of aggregate stock market returns, which proxy for changes in the investment opportunity set, following Maio and Santa-Clara (2012):

$$r_{t,t+q} = a_q + \beta_q z_t + \varepsilon_{t,t+q}, \tag{12}$$

where $r_{t,t+q} \equiv r_{t+1} + \ldots + r_{t+q}$ denotes the continuously compounded market return over $q$ periods, $z_t$ is the candidate state variable in levels (either the TRI or the PRI) at date $t$, $\varepsilon_{t,t+q}$ is the error term, and the forecasting horizon $q \in \{1, 3, 12, 24, 36, 48\}$ months.

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:

$$r_{t,t+q} = a_q + \beta_q^{TRI} TRI_t + \beta_q^{PRI} PRI_t + \boldsymbol{\gamma}_q' \boldsymbol{X}_t + \varepsilon_{t,t+q}, \tag{13}$$

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, the TRI and PRI enter in levels rather than as innovations, consistent with the ICAPM logic: the level of a state variable summarizes the current state of investment opportunities and is thus the relevant forecasting object, whereas its innovations capture the news that matters for cross-sectional pricing.

## 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 5 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 negative in the lowest quintile (Q1) to positive in the highest quintile (Q5). The number of firms is balanced across quintiles, and there is no monotonic pattern in market capitalization, suggesting that the sorting is not mechanically driven by firm size or sample composition.

The firm characteristics reveal a symmetry: the extreme quintiles resemble each other rather than lining up along a gradient. Firms in Q1 and Q5 are on average smaller, younger, less profitable, and faster-growing than firms in the intermediate quintiles, while book-to-market, investment, and cash flow are essentially flat across the sort. Because these differences are common to both legs of the sort, they largely cancel in the long–short portfolio: exposure to transition-risk shocks does not reduce to a monotonic loading on any standard characteristic. The only exceptions are leverage and asset tangibility, which rise modestly from Q1 to Q5.

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 and French (1993, 2015) size, value, profitability, investment and Carhart (1997) momentum factors. Overall, this suggests that the TRI-sorted strategy is not primarily driven by conventional risk factors. Consequently, the abnormal performance documented in Table 6 is unlikely to be fully explained by standard factor exposures, and is instead consistent with compensation related to transition-risk shocks.

[Table 5 about here.]

Table 6 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 and French (1993) three-factor model, the Carhart four-factor model (Carhart, 1997), and the Fama and French (2015) five-factor model.

[Table 6 about here.]

The results reveal that estimated alphas rise from negative values in the low-exposure quintiles to positive values in the top quintile. The strategy that is long the high-exposure quintile and short the low-exposure quintile yields a spread alpha of 0.42% per month ($t = 1.82$) under the Fama and French (1993) three-factor model and 0.73% per month ($t = 1.82$) under the Carhart (1997) four-factor model, statistically significant at the 10% level in two of the three specifications, and economically meaningful, corresponding to roughly 5% to 9% on an annualized basis.

The 2015–2025 period is characterized by frequent adverse transition-risk realizations associated with major climate policy developments, such as the COP21 Paris Agreement and the Inflation Reduction Act. Over this sample, the transition risk indicator therefore predominantly reflects a *bad state* of the world for firms subject to environmental regulation. Consistent with this interpretation, the Q5–Q1 portfolio (long the assets that benefit from transition-risk shocks, short those that suffer from them) exhibits positive abnormal returns: it provides a hedge against increases in transition risk, and therefore outperforms during periods in which adverse shocks materialize frequently. These findings are consistent with the equilibrium mechanism of Pástor et al. (2021), in which brown assets underperform green assets during periods of heightened climate risk, and with the evidence in Pástor et al. (2022) that the realized outperformance of green assets is driven by the realization of climate-concern shocks rather than by their expected returns.

The fact that Q5 alphas are statistically significant only under the three-factor specification suggests that this hedging channel is present but not dominant, part of the performance being absorbed by standard risk factors.

The sub-period from November 2016 to December 2018—which begins with the election of Donald Trump and ends with the loss of the Republican House majority at the November 2018 midterms—is characterized by a temporary easing of U.S. climate policy. During this window, the transition risk indicator predominantly reflects a *good state* of the world from the perspective of a polluting firm, with declining transition risk. Consistent with the intertemporal hedging mechanism, the sign of abnormal returns reverses: the high-exposure portfolio (Q5) exhibits negative and significant 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. This sub-period analysis nevertheless relies on a limited number of observations (26 months) and should be interpreted as illustrative rather than conclusive.

Taken together, the full-sample and sub-period evidence supports a state-dependent interpretation of the abnormal returns on transition-risk-sorted portfolios: the sign of realized abnormal returns depends on whether the period is dominated by the tightening or the rollback of climate policy. This state dependence also helps explain the moderate statistical significance over the full sample: although the 2015–2025 period is dominated by tightening episodes, it includes both states, which pulls the average abnormal return toward zero.

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

Regarding physical risk, Table 7 reports the estimated alphas of the PRI-sorted quintile portfolios and of the corresponding long–short portfolio. The long–short alpha is statistically indistinguishable from zero in all three specifications, and the few marginally significant quintile alphas display no monotonic pattern across the sort. Physical-risk exposure therefore does not generate an economically interpretable abnormal return, consistent with the existing literature (Faccini et al., 2023).

[Table 7 about here.]

### 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 8 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 $-4.348$ ($t = -2.27$) to $-6.941$ ($t = -2.64$) depending on the specification, and the premium is estimated more precisely when PRI innovations are also controlled for. In contrast, exposure to PRI innovations is not significantly priced in any specification, a result that remains unchanged across all robustness checks reported in Section IA3 of the Internet Appendix.

[Table 8 about here.]

Figure 7 examines the stability of the transition-risk premium across the four test-asset universes and first-stage rolling windows ranging from 6 to 36 months. The point estimate is negative in every universe–window combination, and becomes larger and more precisely estimated as the window lengthens, consistent with the errors-in-variables attenuation that affects betas estimated over short windows, as discussed in Section 4.1. Importantly, significance is not confined to the largest cross-sections: the premium is estimated at least as precisely in the 30- and 49-portfolio universes, where the small-sample distortions documented by Kleibergen and Zhan (2020) are attenuated, as in the 55- and 74-portfolio universes. The pattern that would arise if significance were an artifact of the small-$T$ problem is therefore absent. The estimates are similarly robust to alternative factor controls (FF3 and FF5, with and without EPU innovations), as reported in Section IA3 of the Internet Appendix.

[Figure 7 about here.]

Two diagnostics support the specification. First, the second-stage intercepts are economically small and statistically indistinguishable from zero in all eight columns, indicating that the model does not leave an unexplained level in average returns. Second, the cross-sectional fit is moderate but honest by the standards of Lewellen et al. (2010): the GLS $R^2$, the more demanding metric, ranges from 0.27 to 0.34 in the full specification, and adding the TRI beta raises it from 0.21 to 0.27 in the 74-portfolio universe.

These findings support an intertemporal hedging interpretation. Portfolios whose returns 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 in the full specification *(iv)*, at approximately 9.6% and 6% for the 74- and 55-portfolio universes. 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 2.33 to 3.73, within an economically reasonable range and consistent with the third ICAPM restriction of their framework.

Finally, to assess whether the transition-risk premium reflects information already contained in existing news-based climate measures, Table 9 re-estimates the second-stage regressions including the two most prominent alternatives: the Media Climate Change Concern index (MCCC) of Ardia et al. (2023) and the U.S. Climate Policy index (UCP) of Faccini et al. (2023). All indicator innovations are standardized to unit variance and estimated on a common daily sample, so that the risk premia are directly comparable across measures.

Neither competing measure commands a statistically significant premium on this sample, whether included alone (specifications (*ii*) and (*iii*)) or jointly with the directional indicators. By contrast, the premium on $\Delta$TRI survives every specification: in the 55-portfolio universe it remains significant at the 1% level throughout, and its precision, if anything, increases as the competing measures are added ($t = -2.61$ in the baseline against $t = -3.05$ with all four indicators); in the 74-portfolio universe it remains significant at least at the 10% level, and at the 5% level when only MCCC is added. Per one-standard-deviation shock, the transition-risk premium is also several times larger in absolute value than the premia attached to the competing measures, and adding the directional indicators to the competitor-only specifications improves the cross-sectional fit (in Panel A, the OLS $R^2$ rises from 0.58 to 0.65 and the GLS $R^2$ from 0.36 to 0.38 when $\Delta$TRI and $\Delta$PRI are added to $\Delta$MCCC).

[Table 9 about here.]

These results indicate that the pricing information contained in the directional transition-risk indicator is not subsumed by existing attention- or policy-based measures. The distinguishing feature of the TRI is its directional dimension: MCCC captures the intensity of media concern and UCP the salience of climate policy, but neither signs the news as increasing or decreasing risk. This distinction matters over a sample in which climate policy moves in both directions: when a given volume of coverage can signal either rising or falling transition risk depending on the episode, a measure that does not sign the news aggregates the two, and the resulting series carries no consistent risk content. This interpretation is consistent with the attention-only ablation reported in Section IA4 of the Internet Appendix, in which the undirected versions of my own indicators (AT and AP) are not priced either. The absence of a significant premium for MCCC and UCP on this sample does not contradict the original findings of Ardia et al. (2023) and Faccini et al. (2023), which are based on different sample periods and specifications; it indicates only that, over 2016–2025 and within this framework, the directional measure carries the priced component of climate news.

Taken together, the stability of the premium across universes, windows, and factor controls, the near-zero intercepts, and the economically plausible implied risk aversion indicate that the negative cross-sectional price of transition risk is robust and economically sensible. Whether this premium also satisfies the remaining ICAPM restriction—that the sign of the price of risk be consistent with the way the state variable forecasts aggregate market returns—is examined in the next subsection.

### 4.2.3 ICAPM Restrictions and Market Return Predictability

Lastly, I find that the transition risk indicator in levels negatively predicts aggregate market returns at long horizons, even after controlling for standard predictors used in the return predictability literature. The physical risk indicator also forecasts market returns, but without a consistent sign across horizons.

Table 10 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 (negative) and 24-month (positive) horizons.

For transition risk, when standard control variables are included, the predictive coefficients become more precisely estimated and retain the same negative sign, as reported in Table 11. 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 horizon. The sign, however, is not consistent across horizons: the PRI predicts a deterioration of the investment opportunity set at one month but an improvement at 48 months. This lack of sign consistency mirrors the absence of a significant cross-sectional price of physical risk, and reinforces the transition–physical asymmetry documented throughout the paper.

[Table 10 about here.]

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 negative 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 plausible 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 8).

When transition risk shocks materialize, these same assets deliver positive abnormal returns (Table 6), consistent with their role as hedging assets. Overall, this evidence supports the interpretation of transition risk as a relevant ICAPM state variable.

[Table 11 about here.]

Moreover, Section IA4 of the Internet Appendix 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 a coherent intertemporal transition-risk premium.

# 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 a large language model, 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, the directional indicator avoids the interpretive ambiguity of prior unidirectional approaches. This advantage is not merely conceptual: when the competing measures are evaluated within a common asset-pricing framework, the directional indicator is the only one that carries a significant risk premium; neither the climate news indices of Ardia et al. (2023) and Faccini et al. (2023) nor my own attention-based counterpart are priced in this setting.

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 eliminates both legs of the ICAPM evidence: exposure to attention shocks commands no significant risk premium, and the attention measure no longer forecasts aggregate market returns with a sign consistent with the cross-sectional estimates. Direction is thus essential for extracting a transition-risk signal from climate news coverage: without it, the indicator captures attention rather than risk.

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., 2026). 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's website (https://clementtoussaint.com). Firm-level data are from Sharadar Nasdaq Data Link. Fama-French factors and portfolios 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.

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.

Bali, Turan G, Robert F Engle, and Scott Murray, 2016, *Empirical Asset Pricing: The Cross Section of Stock Returns* (John Wiley & Sons).

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, *The 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, 2026, Extreme-weather risk and the cross-section of stock returns, *Journal of Risk and Insurance* 93, 163–198.

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* 104, 298–345.

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.

Chu, Liya, Xinrui Duan, Jun Tu, and Luying Wang, 2025, Media climate change concern and stock returns, *Financial Management* .

Cochrane, John H., 2009, *Asset Pricing: Revised Edition* (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.

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.

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.

Gospodinov, Nikolay, and Cesare Robotti, 2021, Common pricing across asset classes: Empirical evidence revisited, *Journal of Financial Economics* 140, 292–324.

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.

Jha, Manish, Jialin Qian, Michael Weber, and Baozhong Yang, 2024, ChatGPT and Corporate Policies.

Kleibergen, Frank, and Zhaoguo Zhan, 2020, Robust Inference for Consumption-Based Asset Pricing, *The Journal of Finance* 75, 507–550.

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.

Lopez-Lira, Alejandro, and Yuehua Tang, 2026, Can chatgpt forecast stock price movements? return predictability and large language models, *Journal of Financial Economics* 184, 104335.

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.

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.

Petkova, Ralitsa, 2006, Do the Fama–French Factors Proxy for Innovations in Predictive Variables?, *The Journal of Finance* 61, 581–612.

Shanken, Jay, 1992, On the Estimation of Beta-Pricing Models, *The Review of Financial Studies* 5, 1–33.

Tankov, Peter, and Alexis Tantet, 2019, Climate Data for Physical Risk Assessment in Finance.

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.

Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang, 2024, Can Large Language Models Transform Computational Social Science?

# A LLM Robustness Checks

[Table 12 about here.]

[Table 13 about here.]

# B Decomposition of the Climate Risk Indices by News Outlet

[Figure 8 about here.]

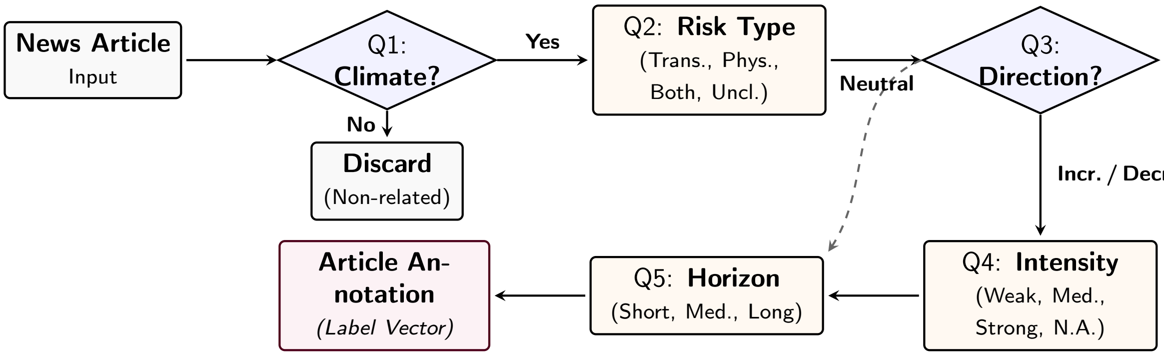



**Figure 1. LLM Annotation Workflow.** Each article is processed sequentially through five classification tasks. Articles unrelated to climate change (Q1 = No) are discarded. For climate-related articles, the model identifies the risk type (Q2), its direction (Q3), intensity (Q4), and time horizon (Q5). 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.

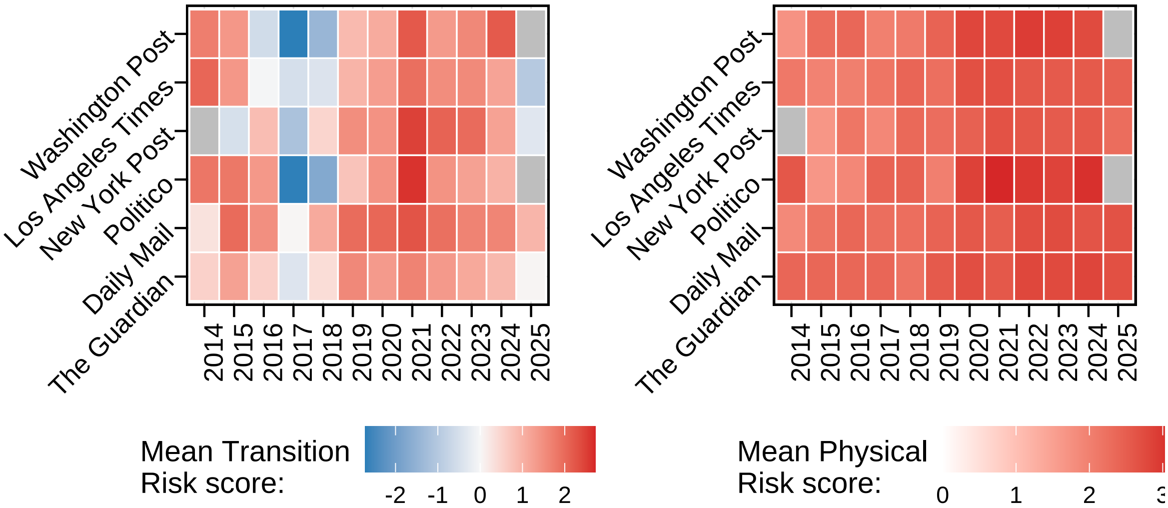



**Figure 2. Average Transition and Physical Risk Scores by Outlet and Year.** This figure displays the average article-level transition risk score (left panel) and physical risk score (right panel) by outlet and year. For the transition risk score, blue indicates negative values (risk-decreasing news) and red indicates positive values (risk-increasing news). For the physical risk score, darker red indicates higher values. Gray cells indicate missing data.

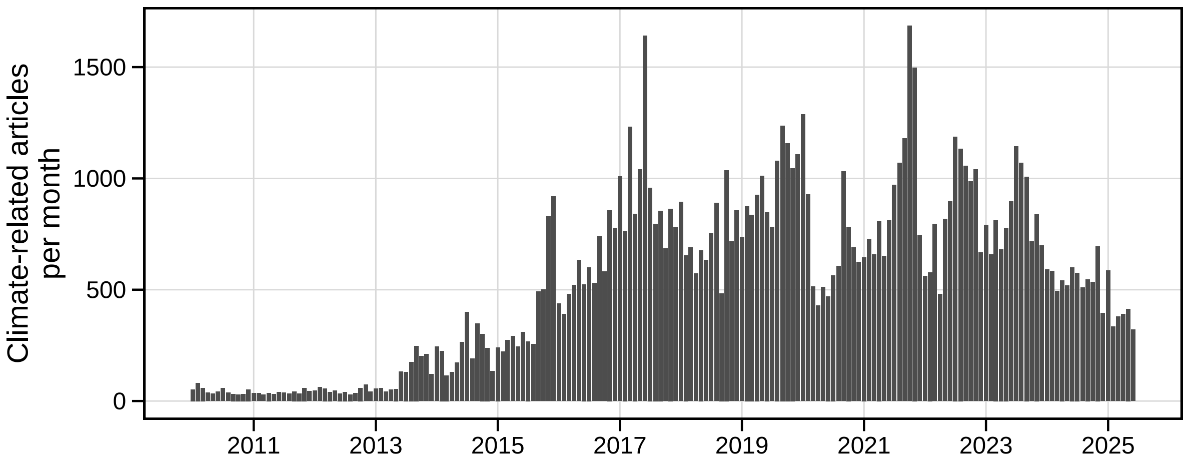



**Figure 3. Monthly Volume of Climate-Related Articles.** This figure displays the number of articles classified as climate-related by the LLM (Q1 = Yes) per month, aggregated across the six outlets of the corpus, from January 2010 to June 2025.

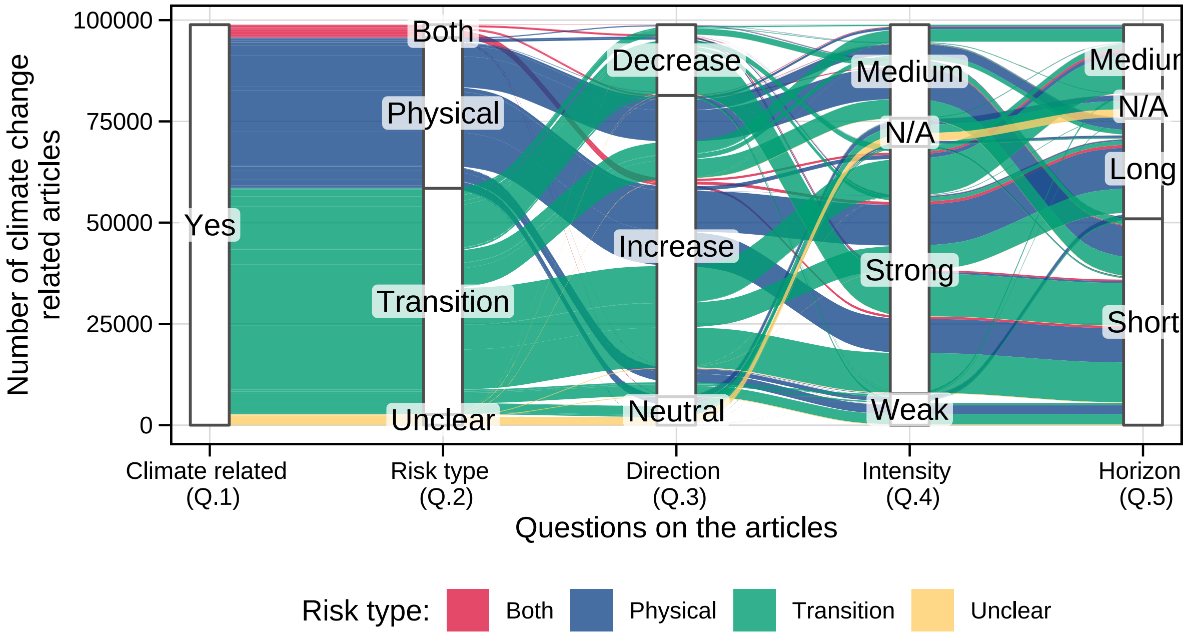
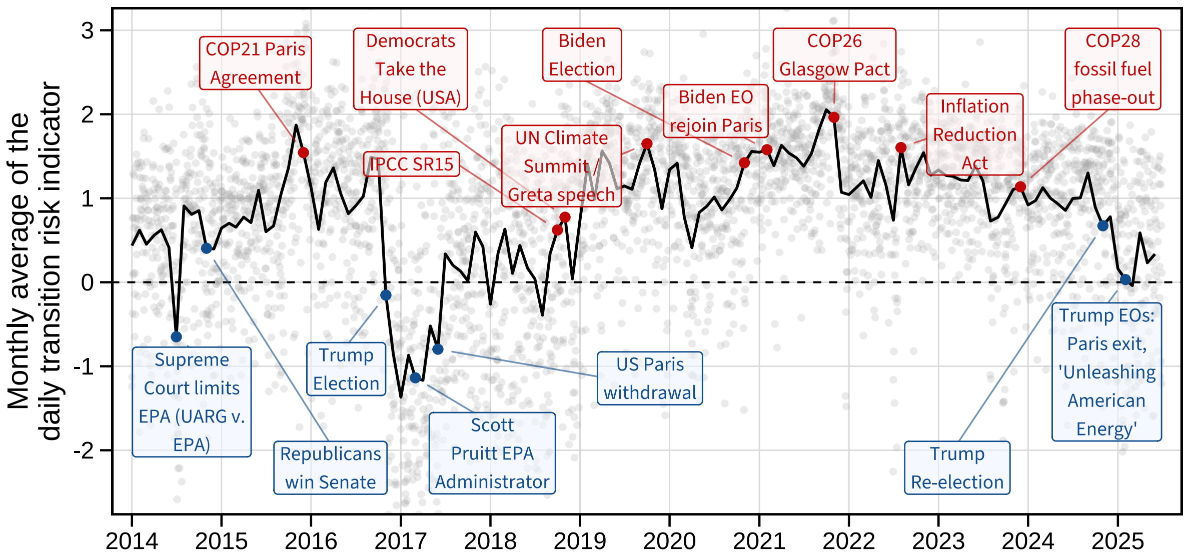
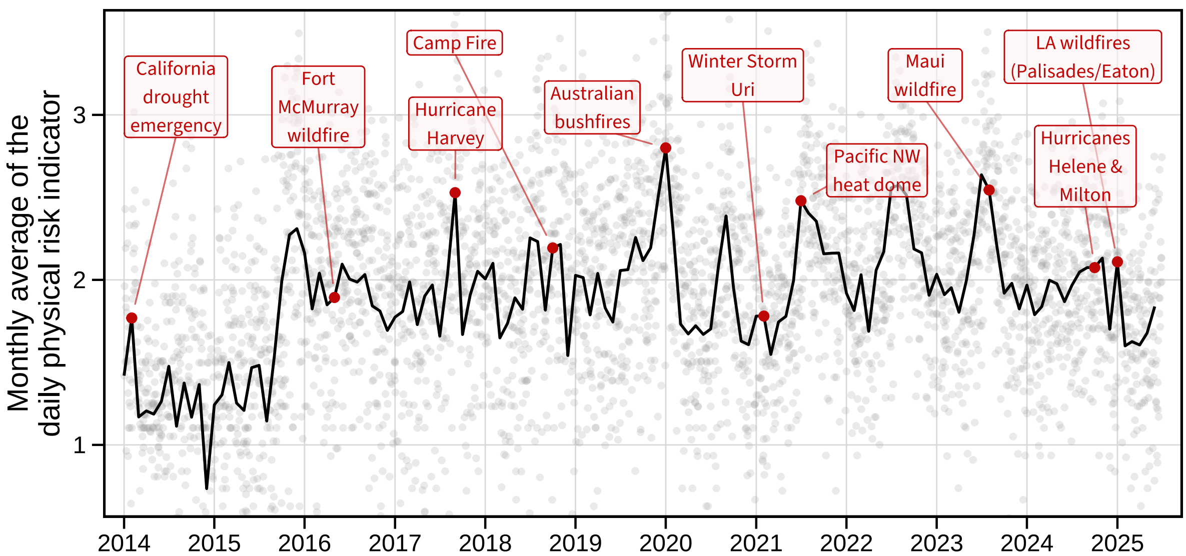



**Figure 4. Alluvial Representation of the LLM-Based Classification of News Articles.** This alluvial diagram illustrates the classification of articles across the five LLM questions, read from left to right. Each vertical bar represents a question: climate relevance (Q1), risk type (Q2), risk direction (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).

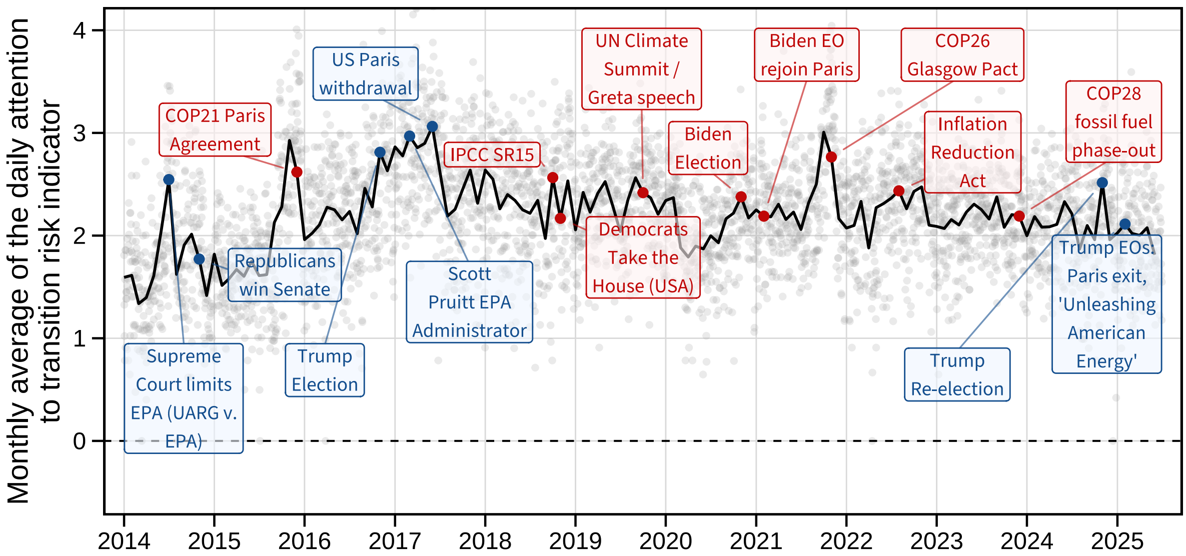



**(a) Transition Risk Indicator (TRI)**

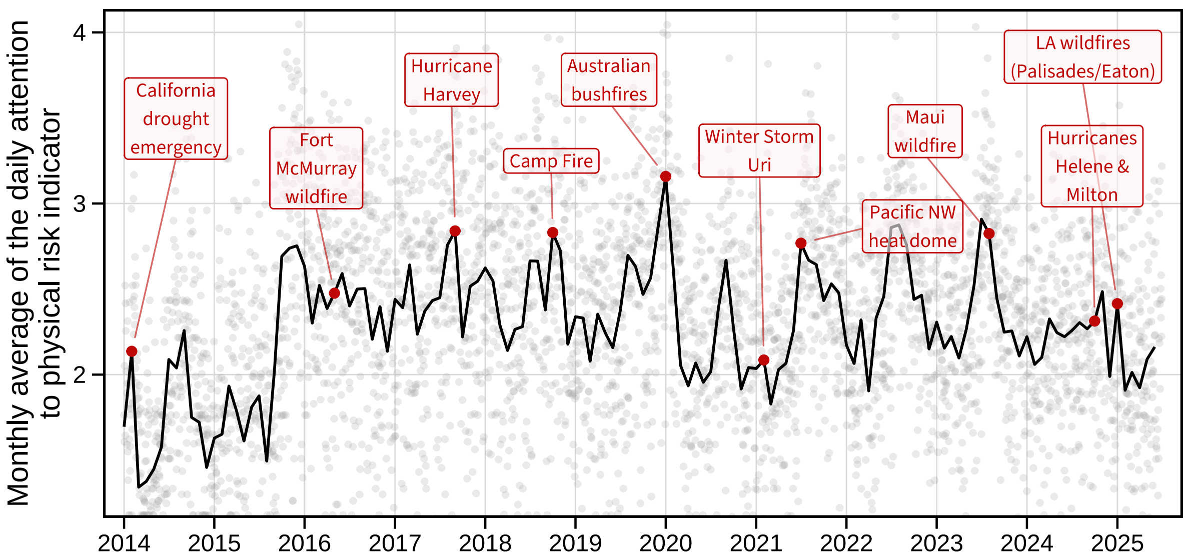



**(b) Physical Risk Indicator (PRI)**

**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 June 2025. Gray dots represent daily values; the black line shows the monthly average. Annotated markers indicate major climate-related events, colored by their expected effect on risk: red for risk-increasing events, blue for risk-decreasing events.




(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 June 2025. These indicators are constructed by setting the direction weight to one in Eq. (3), 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. Annotated markers indicate major climate-related events, colored by their expected effect on risk: red for risk-increasing events, blue for risk-decreasing events.

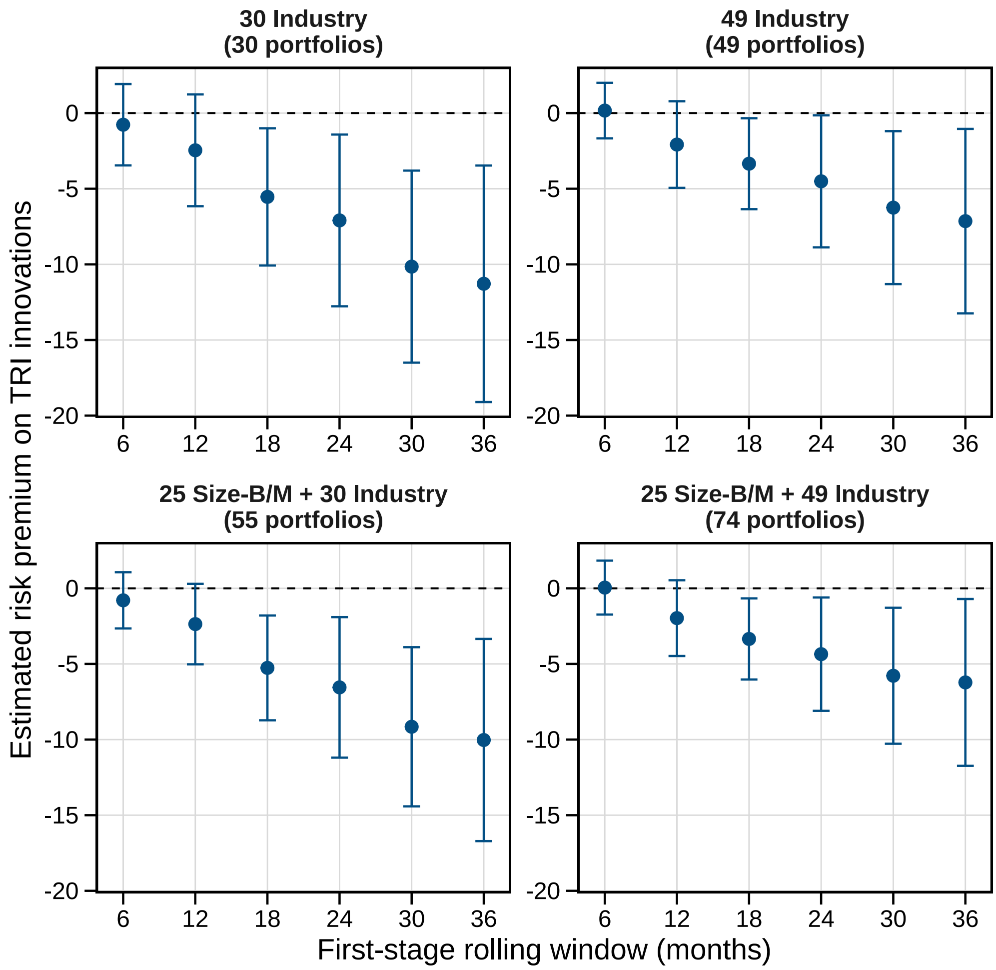
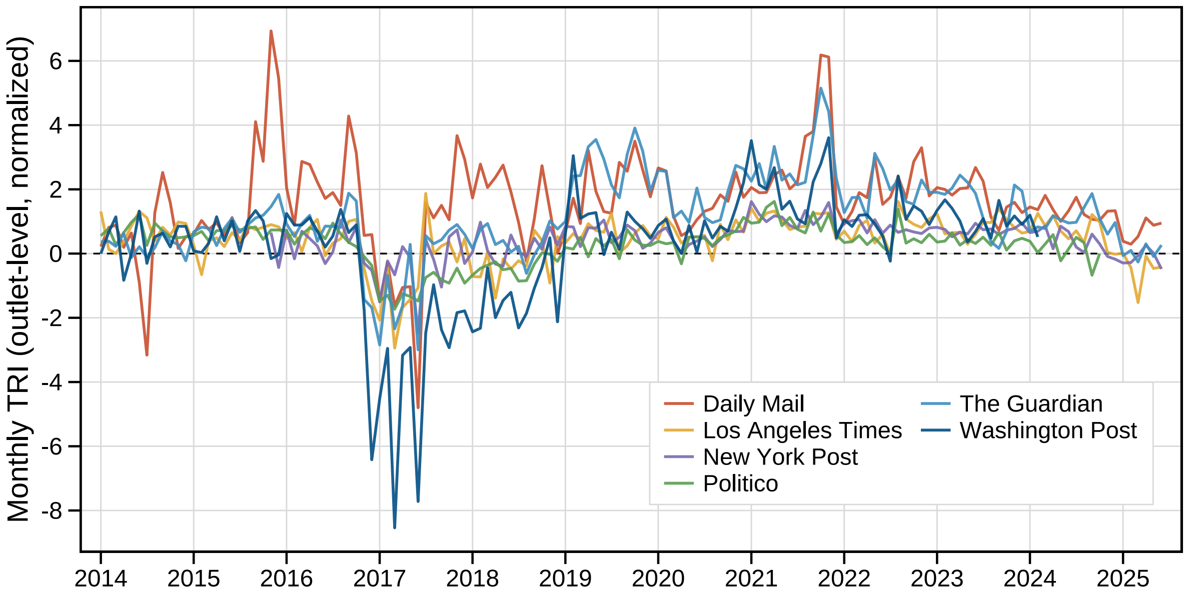
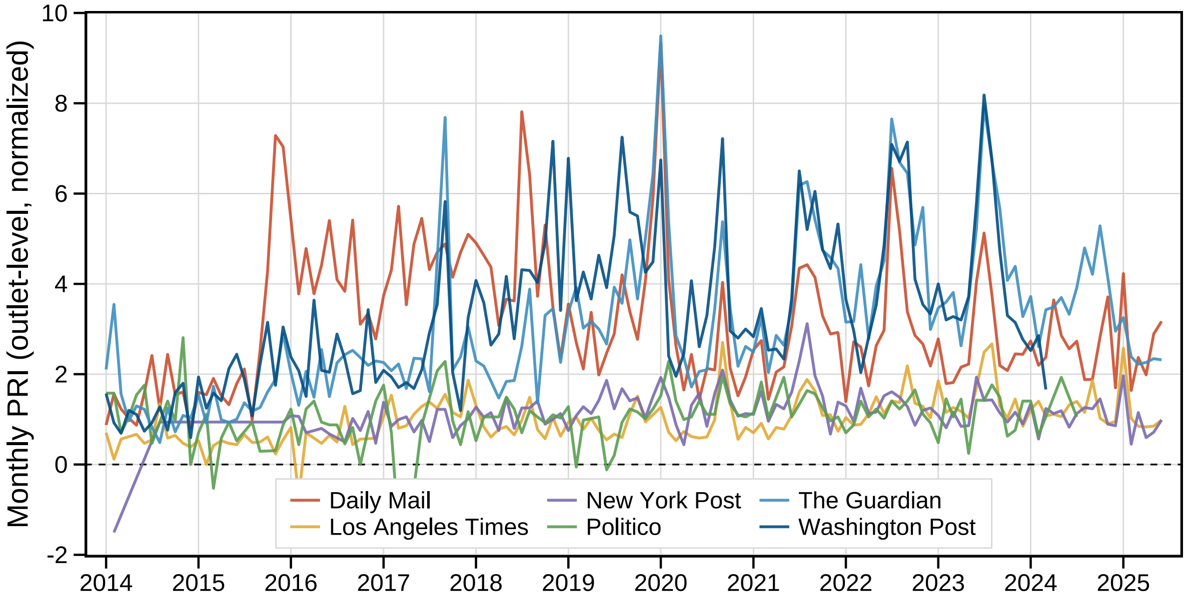



**Figure 7. Robustness of the TRI-Innovation Risk Premium to the Rolling-Window Length and Test-Asset Universe.** This figure plots the second-stage Fama and MacBeth (1973) estimate of the price of risk associated with exposure to transition-risk innovations as a function of the first-stage rolling-window length, for four test-asset universes. Each panel corresponds to a set of test portfolios: the 30 and 49 Fama–French industry portfolios, and their combinations with the 25 size-B/M portfolios (55 and 74 assets). The specification corresponds to column *(iv)* of Table 8: first-stage betas are estimated from daily excess returns over rolling windows of 6 to 36 months on $\Delta$TRI, $\Delta$PRI, the Fama and French (1993) and momentum (Carhart, 1997) factors, and $\Delta$EPU (Baker et al., 2016); the second stage includes all corresponding betas and a constant. $\Delta$ denotes daily first differences, as in Table 8. Each portfolio's betas, estimated over the window ending in month $t$, are paired with its excess return in month $t+1$. 95% confidence intervals are reported, based on Newey and West (1987) standard errors with 4 lags. For each window length, the second-stage sample begins with the first month for which a full window of first-stage betas is available (January 2015 for the 12-month window) and ends in July 2025.




(a) **Transition Risk Indicator (TRI) per Outlet**




(b) **Physical Risk Indicator (PRI) per Outlet**

**Figure 8. Monthly Climate Risk Indicators per Outlet.** Panels (a) and (b) display the monthly transition (TRI) and physical (PRI) risk indicators for each of the six outlets, from January 2014 to June 2025. Daily article scores are summed per outlet and normalized by the outlet's pre-2014 standard deviation; monthly values are averages of the normalized daily scores. The aggregate indicators are the equal-weighted means of these series, after the Kalman filter of Section 2. Series are thus in comparable standard-deviation units across outlets.

**Table 1**
**LLM Classification Performance on Manually Annotated Articles**

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.

| 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 |

**Table 2**
**Descriptive Statistics per Outlet**

This table reports outlet-level descriptive statistics for the news corpus between January 2010 and June 2025. *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.

| Outlet | Country | Articles | | | Indicators | |
|---|---|---|---|---|---|---|
| | | Total | Green | % Green | Av. Transition Risk Score | Av. Physical Risk Score |
| Los Angeles Times | USA | 13548 | 8539 | 63.03 | 0.71 | 2.48 |
| New York Post | USA | 5850 | 3630 | 62.05 | 1.45 | 2.43 |
| Politico | USA | 11724 | 6325 | 53.95 | 0.93 | 2.59 |
| Washington Post | USA | 25148 | 14561 | 57.90 | 0.31 | 2.55 |
| Daily Mail | UK | 53462 | 28905 | 54.07 | 1.39 | 2.41 |
| The Guardian | UK | 51182 | 36917 | 72.13 | 0.98 | 2.59 |
| Total | | 160914 | 98877 | 61.45 | | |

## Table 3
## Correlations Between Climate Risk Indicators

This table reports pairwise correlations between my news-based climate risk indicators and existing alternatives. TRI and PRI denote my directional transition and physical climate risk indices; AT and AP denote their attention-based counterparts, constructed from the same classified articles without the directional component. MCCC is the Media Climate Change Concern index of Ardia et al. (2023). UCP (*U.S. Climate Policy*), IS (*International Summits*), GW (*Global Warming*), and ND (*Natural Disasters*) are the topic-specific climate news indices of Faccini et al. (2023). Correlations below the main diagonal are computed on the indices in levels; correlations above the main diagonal are computed on index innovations, measured as daily first differences. All series are observed at the daily frequency from January 2, 2014 to June 30, 2025.

| | MCCC | UCP | IS | GW | ND | TRI | PRI | AT | AP |
|---|---|---|---|---|---|---|---|---|---|
| MCCC | | 0.132 | 0.099 | 0.147 | 0.107 | -0.002 | 0.055 | 0.101 | 0.063 |
| UCP | 0.371 | | 0.241 | 0.352 | 0.195 | -0.014 | 0.098 | 0.195 | 0.118 |
| IS | 0.333 | 0.317 | | 0.410 | 0.154 | 0.044 | 0.073 | 0.159 | 0.083 |
| GW | 0.370 | 0.500 | 0.528 | | 0.179 | 0.071 | 0.119 | 0.192 | 0.146 |
| ND | 0.366 | 0.324 | 0.145 | 0.234 | | 0.009 | 0.120 | 0.096 | 0.125 |
| TRI | 0.124 | 0.057 | 0.116 | 0.207 | 0.096 | | 0.037 | 0.171 | 0.047 |
| PRI | 0.247 | 0.166 | 0.111 | 0.172 | 0.252 | 0.129 | | 0.119 | 0.801 |
| AT | 0.238 | 0.266 | 0.265 | 0.270 | 0.026 | 0.025 | 0.268 | | 0.135 |
| AP | 0.202 | 0.131 | 0.103 | 0.149 | 0.196 | 0.069 | 0.850 | 0.333 | |

**Table 4**
**Correlation Matrix of Daily TRI/PRI Innovations and Financial Variables**

This table reports pairwise correlations among the daily time-series variables from January 3, 2014 to June 30, 2025. Variables marked with $\Delta$ denote innovations, obtained as first difference. MKT, SMB, and HML are the Fama and French (1993) three factors, 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 returns (Fred St Louis, DCOILWTICO). Coefficients with an absolute value greater than 0.1 are shown in bold.

| Variable | $\Delta$TRI | $\Delta$PRI | MKT | SMB | HML | RMW | CMA | MOM | $\Delta$EPU | WTI | VIX |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\Delta$TRI | **1.00** | 0.05 | -0.00 | -0.02 | 0.00 | -0.01 | 0.01 | 0.02 | 0.01 | 0.00 | -0.01 |
| $\Delta$PRI | | **1.00** | -0.01 | -0.00 | 0.02 | -0.01 | 0.02 | 0.01 | -0.01 | -0.00 | 0.01 |
| MKT | | | **1.00** | **0.22** | -0.09 | **-0.21** | **-0.26** | **-0.12** | 0.01 | **0.12** | **-0.16** |
| SMB | | | | **1.00** | -0.02 | **-0.41** | **-0.13** | **-0.20** | 0.01 | 0.00 | -0.01 |
| HML | | | | | **1.00** | **0.32** | **0.56** | **-0.31** | -0.06 | 0.08 | -0.02 |
| RMW | | | | | | **1.00** | **0.27** | -0.04 | -0.04 | 0.03 | 0.03 |
| CMA | | | | | | | **1.00** | 0.03 | -0.04 | 0.04 | 0.05 |
| MOM | | | | | | | | **1.00** | 0.04 | -0.08 | -0.00 |
| $\Delta$EPU | | | | | | | | | **1.00** | 0.01 | -0.02 |
| WTI | | | | | | | | | | **1.00** | -0.09 |
| VIX | | | | | | | | | | | **1.00** |

## Table 5
## Properties of Portfolios Sorted on TRI Innovation Betas

Each month, stocks are sorted into five value-weighted portfolios (quintiles) based on their estimated exposure to innovations in the TRI, from lowest (1) to highest (5). Betas are estimated from daily excess returns over 12-month rolling windows. Panel A reports time-series averages (January 2015 – July 2025) of the value-weighted monthly return (in percent), the equal-weighted TRI innovation beta, and equal-weighted firm characteristics from Sharadar SF1 fundamentals: size (log market capitalization), book-to-market (book equity/market capitalization), leverage (debt/assets), profitability (net income/assets), tangibility (net PP&E/assets), sales growth (year-over-year quarterly revenue growth), investment (capex/assets), cash flow (operating cash flow/assets), and age (years since first pricing date). Accounting variables enter the sample the month after their filing date and are winsorized at the 1st and 99th percentiles. N is the average number of firms per portfolio. Panel B reports the intercept ($\alpha$, in decimal monthly return units) and the factor loadings from time-series regressions of portfolio excess returns on the Fama and French (2015) five factors augmented with momentum (Carhart, 1997), for each quintile and the long–short portfolio (5–1). $t$-statistics (in parentheses) use Newey and West (1987) standard errors with 4 lags. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| | | | **Panel A: Portfolio Characteristics** | | | |
| Return (VW, %) | 0.584 | 0.953 | 1.010 | 0.858 | 1.244 | |
| $\Delta$TRI $\hat{\beta}$ | -0.002 | -0.001 | 0.000 | 0.001 | 0.002 | |
| Size (log) | 22.098 | 22.482 | 22.565 | 22.508 | 22.167 | |
| Book-to-Market | 0.439 | 0.458 | 0.455 | 0.454 | 0.435 | |
| Leverage | 0.301 | 0.293 | 0.298 | 0.311 | 0.322 | |
| Profitability (ROA) | 0.000 | 0.009 | 0.010 | 0.009 | 0.001 | |
| Tangibility | 0.239 | 0.248 | 0.253 | 0.267 | 0.279 | |
| Sales Growth | 0.167 | 0.117 | 0.112 | 0.124 | 0.180 | |
| Investment | 0.009 | 0.008 | 0.008 | 0.009 | 0.010 | |
| Cash Flow | 0.015 | 0.020 | 0.020 | 0.020 | 0.016 | |
| Age (years) | 16.995 | 20.454 | 20.628 | 20.217 | 17.332 | |
| N | 396.135 | 395.929 | 395.754 | 395.548 | 395.373 | |
| | | | **Panel B: Factor Loadings** | | | |
| $\alpha$ | -0.005 | 0.001 | -0.001 | -0.000 | 0.001 | 0.006 |
| | (-1.42) | (0.60) | (-0.67) | (-0.02) | (0.58) | (1.37) |
| MKT | 1.124*** | 1.016*** | 0.960*** | 0.972*** | 1.102*** | -0.022 |
| | (12.01) | (32.80) | (49.03) | (23.00) | (16.89) | (-0.21) |
| SMB | 0.016 | -0.026 | -0.075** | -0.056 | 0.144** | 0.129 |
| | (0.09) | (-0.50) | (-2.28) | (-1.42) | (2.28) | (0.74) |
| HML | -0.193 | -0.045 | 0.006 | 0.083 | -0.010 | 0.183 |
| | (-1.37) | (-0.78) | (0.14) | (1.50) | (-0.16) | (1.09) |
| RMW | -0.145 | 0.057 | 0.105** | -0.014 | -0.105 | 0.039 |
| | (-1.38) | (0.90) | (2.12) | (-0.24) | (-1.19) | (0.24) |
| CMA | 0.131 | 0.047 | -0.123* | -0.045 | 0.109 | -0.022 |
| | (0.64) | (0.49) | (-1.72) | (-0.44) | (0.83) | (-0.08) |
| MOM | -0.036 | 0.024 | 0.005 | -0.036 | 0.103 | 0.139 |
| | (-0.49) | (0.69) | (0.22) | (-1.00) | (1.42) | (1.36) |

## Table 6
## Alphas of Value-Weighted Portfolios Sorted on Transition-Risk Innovation Betas

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 and French (1993) three-factor model, the Carhart four-factor model (Carhart, 1997), and the Fama and French (2015) five-factor model. Alphas are reported in monthly percentage terms. $t$-statistics (in parentheses) are based on Newey and West (1987) standard errors with a lag length of 4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| | **January 2015 – July 2025 (*127 months*)** | | | | | |
| Fama-French 3-factor alpha | -0.15 | -0.25 | 0.03 | -0.06 | 0.27** | 0.42* |
| | (-0.98) | (-1.35) | (0.33) | (-0.53) | (2.04) | (1.82) |
| Fama-French-Carhart alpha | -0.50 | -0.01 | 0.04 | -0.13 | 0.23 | 0.73* |
| | (-1.50) | (-0.08) | (0.39) | (-1.09) | (1.54) | (1.82) |
| Fama-French 5-factor alpha | -0.16 | -0.16 | -0.04 | 0.07 | 0.19 | 0.34 |
| | (-1.22) | (-0.87) | (-0.60) | (0.65) | (1.17) | (1.59) |
| | **November 2016 – December 2018 *(26 months)*** | | | | | |
| Fama-French 3-factor alpha | 0.18 | -0.10 | 0.07 | -0.06 | -0.23* | -0.41 |
| | (0.73) | (-0.68) | (0.81) | (-0.37) | (-1.66) | (-1.31) |
| Fama-French-Carhart alpha | 0.06 | 0.05 | -0.15 | 0.22** | -0.24* | -0.30 |
| | (0.36) | (0.38) | (-0.98) | (1.98) | (-1.81) | (-1.37) |
| Fama-French 5-factor alpha | -0.08 | 0.10 | -0.11 | 0.38*** | -0.44*** | -0.36 |
| | (-0.48) | (0.70) | (-0.82) | (3.03) | (-2.67) | (-1.58) |

## Table 7
## Alphas of Value-Weighted Portfolios Sorted on Physical-Risk Innovation Betas

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 and French (1993) three-factor model, the Carhart (1997) four-factor model, and the Fama and French (2015) five-factor model. Alphas are reported in monthly percentage terms. $t$-statistics (in parentheses) are based on Newey and West (1987) standard errors with a lag length of 4. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| | 1 | 2 | 3 | 4 | 5 | 5-1 |
|---|---|---|---|---|---|---|
| | **January 2015 – July 2025 (*127 months*)** | | | | | |
| Fama-French 3-factor alpha | -0.08 | -0.08 | -0.17* | 0.14* | -0.19 | -0.12 |
| | (-0.54) | (-0.80) | (-1.66) | (1.91) | (-0.59) | (-0.34) |
| Fama-French-Carhart alpha | -0.10 | -0.06 | -0.14 | 0.17* | -0.27 | -0.16 |
| | (-0.73) | (-0.57) | (-1.45) | (1.90) | (-0.82) | (-0.47) |
| Fama-French 5-factor alpha | -0.14 | -0.01 | 0.00 | -0.03 | -0.23 | -0.09 |
| | (-1.01) | (-0.11) | (0.05) | (-0.27) | (-0.74) | (-0.28) |

**Table 8**
**Fama–MacBeth Risk-Premium Estimates for Transition and Physical Risk Indicator Innovations**

This table reports Fama and MacBeth (1973) risk-premium estimates for the second-stage period January 2016 to July 2025. First-stage betas are estimated from daily data over 24-month rolling windows; a portfolio's beta computed through month $t$ is paired with its month $t+1$ excess return. The *74 (55) portfolios* universe combines the 25 Size–BM portfolios with the 49 (30) industry portfolios. Controls are the Fama and French (1993) factors, the momentum factor of Carhart (1997), and the Economic Policy Uncertainty (EPU) index of Baker et al. (2016). A $\Delta$ denotes daily first differences: $\Delta TRI$, $\Delta PRI$, and $\Delta EPU$. Specification *(i)* is the benchmark with factor controls only; *(ii)* and *(iii)* add the transition- and physical-risk innovation beta respectively; *(iv)* includes both. Following Lewellen et al. (2010), OLS and GLS $R^2$ are from a single cross-sectional regression of average excess returns on time-averaged betas, the GLS version weighting by the inverse sample covariance of returns. $t$-statistics use Newey and West (1987) standard errors with 4 lags. *, **, *** denote 10%, 5%, 1% significance.

| | 74 Portfolios | | | | 55 Portfolios | | | |
|---|---|---|---|---|---|---|---|---|
| | *(i)* | *(ii)* | *(iii)* | *(iv)* | *(i)* | *(ii)* | *(iii)* | *(iv)* |
| Cons. | -0.002 | -0.002 | 0.001 | 0.001 | -0.003 | -0.002 | 0.005 | 0.005 |
| | (-0.43) | (-0.42) | (0.20) | (0.18) | (-0.75) | (-0.33) | (0.98) | (0.91) |
| ΔTRI | | -4.417** | | -4.348** | | -6.941*** | | -6.543*** |
| | | (-2.21) | | (-2.27) | | (-2.64) | | (-2.76) |
| ΔPRI | | | -0.599 | -0.997 | | | 0.481 | 0.069 |
| | | | (-0.37) | (-0.63) | | | (0.23) | (0.04) |
| MKT | 0.011* | 0.012* | 0.008 | 0.008 | 0.013** | 0.011* | 0.004 | 0.005 |
| | (1.73) | (1.68) | (1.27) | (1.27) | (2.23) | (1.81) | (0.65) | (0.79) |
| SMB | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.002 | -0.002 |
| | (-0.23) | (-0.33) | (-0.46) | (-0.51) | (-0.32) | (-0.38) | (-0.56) | (-0.57) |
| HML | -0.000 | -0.001 | -0.000 | -0.000 | -0.001 | -0.001 | -0.001 | -0.001 |
| | (-0.06) | (-0.12) | (-0.00) | (-0.04) | (-0.20) | (-0.28) | (-0.12) | (-0.17) |
| MOM | -0.005 | -0.005 | -0.008 | -0.007 | -0.003 | 0.000 | -0.008 | -0.003 |
| | (-0.74) | (-0.77) | (-1.24) | (-1.09) | (-0.34) | (0.01) | (-0.99) | (-0.42) |
| ΔEPU | 90.300 | 48.500 | 154.806 | 100.583 | -4.830 | -145.843 | 93.743 | -66.080 |
| | (0.67) | (0.37) | (1.14) | (0.80) | (-0.04) | (-1.02) | (0.65) | (-0.49) |
| OLS $R^2$ | 0.34 | 0.37 | 0.46 | 0.48 | 0.40 | 0.42 | 0.48 | 0.51 |
| GLS $R^2$ | 0.21 | 0.27 | 0.22 | 0.27 | 0.32 | 0.32 | 0.32 | 0.34 |
| Obs. | 8510 | 8510 | 8510 | 8510 | 6325 | 6325 | 6325 | 6325 |

## Table 9
## Horse-Race Fama–MacBeth Regressions

This table reports second-stage Fama and MacBeth (1973) risk-premium estimates from regressions that include the transition- and physical-risk indicator innovations alongside competing news-based measures. MCCC is the Media Climate Change Concern index of Ardia et al. (2023); UCP is the U.S. Climate Policy index of Faccini et al. (2023). A $\Delta$ denotes daily first differences. $\Delta$TRI, $\Delta$PRI, $\Delta$MCCC, and $\Delta$UCP are standardized to unit variance, so that risk premia are expressed per one-standard-deviation shock. First-stage betas are estimated from daily excess returns over 24-month rolling windows, controlling for the Fama and French (1993) factors, momentum (Carhart, 1997), and $\Delta$EPU (Baker et al., 2016); the same betas enter the second stage. First-stage estimation is restricted to days on which all competing series are available, so that all columns are estimated on a common sample. Specification *(i)* re-estimates specification *(iv)* of Table 8 on this common sample with standardized innovations; it therefore differs from Table 8 in scale and daily coverage. Specifications *(ii)* and *(iii)* replace the transition- and physical-risk innovations with $\Delta$MCCC and $\Delta$UCP respectively; *(iv)* and *(v)* add each competing measure to specification *(i)*; *(vi)* includes all four. OLS and GLS cross-sectional $R^2$ are computed as in Table 8. $t$-statistics (in parentheses) are based on Newey and West (1987) standard errors with 4 lags. The second-stage sample is January 2016 to July 2025. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | *(i)* | *(ii)* | *(iii)* | *(iv)* | *(v)* | *(vi)* |
|---|---|---|---|---|---|---|
| **Panel A: 25 Size-B/M + 30 Industry (55 portfolios)** | | | | | | |
| $\Delta$TRI | -5.640*** | | | -5.436*** | -5.792*** | -5.538*** |
| | (-2.61) | | | (-2.81) | (-2.75) | (-3.05) |
| $\Delta$PRI | -0.036 | | | 0.078 | -0.477 | -0.569 |
| | (-0.01) | | | (0.03) | (-0.20) | (-0.24) |
| $\Delta$MCCC | | 1.522 | | 0.359 | | 0.903 |
| | | (0.51) | | (0.14) | | (0.32) |
| $\Delta$UCP | | | -1.315 | | -0.699 | -0.166 |
| | | | (-0.65) | | (-0.35) | (-0.07) |
| OLS $R^2$ | 0.48 | 0.58 | 0.41 | 0.65 | 0.49 | 0.64 |
| GLS $R^2$ | 0.35 | 0.36 | 0.33 | 0.38 | 0.36 | 0.39 |
| Months | 115 | 115 | 115 | 115 | 115 | 115 |
| **Panel B: 25 Size-B/M + 49 Industry (74 portfolios)** | | | | | | |
| $\Delta$TRI | -3.577** | | | -3.199** | -3.108* | -2.709* |
| | (-2.00) | | | (-2.05) | (-1.70) | (-1.75) |
| $\Delta$PRI | -1.361 | | | -1.533 | -1.931 | -2.159 |
| | (-0.68) | | | (-0.80) | (-1.02) | (-1.21) |
| $\Delta$MCCC | | -0.477 | | -1.206 | | -0.686 |
| | | (-0.17) | | (-0.49) | | (-0.26) |
| $\Delta$UCP | | | -1.429 | | -1.482 | -1.176 |
| | | | (-0.71) | | (-0.70) | (-0.50) |
| OLS $R^2$ | 0.46 | 0.42 | 0.35 | 0.53 | 0.45 | 0.54 |
| GLS $R^2$ | 0.27 | 0.23 | 0.24 | 0.29 | 0.27 | 0.28 |
| Months | 115 | 115 | 115 | 115 | 115 | 115 |

**Table 10**
**Single Predictive Regressions of Aggregate Market Returns on Transition and Physical Risk Indicators**

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 months on the current level of the Transition Risk Indicator (TRI) or the Physical Risk Indicator (PRI). The sample runs from January 2014 to June 2025. For each horizon $q$, $q$ 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 percent. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| $q =$ | 1 | 3 | 12 | 24 | 36 | 48 |
|---|---|---|---|---|---|---|
| TRI | -0.442 | -0.581 | -2.466 | 1.875 | -2.908 | -5.932*** |
| t-stat$_{\text{NW}}$ | (-1.04) | (-0.63) | (-0.56) | (0.24) | (-0.89) | (-4.83) |
| t-stat$_{\text{HH}}$ | (-1.01) | (-0.68) | (-0.82) | (0.47) | (-0.84) | (-2.48) |
| Adj. $R^2$ (%) | -0.32 | -0.42 | 0.76 | -0.14 | 3.11 | 14.79 |
| PRI | -2.381** | -3.587 | 3.733 | 12.056** | 0.218 | 6.630 |
| t-stat$_{\text{NW}}$ | (-1.99) | (-1.53) | (0.55) | (1.99) | (0.06) | (1.24) |
| t-stat$_{\text{HH}}$ | (-2.01) | (-1.49) | (0.63) | (3.17) | (0.05) | (1.34) |
| Adj. $R^2$ (%) | 2.92 | 2.98 | 0.28 | 7.96 | -0.99 | 4.20 |

## Table 11
## Multiple Predictive Regressions of Aggregate Market Returns

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, and coefficients are expressed in decimal return units: a coefficient of $-0.041$ implies a 4.1 percentage point lower compounded market return per one-standard-deviation increase in the indicator. Sample: January 2014 – November 2024 (the sample ends earlier than in Table 10 because some control variables are not available beyond that date). $q$ 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.

| | TRI | PRI | TERM | DEF | DY | RF | WTI | VIX | Adj. $R^2$ (%) |
|---|---|---|---|---|---|---|---|---|---|
| **Panel A:** $q = 1$ | | | | | | | | | |
| (i) | -0.006 | -0.014*** | -0.018** | 0.005 | -0.001 | -0.008 | | | 5.77 |
| | (-1.50) | (-2.93) | (-2.04) | (0.77) | (-0.09) | (-0.99) | | | |
| | [-1.44] | [-2.88] | [-2.02] | [0.74] | [-0.09] | [-0.97] | | | |
| (ii) | -0.005 | -0.011** | -0.011 | -0.008 | 0.011 | 0.004 | | 0.016*** | 10.06 |
| | (-1.22) | (-2.22) | (-1.26) | (-1.01) | (1.31) | (0.45) | | (2.69) | |
| | [-1.16] | [-2.19] | [-1.27] | [-0.97] | [1.27] | [0.45] | | [2.74] | |
| (iii) | -0.006 | -0.014*** | -0.020** | 0.005 | -0.001 | -0.010 | -0.005 | | 6.12 |
| | (-1.54) | (-2.88) | (-2.13) | (0.77) | (-0.16) | (-1.20) | (-1.12) | | |
| | [-1.46] | [-2.81] | [-2.12] | [0.74] | [-0.16] | [-1.21] | [-1.13] | | |
| (iv) | -0.005 | -0.011** | -0.012 | -0.007 | 0.010 | 0.002 | -0.002 | 0.015*** | 9.49 |
| | (-1.26) | (-2.23) | (-1.30) | (-0.91) | (1.24) | (0.25) | (-0.53) | (2.61) | |
| | [-1.19] | [-2.19] | [-1.30] | [-0.88] | [1.19] | [0.25] | [-0.52] | [2.60] | |
| **Panel B:** $q = 12$ | | | | | | | | | |
| (i) | -0.030** | -0.009 | -0.092* | 0.034*** | 0.052** | -0.017 | | | 48.16 |
| | (-2.56) | (-0.63) | (-1.96) | (2.66) | (2.48) | (-0.38) | | | |
| | [-2.38] | [-0.59] | [-1.81] | [2.66] | [2.26] | [-0.35] | | | |
| (ii) | -0.026** | 0.003 | -0.064* | -0.016 | 0.095*** | 0.031 | | 0.062*** | 56.70 |
| | (-2.54) | (0.21) | (-1.75) | (-0.94) | (4.22) | (0.81) | | (3.70) | |
| | [-2.32] | [0.19] | [-1.59] | [-0.92] | [3.96] | [0.73] | | [3.69] | |
| (iii) | -0.030** | -0.011 | -0.098** | 0.034** | 0.051** | -0.024 | -0.013* | | 48.77 |
| | (-2.50) | (-0.67) | (-2.01) | (2.51) | (2.36) | (-0.50) | (-1.76) | | |
| | [-2.34] | [-0.62] | [-1.84] | [2.51] | [2.19] | [-0.45] | [-1.98] | | |
| (iv) | -0.026** | 0.003 | -0.065* | -0.015 | 0.094*** | 0.029 | -0.001 | 0.061*** | 56.34 |
| | (-2.53) | (0.18) | (-1.72) | (-0.82) | (3.97) | (0.75) | (-0.18) | (3.28) | |
| | [-2.31] | [0.17] | [-1.56] | [-0.79] | [3.72] | [0.67] | [-0.18] | [3.22] | |
| **Panel C:** $q = 48$ | | | | | | | | | |
| (i) | -0.044*** | 0.030* | -0.086*** | -0.023* | 0.072*** | -0.145*** | | | 39.01 |
| | (-8.83) | (1.79) | (-3.49) | (-1.72) | (4.03) | (-5.41) | | | |
| | [-3.73] | [2.04] | [-3.13] | [-1.08] | [2.30] | [-3.97] | | | |
| (ii) | -0.041*** | 0.042*** | -0.062*** | -0.056*** | 0.105*** | -0.117*** | | 0.044*** | 45.85 |
| | (-9.67) | (3.28) | (-2.72) | (-3.07) | (3.95) | (-4.30) | | (2.76) | |
| | [-3.62] | [3.34] | [-1.85] | [-2.35] | [3.18] | [-2.82] | | [1.93] | |
| (iii) | -0.043*** | 0.029 | -0.095*** | -0.024* | 0.074*** | -0.158*** | -0.013** | | 40.26 |
| | (-8.54) | (1.60) | (-3.36) | (-1.85) | (4.22) | (-4.94) | (-2.03) | | |
| | [-3.85] | [1.86] | [-3.31] | [-1.21] | [2.49] | [-4.27] | [-1.32] | | |
| (iv) | -0.041*** | 0.041*** | -0.067*** | -0.055*** | 0.104*** | -0.124*** | -0.004 | 0.041*** | 45.39 |
| | (-10.02) | (2.88) | (-2.58) | (-3.37) | (4.20) | (-3.97) | (-1.01) | (2.64) | |
| | [-3.65] | [2.99] | [-2.03] | [-2.27] | [3.15] | [-2.99] | [-0.64] | [1.82] | |

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

This table reports classification performance on the manually annotated sample, split by publication date relative to the model's training cut-off (September 1, 2021): Group 1 (Pre-cutoff) and Group 2 (Post-cutoff). For each task (Q1–Q5), the table reports micro- and macro-averaged F1-score, precision (P), and recall (R), in percent; Support is the number of labeled articles, and ordinal tasks additionally report MAE and normalized MAE (lower is better). GPT-3.5-Turbo is used here as its earlier cut-off yields a larger post-cutoff sample, strengthening the contamination test.

| Grp. | Prompt | Q. | Support | Micro (%) | | | Macro (%) | | | MAE | nMAE |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | F1 | P | R | F1 | P | R | | |
| G.1 | P.2 | 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 | | |
| | | 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 |
| G.2 | P.2 | 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 | | |
| | | 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 |

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

This table reports micro- and macro-averaged F1-scores on the manually annotated sample, split by publication date relative to the model's training cut-off (September 1, 2021): Group 1 (Pre-cutoff) and Group 2 (Post-cutoff). For each task (Q1–Q5), the table reports the mean F1-score and a 95% confidence interval from 1,000 article-level bootstrap replications, in percent. GPT-3.5-Turbo is used here as its earlier cut-off yields a larger post-cutoff sample, strengthening the contamination test.

| Grp. | Prompt | Q. | 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] |

# Internet Appendix for "Are Transition and Physical Climate Risks Priced? Evidence from Directional News-Based Measures"

Clément Toussaint

Université Paris 1 Panthéon-Sorbonne,

CNRS, Centre d'Économie de la Sorbonne

This Internet Appendix provides supplementary material for "Are Transition and Physical Climate Risks Priced? Evidence from Directional News-Based Measures". Section IA1 details the outlet selection procedure: the ex-ante eligibility criteria and the resulting corpus composition. Section IA2 provides additional robustness checks for the LLM-based classification, including a comparison across alternative prompt designs and language models. Section IA3 reports additional asset pricing robustness checks based on alternative factor controls. Section IA4 presents return predictability and Fama–MacBeth results based on attention-only versions of the transition and physical risk indicators, which abstract from the directional dimension. Section IA5 reports robustness results obtained when constructing the indicators using only the first three LLM questions (Q1–Q3), thereby excluding the intensity and horizon dimensions. Section IA6 reports the return predictability and Fama–MacBeth results when the indicators are constructed from U.S. outlets only, excluding the *Guardian* and the *Daily Mail*.

# IA1 News Corpus Construction

This section describes how the outlet corpus used in the main paper is built. Following the standard practice for news-based indices (e.g., Baker et al., 2016; Ardia et al., 2023), the corpus is restricted to daily generalist newspapers with national audiences. Newswires are excluded because their content is syndicated into the newspapers already in the corpus, so including them would double-count the same news items. Broadcasters, magazines, aggregators, and platforms are excluded to preserve the homogeneity of the medium: the daily index normalizes and averages article flows across outlets, which requires comparable article formats, editorial structures, and archival coverage across sources. Table IA1.1 applies these criteria to the 50 most-visited news websites in the United States.

The outlets meeting the eligibility criteria (i.e., (i) the outlet is a newspaper, (ii) has a national audience, and (iii) is not an advocacy publication) are: the *New York Times*, *USA Today*, the *New York Post*, the *Wall Street Journal*, the *Guardian*, the *Washington Post*, the *Daily Mail*, *The Sun*, *The Hill*, *Politico*, the *Los Angeles Times*, and *The Independent*.

However, six of the eligible outlets are not part of the final corpus used in the main paper for the following reasons:

- *The New York Times* employs strict anti-automation measures that prevent reliable large-scale web scraping of full-text articles.
- *USA Today* is indexed on MediaCloud, and 21,771 article URLs are retrieved. However, these are RSS-feed URLs that no longer resolve to the underlying articles, so the full text cannot be recovered—including through the Internet Archive's Wayback Machine.
- The *Wall Street Journal* is no longer indexed by MediaCloud and is absent from GDELT, so no URL inventory is available.
- *The Sun* enters the ranking through `the-sun.com`, the U.S. edition of the outlet launched in 2020 (the British edition is `thesun.co.uk`); it therefore provides no content over most of the sample period.
- For *The Hill*, only 189 URLs match the climate keyword list (Section 3.1 of the main paper) over the sample period—insufficient for a daily index.

- *The Independent* ranks $49^{\text{th}}$ in the May 2026 ranking but does not appear consistently in the top 50 across adjacent months; the selection rule requires stable presence in the ranking.

## Table IA1.1
## Ex-Ante Corpus Selection over the 50 Most-Visited US News Outlets

This table reports the ex-ante selection over the 50 most-visited US news websites. Monthly visits are for May 2026, from Similarweb as compiled by Press Gazette. The partisan composition of each outlet's audience (% Rep. and % Dem.) is drawn from the Pew Research Center News Media Tracker (2025 wave). These figures report the share of Republican/Republican-leaning and Democratic/Democratic-leaning US adults who say they *regularly get news* from the outlet (Pew "use" measure); they do not sum to 100% and are not available for outlets absent from the Pew survey (reported as blank). An outlet is classified as *eligible* (in bold) if it is (i) national in audience (as opposed to local or regional), (ii) a newspaper (as opposed to an aggregator, newswire, online publication, magazine, or platform), and (iii) not an advocacy press. Eligible outlets that could not be collected for technical reasons (e.g., full-text unavailability, anti-scraping measures) are documented separately.

| Rank | Outlet | Country | Monthly Visits (M) | % Rep. | % Dem. | Media Type | Content Type | Scope | Eligible |
|---|---|---|---|---|---|---|---|---|---|
| **1** | **nytimes.com** | **USA** | **409.5** | **10** | **29** | **Newspaper** | **Generalist** | **National** | **Yes** |
| 2 | cnn.com | USA | 227.3 | 20 | 48 | Broadcaster | Generalist | National | No |
| 3 | foxnews.com | USA | 207.7 | 57 | 18 | Broadcaster | Generalist | National | No |
| 4 | finance.yahoo.com | USA | 140.5 | | | Aggregator | Business | National | No |
| 5 | people.com | USA | 136 | | | Magazine | Entertainment | National | No |
| 6 | msn.com | USA | 129.4 | | | Aggregator | Generalist | National | No |
| 7 | bbc.com | UK | 92.3 | 13 | 30 | Broadcaster | Generalist | National | No |
| 8 | news.google.com | USA | 90.3 | | | Aggregator | Generalist | National | No |
| 9 | substack.com | USA | 88.4 | | | Platform | Platform | National | No |
| **10** | **usatoday.com** | **USA** | **87.1** | **11** | **17** | **Newspaper** | **Generalist** | **National** | **Yes** |
| **11** | **nypost.com** | **USA** | **85** | **10** | **8** | **Newspaper** | **Generalist** | **National** | **Yes** |
| 12 | cnbc.com | USA | 73.7 | | | Broadcaster | Business | National | No |
| 13 | apnews.com | USA | 68.9 | 11 | 31 | Newswire | Generalist | National | No |
| **14** | **wsj.com** | **USA** | **68** | **12** | **16** | **Newspaper** | **Generalist** | **National** | **Yes** |
| 15 | nbcnews.com | USA | 66.8 | 24 | 47 | Broadcaster | Generalist | National | No |
| **16** | **theguardian.com** | **UK** | **65.7** | **4** | **13** | **Newspaper** | **Generalist** | **National** | **Yes** |
| 17 | news.yahoo.com | USA | 64.9 | | | Aggregator | Generalist | National | No |
| 18 | cbsnews.com | USA | 60.4 | 22 | 39 | Broadcaster | Generalist | National | No |

*Continued on next page*

*Continued from previous page*

| Rank | Outlet | Country | Monthly Visits (M) | % Rep. | % Dem. | Media Type | Content Type | Scope | Eligible |
|---|---|---|---|---|---|---|---|---|---|
| **19** | **washingtonpost.com** | **USA** | **55.3** | **7** | **18** | **Newspaper** | **Generalist** | **National** | **Yes** |
| 20 | npr.org | USA | 55 | 9 | 32 | Broadcaster | Generalist | National | No |
| 21 | forbes.com | USA | 48.5 | 9 | 10 | Magazine | Business | National | No |
| 22 | drudgereport.com | USA | 41.6 | | | Aggregator | Generalist | National | No |
| 23 | businessinsider.com | USA | 36.9 | | | Online publication | Business | National | No |
| **24** | **dailymail.com** | **UK** | **35.2** | | | **Newspaper** | **Generalist** | **National** | **Yes** |
| 25 | reuters.com | USA | 32.5 | | | Newswire | Generalist | National | No |
| 26 | huffpost.com | USA | 32.3 | 2 | 8 | Online publication | Generalist | National | No |
| **27** | **the-sun.com** | **UK** | **31.5** | | | **Newspaper** | **Generalist** | **National** | **Yes** |
| **28** | **thehill.com** | **USA** | **31.1** | **4** | **6** | **Newspaper** | **Politics** | **National** | **Yes** |
| 29 | indiatimes.com | India | 29.6 | | | Aggregator | Generalist | National | No |
| **30** | **politico.com** | **USA** | **29.5** | **4** | **12** | **Newspaper** | **Politics** | **National** | **Yes** |
| 31 | buzzfeed.com | USA | 28.9 | | | Online publication | Generalist | National | No |
| 32 | newsbreak.com | USA | 25.9 | | | Aggregator | Generalist | National | No |
| 33 | abcnews.com | USA | 24.5 | 27 | 46 | Broadcaster | Generalist | National | No |
| 34 | breitbart.com | USA | 23.7 | 7 | 0 | Online publication | Generalist | National | No |
| 35 | newsweek.com | USA | 23 | 6 | 10 | Magazine | Generalist | National | No |
| 36 | axios.com | USA | 23 | 1 | 5 | Online publication | Politics | National | No |
| 37 | aljazeera.com | Qatar | 22.1 | | | Broadcaster | Generalist | National | No |
| 38 | theatlantic.com | USA | 21.4 | 1 | 10 | Magazine | Generalist | National | No |
| 39 | thegatewaypundit.com | USA | 21.1 | | | Online publication | Generalist | National | No |
| 40 | patch.com | USA | 20.3 | | | Online publication | Generalist | Local | No |
| **41** | **latimes.com** | **USA** | **20** | | | **Newspaper** | **Generalist** | **National** | **Yes** |
| 42 | newsmax.com | USA | 19.7 | 15 | 1 | Broadcaster | Generalist | National | No |
| 43 | usnews.com | USA | 18.9 | | | Magazine | Generalist | National | No |
| 44 | sfgate.com | USA | 18.8 | | | Newspaper | Generalist | Local | No |

*Continued on next page*

*Continued from previous page*

| Rank | Outlet | Country | Monthly Visits (M) | % Rep. | % Dem. | Media Type | Content Type | Scope | Eligible |
|---|---|---|---|---|---|---|---|---|---|
| 45 | bloomberg.com | USA | 18.6 | | | Other | Business | National | No |
| 46 | variety.com | USA | 18 | | | Magazine | Entertainment | National | No |
| 47 | thedailybeast.com | USA | 15.9 | | | Online publication | Generalist | National | No |
| 48 | ksl.com | USA | 15.9 | | | Broadcaster | Generalist | Local | No |
| **49** | **independent.co.uk** | **UK** | **14.3** | | | **Newspaper** | **Generalist** | **National** | **Yes** |
| 50 | theepochtimes.com | USA | 14.2 | | | Newspaper | Advocacy Press | National | No |

# IA2 Robustness of the LLM Classification

This section reports additional robustness checks for the LLM-based classification. I evaluate three alternative prompt designs on a manually annotated subset of the newspaper corpus and compare two LLMs (*GPT-5-mini* and *GPT-5-nano*) to select the configuration with the best classification performance.

The three prompts are:

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

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

**P.3** Same as P.2 but with more detailed instructions, definitions, and examples.

Figure IA2.1 reports the F1-scores by question and prompt for GPT-5-mini, along with bootstrapped 95% confidence intervals. Table IA2.1 provides the full performance metrics for both models and all three prompts. Table IA2.2 reports the same results with bootstrapped confidence intervals.

Three patterns emerge. First, classification performance varies substantially across prompts: for Q1, GPT-5-mini's macro-F1 ranges from 66.5% (P.1) to 86.3% (P.2). This sensitivity to prompt design is inconsistent with a pure memorization explanation, which would predict similar performance regardless of how the input is framed.

Second, the structured prompt with concise instructions (P.2) outperforms both alternatives on the classification tasks that drive the index construction (Q1–Q3) and on intensity (Q4); performance on the horizon task (Q5) is low for all prompts. This suggests that overly specific instructions may constrain the model's ability to generalize across the heterogeneous language of news articles.

Third, under the selected prompt P.2, GPT-5-mini achieves a higher macro-F1 than GPT-5-nano on all five tasks. For example, under prompt P.2, GPT-5-mini achieves a macro-F1 of 86.3% on Q1 compared to 79.1% for GPT-5-nano. Based on these results, I use GPT-5-mini with prompt P.2 in the main analysis.

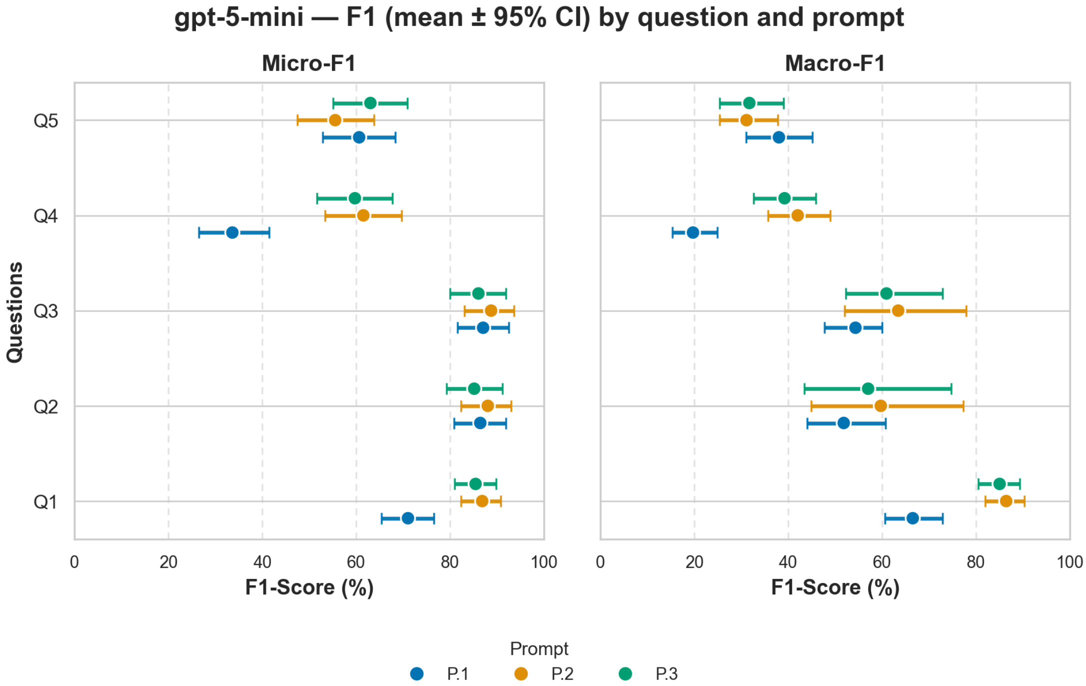



**Figure IA2.1. Sensitivity of LLM Classification Performance to Prompt Design.** This figure reports micro and macro averaged F1-scores for the five annotation tasks (Q1–Q5) obtained with *GPT-5-mini* under three alternative prompt designs (P.1–P.3). Points denote mean F1-scores, and horizontal bars indicate 95% confidence intervals computed from 1,000 bootstrap replications. Differences across prompts highlight the sensitivity of classification performance to instruction design.

## Table IA2.1
## LLM Classification Performance on the Manually Annotated Sample

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.

| 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 |

## Table IA2.2
## Bootstrapped F1-Scores for LLM Classification on the Manually Annotated Newspaper Sample

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.

| 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] |

**Prompt P.1**

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?

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

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

`{{ "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?

`{{ "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?

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

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

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

**Question 6:** Which countries or international organizations are concerned by the climate risk described in the article? Return country names in English or international organizations (e.g., "European Union", "United Nations"), if multiple are mentioned, list them all (and if it concerns the world globally just mention "World"), if none can be identified, return an empty list []. Strictly answer countries names, international organization and/or "World". Do not mention cities or regions.

`{{ "countries_or_orgs": [Entity1, Entity2, ...] }}`

**Question 7:** What is the main topic of the article? If several topics are relevant, choose at most 2 general topics in maximum two or three words without mentioning specific company or event.

`{{ "topic": [] }}`

**Prompt P.2**

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" }}
```

### Prompt P.3

You are an expert researcher in green finance. You will carefully analyze the following article and answer a sequence of questions, step by step. Base your answers strictly on the article's content and on the rules and definitions for each question, not on prior knowledge. For each question, provide the most accurate answer possible. Always output a complete JSON object with all required fields. If the answer to Question 1 is "No," fill every other field with "N/A" (or [] where appropriate).

Article: `{article_text}`

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

Definitions:

- Answer "Yes" if climate change, global warming, greenhouse gas emissions, energy transition policies, extreme weather events, adaptation or mitigation strategies, or their social/economic consequences are a central or substantial part of the article.
- Climate change can be one of several topics, but it must be discussed in detail or play a significant role in the narrative.
- Answer "No" if climate change is absent, mentioned only briefly, or if the article only discusses related topics (e.g., biodiversity, pollution, generic sustainability) without explicitly connecting them to climate change or impacts on humans.

Rules:

- Default to "No" unless the link to climate change is explicit and substantial.
- Do not infer climate relevance if it is not clearly present in the text.

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

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

Definitions:

- "Transition": risks related to the economic, financial, or regulatory changes of moving toward a low-carbon economy (e.g., new climate laws, carbon tax, green technologies, stranded assets, environmental regulations).
- "Physical": risks related to the direct physical effects of climate change (e.g., heatwaves, hurricanes, floods, wildfires, rising sea levels, and their economic or social consequences).
- "Both": the article discusses transition and physical risks with roughly equal weight and importance.
- "Unclear": climate change is mentioned, but the type of risk cannot be confidently identified.

Rules:

- Select the category that reflects the main focus of the article. If one risk type dominates clearly, choose that one.
- Select "Both" only if both transition and physical risks are discussed in comparable detail and importance.
- Select "Unclear" if the link to a specific risk type is vague, indirect, or missing.

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

**Question 3:** From the perspective 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 introduced, so firms must pay additional costs on their emissions.
- Stricter environmental regulations are passed, increasing compliance costs for industries.
- Courts uphold lawsuits against major emitters, raising legal risks.
- Rising global temperatures increase the likelihood of hurricanes or floods, raising physical risks for coastal infrastructure.

- International agreements tighten emission targets (e.g., COP outcomes), creating higher transition risks.

Examples for a "Decrease":

- Lawmakers vote down stricter emission standards, easing regulatory risks.
- The EPA's power is reduced, lowering enforcement risks.
- The EU delays the implementation of stricter climate rules, reducing short-term compliance costs.
- Subsidies or tax breaks are granted for fossil fuel activities, easing transition risks.
- Development of drought-resistant crops lowers physical risks for farmers facing hotter, drier summers.

Examples for "Neutral":

- The article discusses climate change but without clear implications for firms' risks.

```
{{ "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 (the primary criterion is the geographic scope of the risk, with binding level as a secondary criterion):

- "Weak": Local or highly limited scope (e.g., a single city, municipality, or county). Impacts are minor, exploratory, or symbolic. May include pilot projects, consultations, or draft proposals with no binding effect.
- "Medium": Regional or sector-specific scope (e.g., a U.S. state, or a particular industry/sector). Impacts are moderate. Measures may be partially binding or represent serious commitments, but not yet global or fully comprehensive.
- "Strong": National or international/global scope. Involves major actors or organizations (e.g., U.S., EU, China, G7, UN). Impacts are significant and the measures are binding or categorical (e.g., a ban, mandatory phase-out, or declaration of climate emergency).
- "N/A": Use only if the previous answer was "Neutral". When uncertain, choose the closest among Weak, Medium, or Strong.

Examples:

- Weak: A city announces a new recycling pilot program; a draft local ordinance is discussed.
- Medium: California implements stricter emission standards.
- Strong: The EU enforces a binding cap on emissions; the U.S. Congress passes a nationwide climate bill; the UN adopts a global treaty on emissions.

```
{{ "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 imminent or already materializing (0–3 years). Mentions of "now", "immediately", "in the coming years", or near-term policies.
- "Medium": The risk is foreseeable but not immediate (3–10 years). References to the next several years, such as "by 2030" or "within the next decade".
- "Long": The risk unfolds in a distant or structural timeframe (10+ years). Mentions of "by 2050", long-term climate trajectories, or multi-decade technological shifts.
- "N/A": Use only if the article provides no temporal indication.

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

**Question 6:** Which countries or international organizations are concerned by the climate risk described in the article? Return country names in English or international organizations (e.g., "European Union", "United Nations"). If multiple are mentioned, list them all (and if it concerns the world globally just mention "World"). If none can be identified, return an empty list []. Strictly answer country names, international organizations and/or "World". Do not mention cities or regions.

```
{{ "countries_or_orgs": [Entity1, Entity2, ...] }}
```

**Question 7:** What is the main topic of the article? If several topics are relevant, choose at most 2 general topics in maximum two or three words without mentioning specific companies

```
or events.


{{ "topic": [] }}
```

## IA3 Alternative Factor Controls

This section reports additional Fama–MacBeth risk premium estimates to assess the robustness of the cross-sectional pricing results documented in the main paper to alternative specifications. I consider three models: the three-factor model of Fama and French (1993) (FF3), the four-factor model of Carhart (1997) (FFC), and the five-factor model of Fama and French (2015) (FF5) (with and without the first difference of the Economic Policy Uncertainty ($\Delta$EPU) indicator of Baker et al. (2016)).

Table IA3.1 reports the results. The negative risk premium associated with transition risk innovations ($\Delta$TRI) remains negative and significant across all specifications and asset universes. The point estimates are also economically stable, ranging from $-3.593$ to $-6.551$ across specifications. The physical risk innovations ($\Delta$PRI) are not significantly priced, consistent with the main results in the paper.

These results confirm that the negative and significant cross-sectional price of transition risk documented in the main paper is robust to alternative factor controls.

## Table IA3.1
## Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls

This table reports second-pass Fama–MacBeth risk premium estimates over the period January 2016–July 2025. 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 (TRI) and physical (PRI) climate risk factors enter as innovations. Newey and West (1987) t-statistics with 4 lags are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | 74 Portfolios | | | | | | 55 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FF5 | | FFC | | FF3 | | FF5 | | FFC | | FF3 | |
| Term | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) |
| ΔTRI | -4.111** | -3.593* | -4.354** | -4.292** | -3.782** | -3.716* | -5.164** | -4.818** | -6.551*** | -6.213*** | -5.648** | -5.847*** |
| | (-2.12) | (-1.81) | (-2.28) | (-2.33) | (-1.99) | (-1.80) | (-2.38) | (-2.34) | (-2.76) | (-2.77) | (-2.57) | (-2.63) |
| ΔPRI | -1.585 | -1.426 | -1.003 | -1.225 | -1.177 | -1.115 | -0.363 | -0.370 | 0.063 | -0.197 | -0.020 | -0.167 |
| | (-0.96) | (-0.84) | (-0.63) | (-0.76) | (-0.76) | (-0.72) | (-0.17) | (-0.18) | (0.03) | (-0.10) | (-0.01) | (-0.09) |
| MKT | 0.011* | 0.009 | 0.008 | 0.007 | 0.008 | 0.006 | 0.011** | 0.011** | 0.005 | 0.005 | 0.004 | 0.005 |
| | (1.95) | (1.48) | (1.27) | (1.00) | (1.24) | (0.89) | (2.04) | (2.08) | (0.79) | (0.77) | (0.69) | (0.79) |
| SMB | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 | -0.001 | -0.002 | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 |
| | (-0.61) | (-0.59) | (-0.51) | (-0.48) | (-0.50) | (-0.42) | (-0.64) | (-0.63) | (-0.57) | (-0.53) | (-0.51) | (-0.49) |
| HML | 0.001 | 0.001 | -0.000 | -0.000 | 0.001 | 0.001 | -0.000 | -0.001 | -0.001 | -0.001 | -0.000 | -0.001 |
| | (0.22) | (0.28) | (-0.04) | (-0.03) | (0.16) | (0.23) | (-0.11) | (-0.12) | (-0.17) | (-0.22) | (-0.09) | (-0.12) |
| ΔEPU | 201.677 | | 100.429 | | 128.909 | | 108.326 | | -66.315 | | 7.991 | |
| | (1.47) | | (0.80) | | (0.96) | | (0.67) | | (-0.49) | | (0.05) | |
| MOM | | | -0.007 | -0.007 | | | | | -0.003 | -0.002 | | |
| | | | (-1.09) | (-1.03) | | | | | (-0.42) | (-0.33) | | |
| RMW | 0.000 | -0.001 | | | | | 0.000 | 0.000 | | | | |
| | (0.05) | (-0.25) | | | | | (0.13) | (0.00) | | | | |
| CMA | 0.001 | 0.001 | | | | | 0.001 | 0.000 | | | | |
| | (0.32) | (0.26) | | | | | (0.39) | (0.07) | | | | |
| Cons. | -0.002 | 0.001 | 0.001 | 0.003 | 0.002 | 0.004 | -0.001 | -0.001 | 0.005 | 0.005 | 0.005 | 0.005 |
| | (-0.32) | (0.11) | (0.18) | (0.45) | (0.33) | (0.64) | (-0.19) | (-0.19) | (0.92) | (1.01) | (1.01) | (0.96) |
| OLS $R^2$ | 0.43 | 0.39 | 0.48 | 0.38 | 0.43 | 0.33 | 0.51 | 0.49 | 0.51 | 0.48 | 0.44 | 0.43 |
| GLS $R^2$ | 0.20 | 0.20 | 0.27 | 0.27 | 0.16 | 0.16 | 0.21 | 0.20 | 0.34 | 0.33 | 0.19 | 0.19 |
| Observations | 8510 | 8510 | 8510 | 8510 | 8510 | 8510 | 6325 | 6325 | 6325 | 6325 | 6325 | 6325 |

# IA4 Additional Results for the Attention Indicators

This section reports return predictability and Fama–MacBeth results for the attention-based indicators (AT and AP), which are constructed by setting the directional weight to one in Equation (3) of the main paper while keeping intensity and horizon weights unchanged. These indicators capture the volume and intensity of climate-related news coverage without distinguishing whether the news suggests risk is increasing or decreasing.

Table IA4.1 reports predictive regressions of aggregate market returns on the AT and AP indicators. Table IA4.2 reports second-pass Fama–MacBeth risk premium estimates.

Once the directional dimension is removed, the joint pricing-and-predictability evidence no longer supports the ICAPM-based hedging interpretation documented in the main paper. In particular, the AT no longer predicts negative aggregate market returns with a sign consistent with the cross-sectional price of risk, and the Fama–MacBeth risk premium associated with ΔAT is not statistically significant in any of the specifications.

**Table IA4.1**
**Single Predictive Regressions of Aggregate Market Returns on the Transition and Physical Attention Indices**

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 months on the current level of the Transition Attention Indicator (AT) or the Physical Attention Indicator (AP). The sample runs from January 2014 to June 2025. For each horizon $q$, $q$ months of data are lost by construction. Newey and West (1987) $q$-lag t-statistics and Hansen and Hodrick (1980) $q$-lag t-statistics are reported below the coefficient estimates. The estimates are in percent. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| $q =$ | 1 | 3 | 12 | 24 | 36 | 48 |
|---|---|---|---|---|---|---|
| AT | -0.539 | -1.210 | -0.176 | 4.179 | 0.252 | 11.131** |
| t-stat$_{\text{NW}}$ | (-0.56) | (-0.63) | (-0.03) | (0.67) | (0.10) | (2.46) |
| t-stat$_{\text{HH}}$ | (-0.58) | (-0.62) | (-0.02) | (0.71) | (0.07) | (2.20) |
| Adj. $R^2$ (%) | -0.57 | -0.37 | -0.80 | 0.12 | -0.99 | 15.56 |
| AP | -2.659** | -3.907 | 3.051 | 12.199*** | 2.571 | 8.160* |
| t-stat$_{\text{NW}}$ | (-2.18) | (-1.48) | (0.62) | (3.41) | (0.95) | (1.74) |
| t-stat$_{\text{HH}}$ | (-2.27) | (-1.59) | (0.63) | (4.59) | (0.64) | (1.60) |
| Adj. $R^2$ (%) | 3.66 | 3.51 | -0.11 | 8.01 | -0.12 | 7.74 |

**Table IA4.2**

**Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls for Attention Indicators**

This table reports second-pass Fama–MacBeth risk premium estimates over the period January 2016–July 2025. 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 indicator factors enter as innovations. Newey and West (1987) t-statistics with 4 lags are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | 74 Portfolios | | | | | | 55 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FF5 | | FFC | | FF3 | | FF5 | | FFC | | FF3 | |
| Term | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) |
| ΔAT | 0.395 | 0.049 | 0.778 | 0.513 | 0.441 | -0.058 | 1.595 | 0.925 | 1.875 | 0.951 | 1.385 | 0.185 |
| | (0.56) | (0.07) | (1.06) | (0.62) | (0.61) | (-0.08) | (1.39) | (0.84) | (1.61) | (0.86) | (1.11) | (0.15) |
| ΔAP | 0.238 | 0.444 | 0.668 | 0.381 | 0.646 | 0.948 | 0.030 | -0.636 | 0.203 | -1.080 | 0.393 | -0.833 |
| | (0.18) | (0.31) | (0.54) | (0.33) | (0.50) | (0.73) | (0.02) | (-0.37) | (0.11) | (-0.63) | (0.21) | (-0.49) |
| MKT | 0.014** | 0.011* | 0.010 | 0.007 | 0.010* | 0.007 | 0.016*** | 0.015*** | 0.008 | 0.006 | 0.008 | 0.006 |
| | (2.36) | (1.89) | (1.54) | (0.96) | (1.68) | (1.05) | (3.13) | (3.00) | (1.38) | (0.88) | (1.61) | (1.10) |
| SMB | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 | -0.001 |
| | (-0.49) | (-0.46) | (-0.34) | (-0.33) | (-0.35) | (-0.30) | (-0.58) | (-0.57) | (-0.49) | (-0.40) | (-0.45) | (-0.38) |
| HML | 0.001 | 0.002 | -0.000 | 0.000 | 0.001 | 0.001 | 0.000 | 0.001 | -0.000 | -0.000 | 0.000 | 0.000 |
| | (0.30) | (0.40) | (-0.04) | (0.04) | (0.19) | (0.28) | (0.07) | (0.11) | (-0.09) | (-0.06) | (0.04) | (0.05) |
| ΔEPU | 260.736* | | 198.114 | | 235.951 | | 245.981 | | 203.849 | | 276.606* | |
| | (1.83) | | (1.43) | | (1.58) | | (1.58) | | (1.33) | | (1.70) | |
| MOM | | | -0.006 | -0.007 | | | | | -0.003 | -0.004 | | |
| | | | (-1.00) | (-1.05) | | | | | (-0.47) | (-0.54) | | |
| RMW | -0.001 | -0.001 | | | | | -0.000 | -0.000 | | | | |
| | (-0.21) | (-0.45) | | | | | (-0.09) | (-0.07) | | | | |
| CMA | 0.002 | 0.002 | | | | | 0.003 | 0.002 | | | | |
| | (0.54) | (0.47) | | | | | (0.78) | (0.51) | | | | |
| Cons. | -0.004 | -0.002 | -0.000 | 0.002 | -0.000 | 0.002 | -0.007 | -0.006 | 0.001 | 0.004 | 0.001 | 0.003 |
| | (-0.79) | (-0.36) | (-0.07) | (0.42) | (-0.08) | (0.44) | (-1.49) | (-1.32) | (0.30) | (0.73) | (0.17) | (0.62) |
| OLS $R^2$ | 0.42 | 0.37 | 0.50 | 0.40 | 0.44 | 0.35 | 0.47 | 0.47 | 0.54 | 0.50 | 0.48 | 0.45 |
| GLS $R^2$ | 0.18 | 0.18 | 0.23 | 0.23 | 0.18 | 0.18 | 0.24 | 0.23 | 0.34 | 0.33 | 0.22 | 0.22 |
| Observations | 8510 | 8510 | 8510 | 8510 | 8510 | 8510 | 6325 | 6325 | 6325 | 6325 | 6325 | 6325 |

# IA5 Robustness to Excluding Intensity and Horizon Labels (Q4-Q5 Ablations)

This section reports the main asset pricing results when the transition and physical risk indicators are constructed using only the first three LLM questions (Q1: climate relevance, Q2: risk type, Q3: risk direction). The intensity (Q4) and horizon (Q5) weights are set to one in Equation (3) of the main paper, which removes the dimensions most susceptible to look-ahead bias.

Table IA5.1 reports single-predictive regressions of aggregate market returns on the Q1-Q3 only indicators. Table IA5.2 reports second-pass Fama–MacBeth risk premium estimates with alternative factor controls. The negative price of transition risk remains statistically significant in eleven of the twelve specifications. The long-horizon predictive relation is also preserved, with a negative and significant coefficient at the 48-month horizon, and no horizon displays a significant coefficient of opposite sign. These results confirm that the main findings do not depend on the dimensions most vulnerable to look-ahead concerns.

**Table IA5.1**
**Single Predictive Regressions of Aggregate Market Returns – Q1-Q3 Only Indicators (Robustness)**

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 months 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 June 2025. For each horizon $q$, $q$ 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 percent. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| $q =$ | 1 | 3 | 12 | 24 | 36 | 48 |
|---|---|---|---|---|---|---|
| $\text{TRI}_{\text{Q1–Q3}}$ | -0.475 | -0.540 | -2.112 | 2.890 | -2.944 | -5.582*** |
| $\text{t-stat}_{\text{NW}}$ | (-0.98) | (-0.54) | (-0.45) | (0.35) | (-0.86) | (-3.48) |
| $\text{t-stat}_{\text{HH}}$ | (-0.96) | (-0.59) | (-0.67) | (0.70) | (-0.80) | (-2.01) |
| Adj. $R^2$ (%) | -0.32 | -0.51 | 0.17 | 0.63 | 2.57 | 10.88 |
| $\text{PRI}_{\text{Q1–Q3}}$ | -2.695* | -3.962 | 3.189 | 14.052* | 1.451 | 7.803 |
| $\text{t-stat}_{\text{NW}}$ | (-1.94) | (-1.54) | (0.40) | (1.96) | (0.35) | (1.17) |
| $\text{t-stat}_{\text{HH}}$ | (-1.97) | (-1.50) | (0.45) | (3.25) | (0.31) | (1.29) |
| Adj. $R^2$ (%) | 2.85 | 2.73 | -0.20 | 8.33 | -0.79 | 4.70 |

## Table IA5.2
## Fama–MacBeth Risk Premia with Alternative Factor Controls – Q1-Q3-Only Indicators

This table reports second-pass Fama–MacBeth risk premium estimates over the period January 2016–July 2025. 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. The transition and physical risk factors constructed using Q1–Q3 only—$\Delta\mathrm{TRI}_{\mathrm{Q1-Q3}}$ and $\Delta\mathrm{PRI}_{\mathrm{Q1-Q3}}$—enter as innovations. Newey and West (1987) t-statistics with lag length 4 are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | 74 Portfolios | | | | | | 55 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FF5 | | FFC | | FF3 | | FF5 | | FFC | | FF3 | |
| Term | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) |
| $\Delta\mathrm{TRI}_{\mathrm{Q1-Q3}}$ | -4.103** | -3.233* | -3.863** | -3.620** | -3.556* | -3.243 | -6.034*** | -5.327*** | -6.650*** | -6.232*** | -6.233*** | -6.205*** |
| | (-2.21) | (-1.67) | (-2.10) | (-2.04) | (-1.92) | (-1.60) | (-2.87) | (-2.62) | (-3.09) | (-2.97) | (-2.94) | (-2.83) |
| $\Delta\mathrm{PRI}_{\mathrm{Q1-Q3}}$ | -1.363 | -1.491 | -0.184 | -0.490 | -0.554 | -0.773 | -0.326 | -0.321 | 0.391 | 0.230 | 0.225 | 0.135 |
| | (-0.85) | (-0.92) | (-0.12) | (-0.31) | (-0.38) | (-0.51) | (-0.16) | (-0.16) | (0.21) | (0.12) | (0.12) | (0.08) |
| MKT | 0.011* | 0.009 | 0.009 | 0.007 | 0.008 | 0.006 | 0.011** | 0.012** | 0.005 | 0.005 | 0.004 | 0.005 |
| | (1.93) | (1.46) | (1.37) | (1.05) | (1.24) | (0.87) | (2.19) | (2.23) | (0.88) | (0.87) | (0.72) | (0.81) |
| SMB | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 | -0.001 | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 | -0.001 |
| | (-0.61) | (-0.59) | (-0.51) | (-0.47) | (-0.50) | (-0.42) | (-0.63) | (-0.62) | (-0.55) | (-0.52) | (-0.49) | (-0.48) |
| HML | 0.001 | 0.002 | 0.000 | 0.000 | 0.001 | 0.001 | -0.000 | -0.000 | -0.000 | -0.001 | -0.000 | -0.000 |
| | (0.30) | (0.36) | (0.03) | (0.02) | (0.21) | (0.27) | (-0.03) | (-0.04) | (-0.10) | (-0.16) | (-0.03) | (-0.07) |
| $\Delta$EPU | 199.031 | | 97.612 | | 114.734 | | 96.677 | | -60.632 | | -4.167 | |
| | (1.46) | | (0.78) | | (0.85) | | (0.59) | | (-0.44) | | (-0.03) | |
| MOM | | | -0.007 | -0.006 | | | | | -0.004 | -0.003 | | |
| | | | (-1.05) | (-0.96) | | | | | (-0.53) | (-0.37) | | |
| RMW | 0.000 | -0.001 | | | | | 0.000 | 0.000 | | | | |
| | (0.09) | (-0.20) | | | | | (0.14) | (0.02) | | | | |
| CMA | 0.001 | 0.001 | | | | | 0.001 | 0.000 | | | | |
| | (0.26) | (0.21) | | | | | (0.40) | (0.07) | | | | |
| Cons. | -0.001 | 0.001 | 0.001 | 0.002 | 0.002 | 0.004 | -0.001 | -0.001 | 0.004 | 0.005 | 0.005 | 0.005 |
| | (-0.29) | (0.13) | (0.10) | (0.41) | (0.36) | (0.69) | (-0.29) | (-0.31) | (0.85) | (0.91) | (1.06) | (0.98) |
| OLS $R^2$ | 0.41 | 0.36 | 0.42 | 0.32 | 0.41 | 0.29 | 0.50 | 0.48 | 0.48 | 0.45 | 0.44 | 0.42 |
| GLS $R^2$ | 0.19 | 0.19 | 0.27 | 0.27 | 0.16 | 0.16 | 0.21 | 0.20 | 0.33 | 0.32 | 0.19 | 0.19 |
| Observations | 8510 | 8510 | 8510 | 8510 | 8510 | 8510 | 6325 | 6325 | 6325 | 6325 | 6325 | 6325 |

# IA6 Robustness to Excluding Non-U.S. Outlets

This section reports the main asset pricing results when the indicators are constructed from U.S. outlets only, excluding *the Guardian* and the *Daily Mail*. Table IA6.1 reports the single predictive regressions and Table IA6.2 the second-pass Fama–MacBeth estimates. The negative and significant price of transition risk is preserved, indicating that the main results are not driven by the non-U.S. outlets.

**Table IA6.1**
**Single Predictive Regressions of Aggregate Market Returns on the Transition and Physical U.S.-Only Indices**

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 months on the current level of the Transition Risk Indicator (TRI) or the Physical Risk Indicator (PRI), constructed with only U.S. outlets. The sample runs from January 2014 to June 2025. For each horizon $q$, $q$ months of data are lost by construction. Newey and West (1987) $q$-lag t-statistics and Hansen and Hodrick (1980) $q$-lag t-statistics are reported below the coefficient estimates. The estimates are in percent. *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

| $q=$ | 1 | 3 | 12 | 24 | 36 | 48 |
|---|---|---|---|---|---|---|
| $\text{TRI}_{\text{US}}$ | -0.303 | -0.119 | -1.816 | -0.270 | -3.834 | -6.796*** |
| t-stat$_{\text{NW}}$ | (-0.92) | (-0.13) | (-0.46) | (-0.04) | (-1.29) | (-5.33) |
| t-stat$_{\text{HH}}$ | (-0.88) | (-0.16) | (-0.61) | (-0.06) | (-1.15) | (-3.10) |
| Adj. $R^2$ (%) | -0.50 | -0.73 | 0.24 | -0.87 | 7.66 | 25.14 |
| $\text{PRI}_{\text{US}}$ | -1.051 | -1.851 | 2.233 | 8.968 | 2.561 | -2.446 |
| t-stat$_{\text{NW}}$ | (-1.24) | (-1.06) | (0.25) | (0.88) | (0.46) | (-0.37) |
| t-stat$_{\text{HH}}$ | (-1.18) | (-1.32) | (0.44) | (1.46) | (0.47) | (-0.35) |
| Adj. $R^2$ (%) | 0.04 | 0.28 | -0.43 | 3.60 | -0.25 | -0.53 |

## Table IA6.2

## Fama–MacBeth Risk-Premium Estimates with Alternative Factor Controls – U.S.-Only Outlets

This table reports second-pass Fama–MacBeth risk premium estimates over the period January 2016–July 2025. 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 (TRI) and physical (PRI) climate risk factors enter as innovations. Newey and West (1987) t-statistics with 4 lags are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels.

| | 74 Portfolios | | | | | | 55 Portfolios | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | FF5 | | FFC | | FF3 | | FF5 | | FFC | | FF3 | |
| Term | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) | (i) | (ii) |
| $\Delta\text{TRI}_{\text{US}}$ | -5.210** | -5.695*** | -6.054*** | -6.617*** | -5.640** | -6.315*** | -7.995*** | -8.243*** | -9.240*** | -10.002*** | -8.518*** | -9.422*** |
| | (-2.43) | (-2.64) | (-2.66) | (-2.87) | (-2.57) | (-2.89) | (-3.17) | (-3.26) | (-3.48) | (-3.67) | (-3.16) | (-3.50) |
| $\Delta\text{PRI}_{\text{US}}$ | -2.228 | -2.087 | -1.735 | -1.466 | -3.037 | -2.586 | -2.603 | -1.623 | -1.559 | -0.721 | -3.167 | -1.743 |
| | (-0.90) | (-0.86) | (-0.62) | (-0.52) | (-1.23) | (-1.02) | (-0.80) | (-0.52) | (-0.43) | (-0.21) | (-0.95) | (-0.54) |
| MKT | 0.013** | 0.010 | 0.010 | 0.008 | 0.009 | 0.007 | 0.017*** | 0.015*** | 0.012** | 0.012** | 0.012** | 0.011** |
| | (2.12) | (1.54) | (1.47) | (1.19) | (1.47) | (1.10) | (3.23) | (3.00) | (2.14) | (2.08) | (2.33) | (2.31) |
| SMB | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.001 | -0.002 | -0.002 | -0.001 | -0.001 | -0.001 | -0.001 |
| | (-0.50) | (-0.43) | (-0.35) | (-0.29) | (-0.34) | (-0.23) | (-0.58) | (-0.55) | (-0.49) | (-0.46) | (-0.43) | (-0.41) |
| HML | 0.001 | 0.002 | 0.000 | 0.000 | 0.001 | 0.001 | 0.000 | 0.000 | -0.001 | -0.000 | -0.000 | -0.000 |
| | (0.33) | (0.43) | (0.01) | (0.05) | (0.22) | (0.32) | (0.03) | (0.06) | (-0.12) | (-0.10) | (-0.04) | (-0.03) |
| $\Delta$EPU | 159.102 | | 82.232 | | 83.912 | | 133.156 | | 8.904 | | 54.867 | |
| | (1.11) | | (0.63) | | (0.61) | | (0.80) | | (0.06) | | (0.38) | |
| MOM | | | -0.005 | -0.006 | | | | | 0.001 | 0.002 | | |
| | | | (-0.84) | (-0.86) | | | | | (0.17) | (0.25) | | |
| RMW | -0.001 | -0.001 | | | | | -0.001 | 0.000 | | | | |
| | (-0.21) | (-0.27) | | | | | (-0.26) | (0.03) | | | | |
| CMA | 0.002 | 0.002 | | | | | 0.002 | 0.001 | | | | |
| | (0.70) | (0.54) | | | | | (0.53) | (0.16) | | | | |
| Cons. | -0.003 | -0.000 | -0.000 | 0.001 | 0.001 | 0.002 | -0.007 | -0.005 | -0.002 | -0.002 | -0.002 | -0.001 |
| | (-0.58) | (-0.00) | (-0.07) | (0.21) | (0.11) | (0.48) | (-1.45) | (-1.07) | (-0.52) | (-0.38) | (-0.58) | (-0.39) |
| OLS $R^2$ | 0.46 | 0.41 | 0.42 | 0.33 | 0.43 | 0.34 | 0.62 | 0.61 | 0.56 | 0.56 | 0.55 | 0.55 |
| GLS $R^2$ | 0.27 | 0.27 | 0.32 | 0.32 | 0.28 | 0.28 | 0.28 | 0.27 | 0.36 | 0.35 | 0.26 | 0.25 |
| Observations | 8510 | 8510 | 8510 | 8510 | 8510 | 8510 | 6325 | 6325 | 6325 | 6325 | 6325 | 6325 |

# References

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.

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

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

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.

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.

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