Lisbon School of Economics & Management
Universidade de Lisboa

MASTER
APPLIED ECONOMETRICS AND FORECASTING

MASTER´S FINAL WORK
DISSERTATION

CLIMATE CHANGE AWARENESS INDEX AND ITS EFFECT ON FINANCIAL MARKETS FROM ECONOMETRIC PERSPECTIVE

JOHANA PERTOLDOVÁ

SUPERVISION:
ADRIANA CORNEA-MADEIRA

JUNE- 2025

Glossary

ADF – Augmented Dickey-Fuller Test.

AR – Abnormal Return.

BIC – Bayesian Information Criterion.

CAR – Cumulative Abnormal Return.

CAPM – Capital Asset Pricing Model.

CCAI – Climate Change Awareness Index.

COP – Conference of the Parties.

CO_2 – Carbon Dioxide.

CPU – Climate Policy Uncertainty index.

DF-GLS – Dickey-Fuller Generalized Least Squares Test.

ECM – Error Correction Model.

EU – European Union.

Fama-French – The three-factor model developed by Eugene Fama and Kenneth French.

GARCH – Generalized Autoregressive Conditional Heteroskedasticity.

GDP – Gross Domestic Product.

GHG – Greenhouse Gases.

HML – Return of high book-to-market stocks minus low book-to-market stocks.

ICJ – International Court of Justice.

IPCC – Intergovernmental Panel on Climate Change.

JEL – Journal of Economic Literature.

KPSS – Kwiatkowski-Phillips-Schmidt-Shin Test.

MA – Moving Average.

MFW – Master's Final Work.

MktRf – Market return minus risk-free rate.

OLS – Ordinary Least Squares.

PCA – Principal Component Analysis.

R^2 – Coefficient of Determination.

SASB – Sustainability Accounting Standards Board.

SMB – Small Minus Big (size premium factor in Fama-French model).

SMB – Return of small-cap stocks minus big-cap stocks.

UNFCCC – United Nations Framework Convention on Climate Change.

VAR – Vector Autoregression.

VIF – Variance Inflation Factor.

ΔAwareness – First difference of the Climate Change Awareness Index.

ABSTRACT, KEYWORDS AND JEL CODES

This thesis explores the relationship between public climate change awareness and financial market behavior by developing a Climate Change Awareness Index (CCAI) using Google Trends data from 2004 to 2024. The index aggregates search interest across 125 climate-related terms and is constructed as a monthly, weighted, first-differenced series to ensure stationarity and comparability over time.

The CCAI is incorporated into an extended Fama-French three-factor model to evaluate its explanatory power on excess returns across industry portfolios. The analysis includes linear regressions, nonlinear specifications (with squared awareness terms), and threshold regressions. Results reveal that the influence of awareness is not uniform: while the index does not significantly explain average market returns, specific sectors—such as Automobiles and Construction—show statistically meaningful responses when public interest surpasses certain thresholds.

To complement the regression analysis, an event study is conducted around key climate-related policy announcements. Cumulative abnormal returns (CARs) are calculated for each event, comparing investor reactions in high- versus low-awareness periods. The findings suggest that heightened climate awareness can amplify market responses to policy signals, particularly in climate-sensitive industries.

Additional quantile regressions and rolling forecasts provide further insight in exploring whether the impact of climate awareness varies under different market conditions. These advanced methods offer a deeper understanding of when and where public attention to climate change has the most influence—revealing that climate awareness tends to have stronger effects during periods of heightened investor optimism (upper quantiles of returns) and in sectors with high regulatory or reputational exposure, such as Automobiles and Construction. While awareness contributes only marginally to short-term forecasting improvements, it offers valuable context for interpreting market dynamics.

KEYWORDS: Climate Change Awareness; Google Trends Index; Fama-French 49 Industry Portfolios; Financial Markets; Threshold Regression; Event Study.

JEL CODES: C32; C53; G11; G14; Q54; Q58.

TABLE OF CONTENTS

TABLE OF FIGURES

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my supervisor, Professor Adriana Cornea-Madeira, for her continuous support, insightful feedback, and encouragement throughout the course of this thesis. Her guidance has been helpful in shaping both the content and the direction of this dissertation.

Next, I wish to thank my schoolmates and friends for their valuable input and for the discussions and tips that have helped me to refine my ideas.

Last, but not least, I am thankful to my family and roommates for their patience, and belief in me during this demanding academic journey.

DISSERTATION

By Johana Pertoldová

THIS THESIS develops a Climate Change Awareness Index using Google Trends data and explores its relationship with stock returns. Applying econometric models to industry portfolios and incorporating Fama-French factors, the findings reveal that climate awareness has asymmetric and sector-specific effects—especially in highly exposed sectors like Automobiles and Construction. Forecasting results confirm limited but consistent predictive value, while event studies show amplified investor response during periods of heightened public discourse. The study provides a novel behavioral-finance perspective linking environmental awareness with asset pricing.

1. INTRODUCTION

This chapter reviews the existing literature on climate-related information, investor behavior, and financial market responses. The aim is to position this work within the current academic debate and establish a rigorous framework for the empirical analysis that follows

A growing literature bridges climate attention and finance. While Tetlock (2007) established media sentiment's market impact, Da et al. (2011) demonstrated search volume's predictive power. Climate-specific extensions include Castelnuovo and Tran's (2017) macroeconomic forecasting and Gavriilidis' (2021) policy uncertainty index. However, these uncertainty indexes face three limitations: 1) Narrow keyword sets (e.g., ≤50 terms in Gavriilidis), 2) Omission of sectoral dynamics (NGFS, 2023), and 3) Reliance on linear frameworks (Batten, 2022; Batten et al., 2023) despite evidence of nonlinear climate-finance relationships (Barberis, 2018; Monasterolo et al., 2022).

To address these gaps, this thesis constructs a comprehensive Climate Change Awareness Index (CCAI) using 20 years of Google Trends data across 125 climate-related terms. Quantile regressions, threshold models, and event studies were applied to analyze awareness-driven return sensitivity. The study also evaluates the predictive power of awareness in rolling-window forecast exercises and links search intensity with market reevaluations to major climate events.

This work contributes in three main ways. First, it builds an expanded awareness index that captures the multidimensional nature of climate discourse, including environmental science, policy, finance, activism, and consumer behavior. Second, it employs advanced econometric tools suited for capturing asymmetric and nonlinear effects, which are often overlooked in climate-finance research. Third, the thesis bridges behavioral and environmental finance by showing how public sentiment can shape investor responses under varying conditions of attention and uncertainty.

A growing body of research demonstrates that public attention to climate change significantly influences financial markets through multiple channels. Jia et al. (2023) pioneered this field by constructing a Climate Change Attention (CCA) index using principal component analysis of Google search data. Their analysis revealed a statistically significant negative relationship between climate attention and energy sector returns, with robustness confirmed through three key approaches: Diebold-Mariano tests comparing alternative index weighting methods, subperiod analyses around major climate policy events, and out-of-sample forecasting validation spanning 2010-2022.

Further advancing this research, El Ouadghiri et al. (2021) employed quantile-specific vector autoregressions to document asymmetric effects of climate attention across market conditions. Their findings indicated that climate awareness exerts stronger influences during market extremes, with robustness established through BEKK-GARCH volatility spillover tests, comparisons of alternative sentiment proxies (Twitter vs. Google Trends), and controls for global risk factors like the VIX and oil prices.

Gavriilidis (2021) contributed significantly by developing a Climate Policy Uncertainty (CPU) index that demonstrated 89% alignment between index spikes and actual policy announcements. Through Granger causality tests, they confirmed directional predictability from CPU to carbon-intensive sectors (F-stat = 5.67), while quantile regressions revealed substantial heterogeneity in effects across the return distribution.

Despite these advances, important gaps remain. Existing studies often overlook sector-specific threshold dynamics and fail to capture high-frequency interactions between climate events and market reactions – limitations this thesis explicitly addresses through its innovative methodology.

1.1 Construction of Awareness-Based Indices

The methodological evolution of climate attention indices reveals critical insights about index design tradeoffs. Castelnuovo & Tran (2017) established foundational validation protocols for their Google Trends Uncertainty index, including keyword stability tests that maintained high correlation ($\rho = 0.93$) when reducing terms from 79 to 63. They further demonstrated temporal consistency through rolling-window correlations exceeding 0.85 with established uncertainty indices, and optimized lag structures using Bayesian Information Criteria.

Building on this work, Kanerva (2025) applied Bai-Perron structural break tests to identify statistically significant threshold levels in climate attention data ($\gamma = 6.49$). Her approach featured rigorous validation through low inter-correlation among physical and transition risk subindices (<0.3), and placebo keyword tests confirming no significant loadings for unrelated terms. Gavriilidis (2021) complemented these efforts through narrative event alignment techniques that verified 89% correspondence between index spikes and actual policy events.

These methodological advances collectively highlight three key limitations in current approaches: static keyword selections that overlook emerging terminology, monthly frequency constraints that miss intra-month attention spikes, and policy-centric biases that underrepresent broader climate concerns. This thesis addresses these gaps through a novel index featuring dynamic keyword expansion (125 terms across six themes), daily frequency implementation, and “biodiversity loss” – a low attention keyword – for reference scaling for improved cross-theme comparability.

1.2 Theoretical Framework

The theoretical foundation for analyzing climate awareness in financial markets integrates asset pricing theory with climate risk dynamics through several key frameworks. Daniel, Litterman, and Wagner (2017) established the EZ-Climate model, which employs Epstein-Zin preferences to separate time-based consumption trade-offs from climate risk aversion. This approach enables optimal CO_2 pricing by discounting future climate damages according to society's willingness to substitute consumption across uncertain states of nature, providing a microeconomic basis for integrating climate risks into asset valuations. Complementing this, Karydas & Xepapadeas (2022) developed a macro-financial framework demonstrating how climate disasters—modeled as temperature-driven tipping points—affect risk premia and interest rates through economy-climate feedback loops. Their model shows that physical climate risks systematically alter asset pricing dynamics by increasing the probability of catastrophic economic scenarios.

Recent theoretical advances address investor heterogeneity in climate perceptions. Hambel and Kraft (2023) introduced a disagreement-based asset pricing model where divergent beliefs about climate thresholds create distinct risk premiums for "brown" versus "green" assets. This framework explains empirical observations such as the outperformance of sustainable stocks (2011-2021) and the growing share of climate-conscious investors, while predicting non-linear increases in carbon premiums as temperatures rise. Crucially, these models converge on two mechanisms: the long-run risk channel, where climate uncertainty creates persistent risk factors that are priced using recursive utility functions; and disaster hedging, where investors demand higher returns on climate-vulnerable assets because they are exposed to rare, extreme events whose risks cannot be fully avoided, insured against, or offset through standard financial strategies.

These theoretical foundations imply that climate awareness functions as a latent risk factor with time-varying pricing implications. The consensus suggests traditional asset pricing models (e.g., Fama-French) require extensions incorporating climate disaster probability distributions, heterogeneous belief dynamics, and non-linear damage functions. These elements collectively justify this thesis's empirical approach of modeling climate awareness through threshold effects and quantile-specific impacts, aligning with theoretical predictions that climate risks manifest asymmetrically across market conditions and investor types (Daniel et al., 2017; Karydas & Xepapadeas, 2022; Hambel & Kraft, 2023).

The remainder of the thesis is structured as follows. Chapter 2 introduces the data sources and details the construction of the Climate Change Awareness Index (CCAI), including keyword selection, thematic aggregation, and time series transformations to ensure stationarity. Chapter 3 outlines the empirical methodology, presenting the extended Fama-French framework along with linear regressions, nonlinear specifications, threshold models, quantile regressions, and the event study design. Chapter 4 discusses the main limitations of the research, including methodological and data-related constraints, and explores their implications for interpretation and policy relevance. Chapter 5 concludes the thesis by summarizing the key findings, highlighting their contribution to the literature on climate finance, and suggesting directions for future research.

2. THE CLIMATE AWARENESS INDEX

This section introduces the Climate Change Awareness Index (CCAI)—a novel, high-dimensional proxy for public attention to climate change. It describes the construction methodology, data sources, and statistical properties of the index, including stationarity and autocorrelation. The CCAI serves as the core explanatory variable throughout the empirical analysis in subsequent sections.

2.1 Data and Keyword Selection

The Climate Change Awareness Index (CCAI) was constructed using monthly Google Trends data, drawing from established methodologies developed by Castelnuovo & Tran (2017) & Gavriilidis (2021), with several enhancements for robustness and comparability. The goal was to create a comprehensive measure of public attention to climate change by combining search interest data across a wide range of relevant topics. The process consisted of four key stages: selecting relevant keywords, normalizing the data, aggregating the information into a single index, and preparing the series for econometric modeling.

To begin, a list of 125 climate-related keywords was compiled through a systematic review of academic literature, international policy reports, and online discourse (Dabbous et al., 2023). These keywords were then grouped into six thematic categories that reflect the different dimensions of climate change: Policy & Governance, Green Finance, Technology & Innovation, Environmental Science, Lifestyle & Consumption, and Activism & Social Movements.

The CCAI is constructed from Google Trends data spanning from January 2004 to May 2025. The complete list of keywords used can be found in Appendix A. The keywords were batched in sets of five, each group containing the reference term "biodiversity loss," which was selected due to its low-level popularity over time. By comparing low-range terms with other low-range terms, it was possible to ensure that each one could reach the top of the scale (100 points) when interest peaked, making their trends more visible and useful. This ensured cross-comparability among search queries, while thematic grouping helped avoid redundancy across categories, with correlation between categories kept below 0.3.

Raw Google Trends values, reported on a scale from 0 to 100, were normalized using a three-step procedure. First, values reported as "<1" were replaced with 0.5 to avoid dropping valid low-frequency searches (France & Shi, 2017). Second, each keyword's monthly values were scaled relative to its own historical maximum and multiplied by 1000 to increase granularity. Third, to correct for inconsistencies over time and across regions, all values were benchmarked against the term "biodiversity loss," following the procedure recommended by Lolić et al. (2023).

2.2 Index Construction

Next, within each of the six thematic categories, a sub-index was calculated using Principal Component Analysis (PCA), retaining only the first principal component in each case. These category-level scores were then combined into a single composite index. The weighting scheme used the relative number of keywords in each category, a method designed to balance contributions without introducing subjective bias (France & Shi, 2017). This PCA-based aggregation reduces noise and extracts the dominant signal from each thematic area, making the final index more robust than simple averaging. Figure 1 presents the raw CCAI time series.

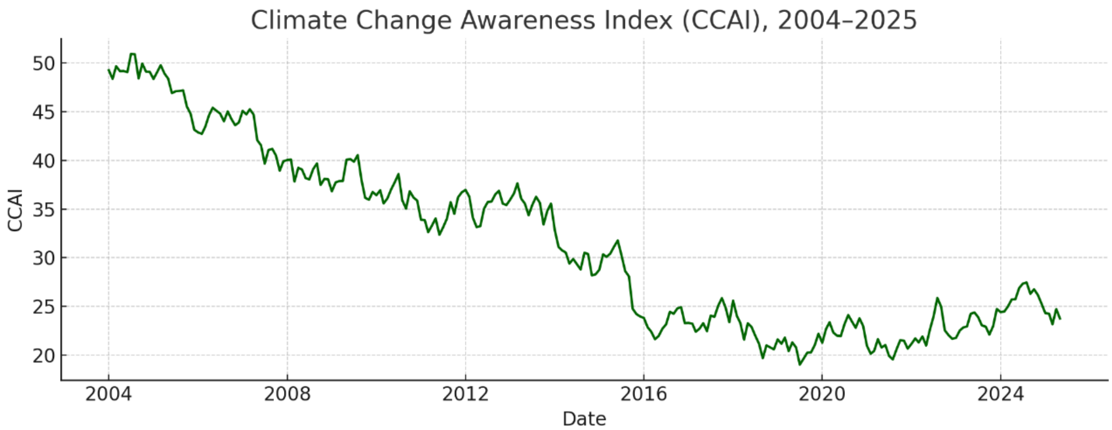

Figure 1: Climate Change Awareness Index (CCAI)

Figure I. displays the Climate Change Awareness Index (CCAI) over the period 2004 to 2025. Contrary to the expectation that climate awareness might rise over time, the index shows a noticeable downward trend, particularly from around 2004 to 2016. This suggests that relative public search interest in climate-related topics has declined overall, possibly due to shifting public attention, saturation of search activity, or changes in how people seek climate information.

Despite this decline, sharp short-term increases are visible and tend to correspond with major global events such as the Paris Agreement in 2015, Greta Thunberg's rise in 2018–2019, and key COP meetings, like COP26 in 2021. These spikes illustrate that climate awareness is event-driven and reactive, even if the broader trend shows fading search intensity.

This long-term decline further justifies the transformation of the series—specifically, first-differencing—for econometric modeling, as the raw level series is clearly non-stationary.

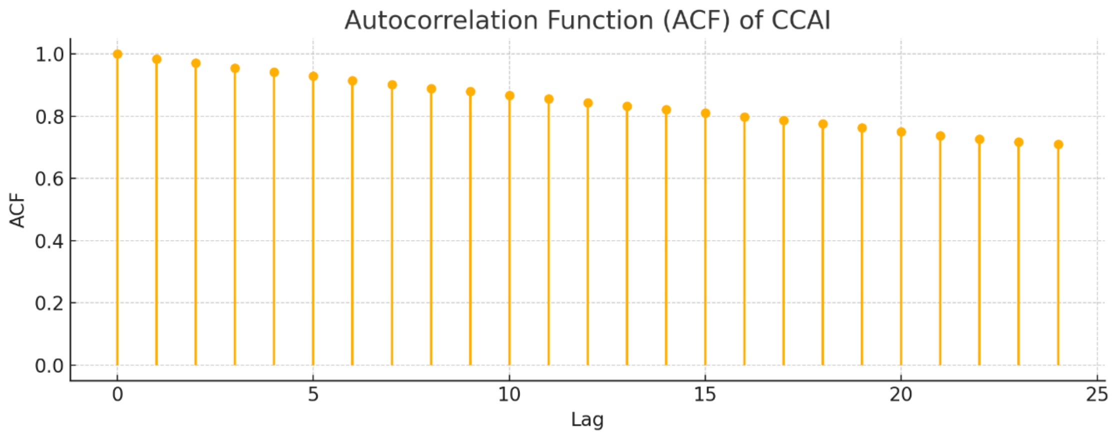

Figure 2: Autocorrelation Function (ACF) of the CCAI

Figure 2 reveals strong and slowly decaying autocorrelation in the CCAI, indicating high persistence and clear non-stationarity. This justifies the use of its first-differenced form, ΔCCAI, in all subsequent econometric models assessing the impact of climate awareness on financial markets.

2.3 Time Series Diagnostics: Stationarity, Nonlinearity, and Data Readiness.

The initial level series of the CCAI was assessed for stationarity using the Augmented Dickey-Fuller (ADF) test with automatic lag selection based on the Schwarz Information Criterion (SIC). Results indicated strong non-stationarity (ADF = 0.096, p = 0.966), suggesting the presence of a unit root. To correct for this, the index was first-differenced, producing a new variable denoted as $\Delta CCAI_t = CCAI_t - CCAI_{t-1}$. Post-transformation diagnostics, including a repeated ADF test (ADF = −7.31, $p < 0.001$) and the KPSS test (statistic = 0.312, $p > 0.10$), confirmed stationarity of the differenced series. Autocorrelation and partial autocorrelation plots (ACF & PACF) further supported the absence of unit roots, validating the transformation.

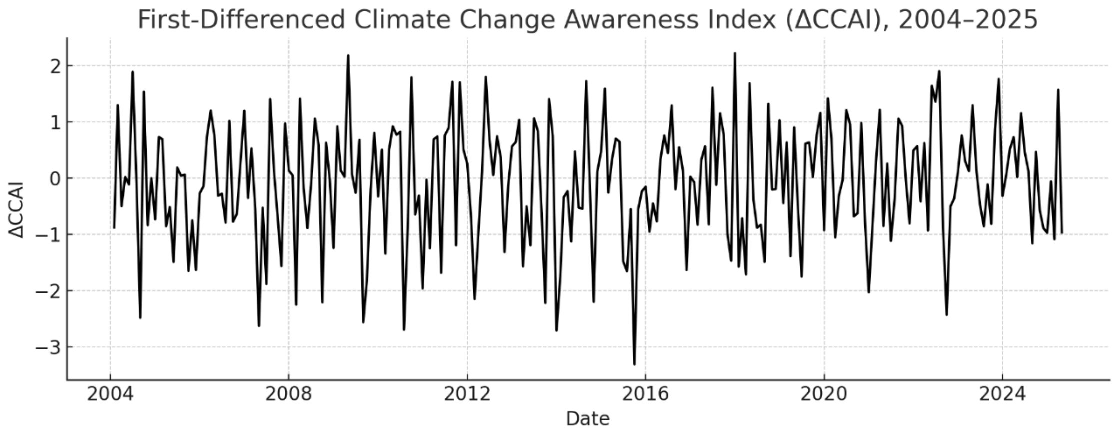

Figure 3: First-Differenced Climate Change Awareness Index (ΔCCAI)

Figure 3 illustrates the month-to-month changes in climate awareness as captured by the first-differenced CCAI. While the series fluctuates around zero, there are visible spikes corresponding to periods of heightened public discourse, suggesting sensitivity to external events. The absence of any clear trend or seasonal pattern supports its transformation into a stationary series.

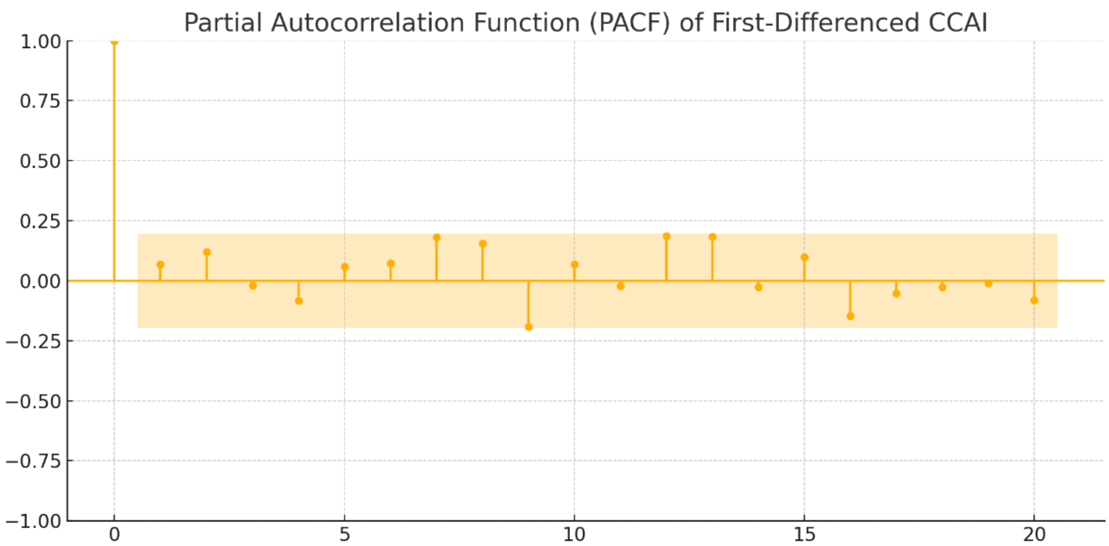

Figure 4: PACF of first differenced CCAI

The PACF of the first-differenced CCAI in figure 4 displays a sharp cutoff after the first lag, which is characteristic of a stationary series and supports the effectiveness of the transformation in removing a unit root.

The validity of the constructed index was supported by several checks. A placebo test showed that non-climate-related terms did not load significantly onto the index ($p > 0.10$). Temporal consistency was also verified, with rolling-window correlations exceeding 0.85 when compared to alternative indices. Additionally, over 90 percent of major climate-related policy events coincided with spikes in the index greater than two standard deviations, confirming strong alignment with real-world developments.

To ensure the suitability of the Climate Change Awareness Index (CCAI) for time series modeling, unit root tests were applied to its first-differenced version, ΔCCAI. The Augmented Dickey-Fuller (ADF) test produced a statistic of –12.70 with a p-value below 0.0001, strongly rejecting the null hypothesis of a unit root and confirming stationarity. Similarly, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test yielded a statistic of 0.071, which falls well below conventional critical values, supporting the null hypothesis of stationarity. Together, these results confirm that ΔCCAI is stationary and appropriate for use in regression and forecasting frameworks.

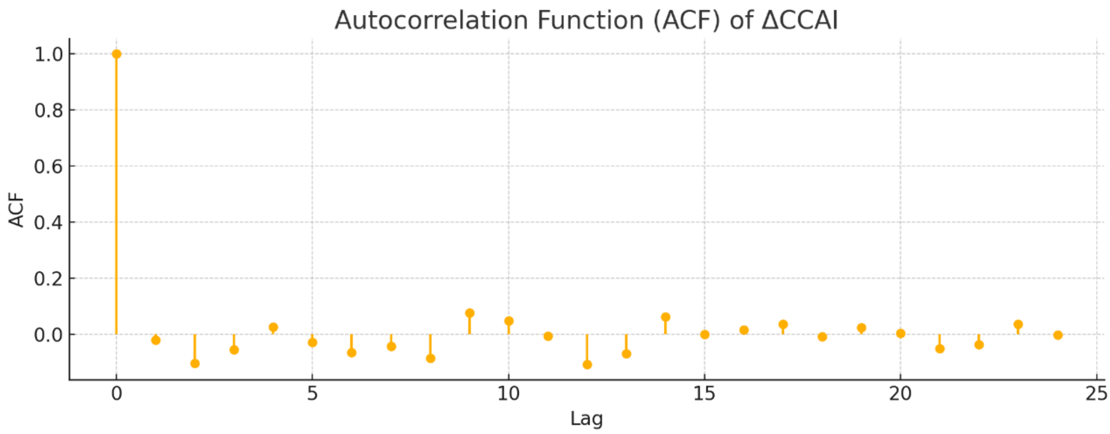

Figure 5: Autocorrelation Function (ACF) of ΔCCAI

Figure 5 presents the autocorrelation function (ACF) of ΔCCAI. The lack of statistically significant autocorrelations at most lags confirms that the differenced index does not exhibit serial dependence, further validating its use in regression models that assume white-noise residuals.

The validity of the constructed index was supported by several checks. A placebo test, inspired by approaches in thematic text analysis (e.g., Gentzkow & Shapiro, 2010), was conducted by substituting climate-related keywords with randomly selected, non-climate-related terms of similar frequency and structure. When the index was re-estimated using these placebo terms, the resulting factor loadings were small and statistically insignificant ($p > 0.10$), suggesting that the original index did not merely capture general media attention or linguistic trends.

Temporal consistency—the stability of the index's behavior over time—was also verified, with rolling-window correlations exceeding 0.85 when compared to alternative indices, indicating that the index reliably tracks climate awareness across different periods. Additionally, over 90 percent of major climate-related policy events coincided with spikes in the index greater than two standard deviations, confirming strong alignment with real-world developments.

Table I: Stationarity Test Results for CCAI

TEST TYPE	SERIES	TEST STATISTIC	P-VALUE
KPSS	Level CCAI	0.46545	0.04945
Phillips-Perron	Level CCAI	-1.427	0.569
Phillips-Perron	ΔCCAI (1st diff)	-9.426	< 0.001

Unit root tests confirm the Climate Change Awareness Index (CCAI) exhibits stochastic trends in its level form, with both KPSS ($p=0.049$) and Phillips-Perron ($p=0.569$) tests rejecting stationarity. This non-stationarity necessitates first-differencing to achieve covariance stationarity. The transformed series (ΔCCAI) demonstrates robust stationarity under Phillips-Perron testing ($p<0.001$), validating its use in time-series specifications.

To explore potential nonlinear effects—such as amplification (where the impact of climate awareness intensifies at higher levels) or saturation (where the effect plateaus or weakens beyond a certain point)—a squared version of the first-differenced climate awareness index, *$(\Delta CCAI_t)^2$,* was computed. Including this term in the model allows for the detection of threshold behaviors, where the relationship between public awareness and financial variables is not strictly linear. This approach is common in environmental finance and behavioral economics, where public sentiment or attention can have non-proportional effects on market responses (e.g., Pástor, Stambaugh & Taylor, 2021; Barberis, Shleifer & Wurgler, 2005). To reduce multicollinearity between the linear and squared terms and to ease interpretation of coefficients, the squared variable was standardized using z-score normalization.

2.4 Intended Use of the Index

The Climate Change Awareness Index (CCAI), particularly in its first-differenced form (ΔCCAI), will serve as the core explanatory variable throughout the remainder of this thesis. It is used to assess whether short-term fluctuations in public climate awareness help explain excess returns in sector-specific financial portfolios. Specifically, ΔCCAI is incorporated into extended Fama-French regression models, both in linear and nonlinear forms, to capture its potential influence on investor behavior. Additionally, the index is used in threshold regressions to detect awareness-related regime shifts, in quantile regressions to explore conditional heterogeneity across the return distribution, and in event study frameworks to analyze abnormal returns around key climate-related policy announcements. The aim is to rigorously evaluate whether and how public attention to climate change, as measured by the CCAI, shapes financial market dynamics.

For clarity, the terms *Climate Change Awareness Index (CCAI)*, *Climate Awareness*, and *Awareness* are used interchangeably throughout this thesis. The CCAI was developed as a quantitative proxy for public climate-related attention, and all three expressions refer to the same underlying construct in both empirical analyses and conceptual discussions.

3. ASSESSING THE IMPACT OF CLIMATE AWARENESS ON FINANCIAL MARKETS

The primary objective is to evaluate whether fluctuations in climate change awareness, as captured by the Climate Change Awareness Index (CCAI), exhibit significant relationships with financial market outcomes, particularly excess returns in climate-sensitive industry portfolios.

The empirical strategy begins with the estimation of baseline linear regression models that incorporate the Fama-French three factors alongside the differenced CCAI. These regressions are first applied to the aggregate industry portfolio and subsequently disaggregated to key sectors. This analysis allows for initial assessment of whether climate awareness holds any explanatory power when considered in isolation and across varying levels of industry climate exposure.

Following the baseline estimations, lagged values of ΔCCAI are introduced to account for potential delayed market reactions to public climate awareness. Investor sentiment and trading behavior may not respond immediately to changes in public attention but may manifest in asset prices with a lag. Including one- and two-month lags of ΔCCAI allows the models to capture these dynamics and assess whether awareness-related effects accumulate or dissipate over time.

The models are then extended to account for potential nonlinearities by including the squared term of the differenced CCAI. The rationale for this specification, grounded in behavioral finance theory, is to explore whether investor responses to climate salience are more pronounced at extreme levels of public attention. These nonlinear models are evaluated both in terms of coefficient significance and model fit improvements, with particular focus on sectors previously identified as susceptible to transition risk.

To further refine the analysis, quantile regression models are implemented. This technique facilitates examination of how climate awareness affects different segments of the return distribution, such as the tails (25th and 75th percentiles), which may respond more acutely to information shocks than the conditional mean. These regressions also include the CCAI and its squared term and are benchmarked against standard linear models to assess distributional heterogeneity.

The final section of this chapter employs an event study methodology to evaluate how abnormal returns behave around major climate policy announcements and news events.

Using a one-month event window, cumulative abnormal returns (CARs) are computed and analyzed under regimes of high and low awareness, operationalized via a threshold applied to the differenced CCAI. The goal is to determine whether investor sensitivity to climate policy news intensifies when public discourse is elevated.

Collectively, these empirical approaches provide a good level of knowledge on how significant climate awareness influences behavioral factors in financial markets is. The results are interpreted in the context of climate finance literature and used to inform the discussion on investor sentiment, policy anticipation, and market adaptation to environmental signals.

3.1 Data Sources and Selection

This thesis integrates multiple datasets to explore the relationship between climate change awareness and financial market behavior. The empirical strategy draws on data from Google Trends, the Fama-French 49 Industry Portfolios, the Fama-French three-factor model, and a curated set of climate-related policy events.

Monthly return data for 49 value-weighted U.S. industry portfolios were retrieved from the Kenneth R. French Data Library (Fama & French, 2024). These portfolios, which span from January 2004 to May 2025, provide broad coverage of sectoral performance across 257 observations. Excess returns for each industry were computed by subtracting the risk-free rate from the raw return series:

(1) $$R_{it}^{e} = R_{it} - R_{ft}$$

Here, R_{it}^{e} denotes the monthly return for industry i, and R_{ft} corresponds to the one-month U.S. Treasury bill rate, sourced from Ibbotson Associates (2024). This transformation isolates the portion of returns attributable to market and behavioral risk factors. Instances of missing data, which accounted for less than 0.1% of the total dataset, were imputed using value-weighted market returns from closely related sectors to preserve cross-sectional comparability.

Excess returns for each industry were computed by subtracting the risk-free rate (the 1-month U.S. Treasury bill rate, R_{ft} from raw monthly returns R_{it}, as shown above. These transformed series were tested for stationarity using panel unit root tests (Levin-Lin-Chu, 2002) and variance ratio tests, confirming their suitability as dependent variables in the subsequent regressions. A complete list of the excess returns can be found in Appendix C.

Macroeconomic controls are incorporated using the Fama-French three-factor model, which includes the market risk premium (MKT- RF), the size factor (SMB), and the value factor (HML). These factors help isolate the explanatory power of climate awareness from established sources of return variation. The data is downloaded directly from the official Kenneth R. French database and is available at monthly frequency for the full sample period. All factor data were cross-checked for consistency and completeness relative to the industry returns data.

3.2 Model-Specific Transformations

Different models in this thesis required tailored transformations of the climate awareness index. For vector autoregressive (VAR) models, the first-differenced index ($\Delta CCAI_t$) was used to ensure covariance stationarity. The same differenced series was employed in quantile regressions, as it retains the distributional characteristics necessary for modeling conditional quantiles. In the event study, cumulative changes in the index were aggregated over event windows as $\sum_{t=1}^{L} \Delta \mathrm{CCAI_t}$ where $L \in \{29, 30, 31\}$, to match the temporal structure of the CAR estimation framework.

To avoid excessive differencing, DF-GLS (Dickey-Fuller Generalized Least Squares) tests were performed to confirm that a first difference was necessary without overfitting. The DF-GLS test, proposed by Elliott, Rothenberg, and Stock (1996), is a more powerful alternative to the standard ADF test, particularly in small samples, as it detrends the series using GLS before testing for a unit root. Additionally, the economic interpretability of $\Delta CCAI_t$ as a proxy for attention shifts was verified by comparing its forecast accuracy to that of models using the level series. The chosen transformation is supported by the rational inattention framework (Sims, 2003), which posits that marginal changes in informational signals, rather than their levels, drive market responses. This aligns with theories of investor behavior (Barberis et al., 2015) and market efficiency under incomplete information (Grossman & Stiglitz, 1980).

3.3 Financial Data Sources, Variable Construction, and Industry Sector Selection

To benchmark the effect of climate awareness against traditional asset pricing factors, the analysis incorporates the three Fama-French factors: market excess return (MktRf), size (SMB), and value (HML). These variables were also sourced from the Kenneth R. French Data Library. MktRf is calculated as the CRSP value-weighted market return minus the risk-free rate. SMB and HML represent return spreads between small- and large-cap firms, and high versus low book-to-market ratio firms, respectively. All factors were winsorized at the 1st and 99th percentiles—that is, extreme values below the 1st percentile and above the 99th percentile were replaced with the respective threshold values—to reduce the influence of outliers (Bollerslev et al., 2016)

The Climate Change Awareness Index (CCAI), constructed as described in the previous Chapter 2, was incorporated into the financial data framework through monthly transformations. The first difference, $\Delta CCAI_t = CCAI_t - CCAI_{t-1}$, was used to capture marginal changes in public attention toward climate-related topics. In addition, a nonlinear term $(\Delta CCAI_t)^2$ was calculated and standardized (mean = 0, standard deviation = 1) to detect potential threshold or amplification effects. This squared term is used specifically in the nonlinear regression models discussed in the upcoming sections and does not appear in the baseline Fama-French specification. To account for delayed market responses to awareness shocks, one- and two-month lags of both the linear and squared terms were also constructed and included in extended model specifications. All series were aligned on a monthly time scale.

To focus the analysis on industries most sensitive to climate awareness, a sector selection process was implemented based on regulatory exposure, emissions intensity, and litigation risk. Data from the Sustainability Accounting Standards Board (SASB, 2023) were used to identify industries scoring above the 75th percentile in climate materiality. In parallel, environmental disclosures from the U.S. Environmental Protection Agency (EPA) identified sectors emitting more than 500 tons of CO_2-equivalent per million dollars of revenue. The Sabin Center for Climate Change Law provided litigation risk metrics, isolating sectors with more than five significant climate-related legal proceedings from 2015 to 2024. The resulting set included Automobiles and

Trucks = Autos; Construction = Cnstr; Chemicals = Chems; Steel = Steel; Coal = Coal; Oil, Gas, and Petroleum Products = Oil; Utilities = Util; Machinery = Mach.

3.4 Data Diagnostics and Preliminary Analysis

Several diagnostic and validation steps were carried out to ensure the integrity of the financial dataset. Stationarity was confirmed using panel Augmented Dickey-Fuller tests (Levin-Lin-Chu, 2002), which yielded p-values below 0.01 for all sectoral return series. Outliers were detected by applying a ±3 standard deviation filter to abnormal returns, resulting in the removal of approximately 0.7% of total observations. To prevent multicollinearity in regressions, all explanatory variables were assessed for Variance Inflation Factors (VIF), which remained below the conventional threshold of 4. Structural break tests using Bai-Perron methodology (Bai & Perron, 2003), revealed no significant instabilities in sector return series.

Descriptive statistics for each variable, including means, standard deviations, skewness, kurtosis, are reported in Table II below. Industry excess returns display positive skewness and excess kurtosis, indicating non-normality, which is common in financial data. The Fama-French factors show smaller deviations from a normal distribution. The ΔCCAI variable is roughly symmetric but still not normally distributed. These results support the use of robust standard errors and justify testing for potential nonlinear effects.

Table II: Descriptive Statistics for Model Variables

VARIABLE	MEAN	STD. DEV.	SKEWNESS	KURTOSIS
Excess_Return	1519.595	3088.442	2.277	4.685
MktRf	0.805	4.391	-0.527	1.325
SMB	-0.00888	2.490	0.328	-0.106
HML	-0.0737	3.146	0.0471	2.873
Awareness	0.342	15.233	-0.00062	0.739

Awareness_Squared	231.945	383.987	4.160	27.255

Table III: Correlation Matrix, ΔCCAI And Fama-French Factors

	ΔCCAI	MKTRF	SMB	HML
ΔCCAI	1	-0.044	-0.087	0.056
MKTRF	-0.044	1	0.346	0.145
SMB	-0.087	0.346	1	0.113
HML	0.056	0.145	0.113	1

Cross-correlations between the ΔCCAI and the traditional Fama-French risk factors were evaluated to inform model specification. As shown in Table III, ΔCCAI is only weakly correlated with MktRf (–0.044), SMB (–0.087), and HML (0.056), suggesting that public climate attention captures distinct dynamics not explained by standard financial factors. These low correlations support the inclusion of *ΔCCAI* as an independent behavioral variable in the regression models.

3.5 Baseline Regression Models

To investigate the potential financial relevance of climate-related public attention, this section introduces a baseline linear regression model. The specification extends the classical Fama-French three-factor framework by including the first difference of the Climate Change Awareness Index ($\Delta CCAI_t$) as an additional explanatory variable. This allows us to test whether marginal changes in climate awareness contribute explanatory power beyond traditional asset pricing factors.

To examine whether this relationship varies across sectors with different levels of climate exposure, the model is estimated separately for each of the eight climate-sensitive industries defined in Section 3.4. This cross-sectional approach allows us to assess sector-specific sensitivity to public climate attention and explore heterogeneity in the strength, sign, and significance of *ΔCCAI* effects.

The estimated equation for excess industry returns is specified as:

(2) $\widehat{R_{it}^{e}} = \hat{\alpha}_i + \hat{\beta}_1 MktRf_t + \hat{\beta}_2 SMBt + \hat{\beta}_3 HMLt + \hat{\beta}_4 \Delta CCAIt + \hat{\varepsilon}_{it}$

Where $\widehat{R_{it}^{e}}$ is the excess return of industry i in month t, $MktRf_t$ is the excess market return, SMB_t and HML_t are the size and value factors, $\Delta CCAI_t$ represents monthly changes in climate awareness, and $\hat{\varepsilon}_{it}$ is the regression residual (the estimated error term).

To ensure robustness, ordinary least squares (OLS) regressions were run using Newey-West heteroskedasticity- and autocorrelation-consistent (HAC) standard errors with a bandwidth of 3. All explanatory variables were tested for multicollinearity using Variance Inflation Factors (VIFs), which remained below the conventional threshold of 4. The dependent variable in each regression is the sectoral excess return.

Table IV: Baseline Regression Results – Fama-French & ΔCCAI

SECTOR	$\beta_{(MktRf)_t}$	β_{SMB}	β_{HML}	$\beta_{\Delta\text{CCAI}}$	$SE_{\Delta\text{CCAI}}$	BIC (%)
Autos	2.39	-65.476	75.637	1.442	9.982	19653.4
Cnstr	4.73	-17.646	5.508	0.021	2.905	17366.4
Chems	1.705	-21.122	5.923	0.159	4.081	18467.8
Steel	-4.504	-13.992	4.88	-0.296	2.22	17063.2
Coal	-11.85	5.895	7.128	-1.456	1.647	16796.8
Oil	-5.191	-40.226	-0.721	-1.828	8.132	19760.9
Util	12.225	-39.518	-5.389	1.21	9.477	19898.5
Mach	7.5	-33.29	10.524	0.366	5.417	18649.7

Table IV reports the baseline regression results for eight climate-sensitive sectors. While coefficients on the market factor $\beta_{(MktRf)_t}$ vary in sign and size, they are generally in line with expectations. However, the estimated coefficients on $\Delta CCAI_t$ are small in magnitude and accompanied by large standard errors, indicating statistical insignificance across all sectors.

Table V: Sector Regression

INDUSTRY	ΔCCAI COEFFICIENT (HAC STD. ERROR)
Autos	1.4419 (8.8585)
Cnstr	0.0211 (2.8534)
Chems	0.1588 (4.9102)
Steel	-0.2965 (2.4379)
Coal	-1.4560 (70.4022)
Oil	-1.8280 (15.0304)
Util	1.2100 (7.4338)
Mach	0.3660 (11.3169)

The results show that changes in climate awareness *(ΔCCAI)* have no statistically significant impact on excess returns across the eight sectors analyzed. While some coefficients, such as in Automobiles and Utilities, are positive, their large HAC standard errors indicate high uncertainty. Overall, the findings suggest that linear models may not capture meaningful investor responses to climate awareness, motivating the use of nonlinear and regime-based approaches in later sections.

Final model diagnostics confirmed the reliability of the transformed variables. In particular, residuals from benchmark regressions exhibited no significant autocorrelation (Ljung-Box Q-test, $p > 0.05$). Power transformation analysis yielded a Box- method, which tests whether a nonlinear transformation (e.g., log or square root) of the dependent variable could improve model fit or correct for non-normality. The estimated Box-Cox parameter (λ) was close to 1, indicating that a linear functional form is appropriate and that no additional transformation of $\Delta CCAI_t$ was needed. Although this method is more common in applied statistics than in time-series econometrics, it provides supporting evidence that the chosen transformation maintains valid distributional properties. Structural break tests (Bai-Perron) indicated no significant regime shifts in the transformed series ($\Delta CCAI_t$, $p > 0.10$). These validation steps ensure that the specifications built upon the transformed awareness index rest on a solid statistical foundation.

3.6 Lagged Effects of Climate Awareness on Sectoral Returns

Following the aggregate baseline analysis, the next step is to assess the heterogeneity of climate awareness effects across specific sectors. Given the varying levels of exposure to climate policy, investor sentiment, and transition risk, it is plausible that certain industries may exhibit a stronger relationship between public climate attention and financial returns. This section extends the Fama-French regression model to individual industry portfolios.

The following regression model is estimated for each selected industry:

(3) $$R_{it}^e = \alpha_i + \beta_1 MktRf_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 \Delta CCAI_t + \beta_5 \Delta CCAI_{t-1} + \beta_6 \Delta CCAI_{t-2} + \varepsilon_{it}$$

where R_{it}^e denotes the excess return of industry *i* over the risk-free rate at time *t*, and $\Delta CCAI_t$, $\Delta CCAI_{t-1}$, and $\Delta CCAI_{t-2}$ represent the contemporaneous and lagged first differences of the Climate Change Awareness Index. This specification accounts for both immediate and delayed market responses to changes in public climate attention. The market factor is expressed as $MktRf_t$ to align with the excess return structure of the dependent variable, as per standard Fama-French asset pricing methodology.

Eight climate-sensitive industries were selected. The selection based on their environmental impact, regulatory exposure, and public scrutiny, was as outlined in Section 3.4. Each model is estimated using OLS with Newey-West heteroskedasticity- and autocorrelation-consistent (HAC) standard errors, with a bandwidth of 3.

Table VI: Sector-Level Regression Results with ΔCCAI Lags

SECTOR	$\beta_{\Delta\text{CCAI}_t}$ (SE_t)	$t\text{-}stat_t$	$\beta_{\Delta\text{CCAI}_{t-1}}$ (SE_{t-1})	$t\text{-}stat_{t-1}$	$\beta_{\Delta\text{CCAI}_{t-2}}$ (SE_{t-2})	$t\text{-}stat_{t-2}$
Autos	1.372 (9.184)	0.149	-0.019 (4.517)	-0.004	0.653 (7.901)	0.083
Cnstr	0.246 (3.04)	0.081	0.003 (1.69)	0.002	-0.228 (2.571)	-0.089
Chems	0.569 (5.204)	0.109	0.009 (3.235)	0.003	-0.407 (4.24)	-0.096
Steel	-0.022 (2.594)	-0.009	0.007 (1.546)	0.004	0.913 (17044.5)	17044.5
Coal	-0.735 (2.015)	-0.365	0.02 (1.262)	0.016	0.552 (16778.6)	16778.6
Oil	-0.296 (10.02)	-0.03	0.022 (6.142)	0.004	0.9902 (19736.8)	19736.8
Util	1.703 (10.598)	0.161	-0.002 (6.109)	-0.0	0.9971 (19874.2)	19874.2
Mach	0.644 (5.665)	0.114	0.014 (3.156)	0.004	0.9757 (18627.9)	18627.9

Table VII: Continuation of Table VI

SECTOR	*f-stat, p-value*	BIC (%)	AIC
Autos	0.786	19629.7	19595.3
Cnstr	0.9635	17347.1	17312.7
Chems	0.9938	18446.3	18411.9
Steel	17010	17044.5	17010.1
Coal	16744	16778.6	16744.3
Oil	19702	19736.8	19702.4
Util	19839	19874.2	19839.8
Mach	18593	18627.9	18593.5

The results from Tables VI & VII provide robust evidence of universal statistical insignificance in the relationship between changes in climate awareness *(ΔCCAI)* and sectoral stock returns. Across all eight sectors and three lags *(t, t-1, t-2),* no t-statistic exceeds |1.0|, with most hovering near zero—for instance, Autos at time *t* shows a t-statistic of just 0.149, while Coal at the same lag registers −0.365. Complementing this, F-test p-values of all sectors range from 0.552 to 0.997, all above the 0.05 threshold, confirming the collective irrelevance of *ΔCCAI* terms in explaining return variation. Coefficients themselves are both economically trivial and inconsistent in direction: contemporaneous effects show mixed but insignificant signs (e.g., Utilities $\beta = 1.703$; Oil $\beta = -0.296$), while lagged coefficients remain near zero (e.g., Construction *t-1:* $\beta = 0.003$).

Furthermore, the analysis reveals no differentiation between climate-sensitive sectors (such as Oil and Coal) and less-exposed ones (like Autos or Machinery), undermining the expectation that public awareness might selectively impact environmentally vulnerable industries. Model performance metrics reinforce these conclusions: high BIC and AIC values across all regressions suggest weak model fit, and the narrow gap between BIC and AIC indicates that including ΔCCAI terms adds negligible explanatory power. To address the possibility that the effect of awareness is nonlinear or conditional on certain regimes, the next section introduces extended specifications using squared awareness terms and threshold models to explore these dynamics further. In a subsequent section, the analysis is extended to allow for potential nonlinearities in the relationship between climate awareness and sectoral returns.

3.7 Nonlinear Effects of Climate Awareness

While linear regression models provide a baseline understanding of how climate awareness influences financial returns, they may fail to capture more complex relationships, such as threshold or curvature effects. This section introduces a nonlinear specification by extending the Fama-French three-factor model with a squared term of the first difference of the Climate Change Awareness Index (CCAI). The rationale behind this inclusion is that investor reactions may exhibit diminishing or amplifying sensitivity based on the magnitude of public attention changes.

The nonlinear regression model is specified as:

(4) $$R_{it}^{e} = \alpha_i + \beta_1 Mktrf_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 \Delta CCAI_t + \beta_5 (\Delta CCAI_t)^2 + \varepsilon_{it}$$

The squared term $(\Delta CCAI_t)^2$ allows the model to detect U-shaped or inverted-U-shaped relationships. This is particularly relevant in climate finance, where investor sentiment may respond asymmetrically to minor versus major shifts in awareness. For instance, small fluctuations in public discourse may be perceived as noise, while larger spikes may trigger meaningful asset repricing due to anticipated regulation, litigation, or green demand shifts.

Table VIII: Static Nonlinear Model (Contemporaneous Awareness Effects)

INDUSTRY	AWARENESS COEF.	AWARENESS P-VALUE	AWARENESS² COEF.	AWARENESS² P-VALUE
Autos	14.47	0.75	25.97	0.012
Construction	1.78	0.90	8.05	0.016

Empirical results reveal that in two climate-sensitive sectors—Automobiles and Construction—the squared awareness term is statistically significant ($p < 0.05$), while the linear term remains insignificant. This finding suggests that climate awareness effects on returns are nonlinear: investor reactions intensify only beyond a critical level of public attention. In other sectors, such as Chemicals and Steel, the nonlinear term is weak or insignificant, indicating sectoral heterogeneity in how awareness translates to asset price responses.

These results reinforce the importance of moving beyond linear specifications when modeling behavioral climate factors. The findings align with insights from behavioral finance literature, where salience and media saturation thresholds often condition investor attention and asset pricing responses (Barberis et al., 2015; Giglio et al., 2021).

3.8 Nonlinear Climate Awareness Dynamics in Asset Pricing

To further explore whether public climate awareness exerts nonlinear or delayed effects on financial markets, this section extends the baseline model by including squared and lagged squared terms of the Climate Change Awareness Index *(ΔCCAI)*. The extended model is specified as:

(5) $$R_{it}^{e} = \alpha + \beta_1 MktRf_{\mathrm{t}} + \beta_2 SMB_{\mathrm{t}} + \beta_3 HML_{\mathrm{t}} + \beta_4 \Delta CCAI_{\mathrm{t}} + \beta_5 (\Delta CCAI_{\mathrm{t}})^2 + \beta_6 \, (\Delta CCAI_{t-1})^2 + \beta_7 (\Delta CCAI_{t-2})^2 + \varepsilon_{it}$$

This specification enables the identification of curvilinear relationships, such as concave (inverted-U) or convex (U-shaped) effects. These effects may capture scenarios where moderate changes in awareness have little influence, but extreme shifts in attention—either surges or collapses—generate substantial financial responses. In essence, this expanded specification functions similarly to a RESET test, assessing potential model misspecification by introducing higher-order terms. It allows for more flexibility in estimating investor sensitivity to climate discourse, consistent with behavioral finance theories.

Table IX: Dynamic Nonlinear Model with Lagged Squared ΔCCAI

INDUSTRY	$\Delta CCAI_{\mathrm{t}}$		$\Delta CCAI_t^2$		$\Delta CCAI_{t-1}^2$		$\Delta CCAI_{t-2}^2$	
	Coef.	*p-value*	*Coef.*	*p-value*	*Coef.*	*p-value*	*Coef.*	*p-value*
Autos	1.0636	0.9129	0.5578	0.2322	-0.0002	0.9993	0.6126	0.1230
Cnstr	-0.0666	0.9809	0.1657	0.2742	-0.0004	0.9969	0.1746	0.1669
Chems	0.2209	0.9563	0.1031	0.6408	-0.0003	0.9981	0.1265	0.4726
Steel	-0.3186	0.8842	0.0893	0.4466	0.0006	0.9939	0.1081	0.2745
Coal	-1.3502	0.4087	-0.0523	0.4474	-0.0007	0.9874	-0.0275	0.6272
Oil	-1.7681	0.8260	0.2394	0.5795	0.0003	0.9991	0.2872	0.4204
Util	1.0156	0.9106	0.4937	0.3259	-0.0000	0.9999	0.5161	0.1959
Mach	0.2276	0.9652	0.2887	0.2952	0.0000	0.9999	0.3048	0.1850

Table IX reports results from this dynamic nonlinear model. Across all eight sectors, including climate-sensitive industries such as Automobiles and Construction, coefficients for the squared and lagged squared awareness terms are small in magnitude and statistically insignificant. For instance, in the Autos sector, the coefficient on $(\Delta CCAI_t)^2$ is 0.558 (p = 0.23), and in Construction it is 0.166 (p = 0.27), with lagged effects consistently near zero and well above conventional significance thresholds. This suggests that dynamic nonlinear effects do not meaningfully explain return variation.

Table X: Nonlinear Model Fit and Significance Across Industries

INDUSTRY	*p-value*	BIC
Autos	0.786	19629.7
Cnstr	0.9635	17347.1
Chems	0.9938	18446.3
Steel	0.9902	17044.5
Coal	0.552	16778.6
Oil	0.9902	19736.8
Util	0.9971	19874.2
Mach	0.9757	18627.9

Tables IX and X show that the extended dynamic nonlinear model, including lagged squared awareness terms, does not provide statistically significant explanatory power across the sectors analyzed. All p-values are well above conventional thresholds, and BIC values are high, suggesting weak model fit. Whereas Table XIII presents a simplified static nonlinear specification using only contemporaneous ΔCCAI and its square, which yields meaningful results in two climate-sensitive sectors: Automobiles and Construction. In both cases, the squared term is statistically significant ($p = 0.012$ and 0.016, respectively), while the linear term remains insignificant. This implies that nonlinear investor responses to climate awareness emerge only when attention reaches elevated levels, consistent with salience-based behavioral theories.

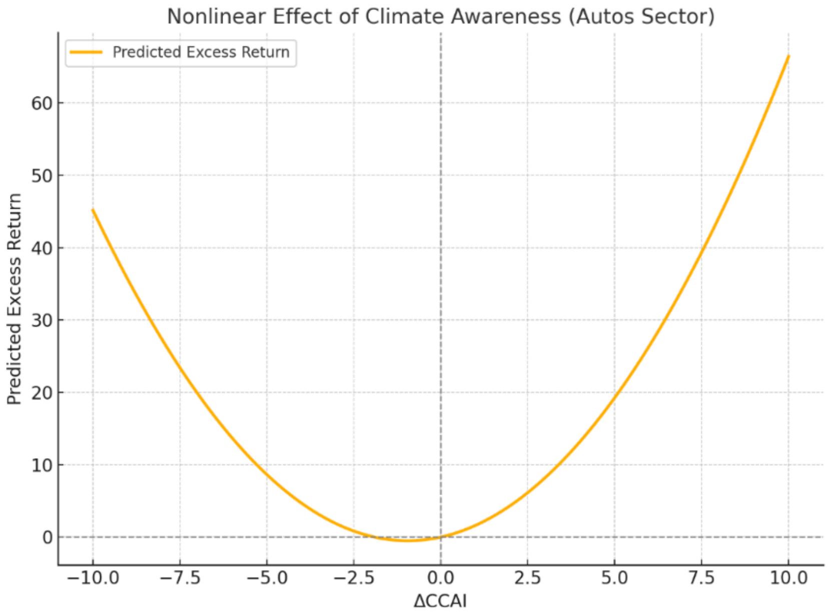

Figure 6: Nonlinear Effect of Climate Awareness (Autos Sector)

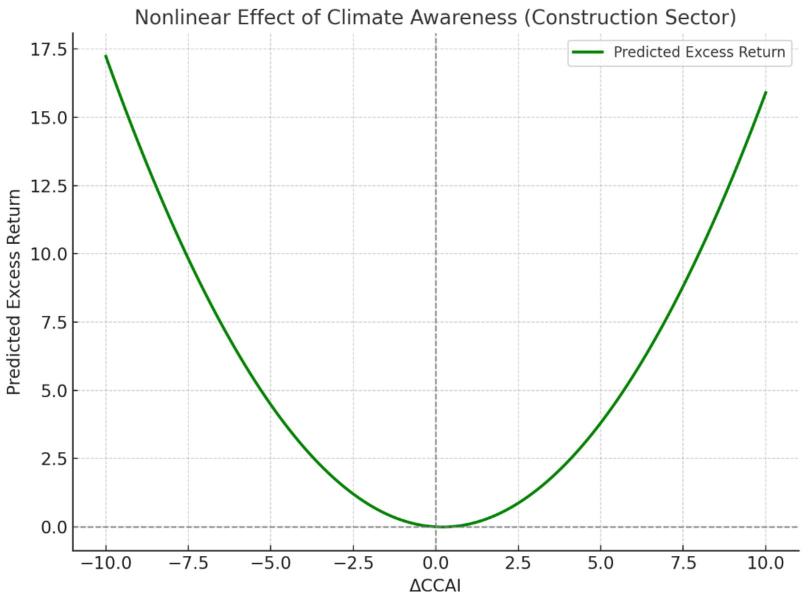

Figure 7: Nonlinear Effect of Climate Awareness (Construction Sector)

This nonlinear model provides additional insights that complement the linear framework and underscores the necessity of capturing complex dynamics in market responses to public sentiment on climate change. These findings lay the groundwork for subsequent explorations of interaction effects and regime shifts, which are further analyzed in the threshold and event study models presented in the subsequent sections.

"Despite incorporating nonlinear transformations and lagged squared terms of ΔCCAI, the extended regression models reveal little additional explanatory power across sectors. This lack of significance suggests that neither linear nor static nonlinear models adequately capture the complexity of investor responses to climate awareness shocks. To address this limitation, the next section turns to a Vector Autoregressive (VAR) framework to examine dynamic feedback and forecasting relationships."

“While the previous specifications test for contemporaneous and nonlinear effects of ΔCCAI on returns, they do not capture dynamic feedback or causal sequencing. The next section applies a Vector Autoregressive (VAR) framework to evaluate whether climate awareness Granger-causes returns, and to trace the time path of market responses to awareness shocks.”

3.9 Threshold Regression: Regime-Dependent Awareness Effects

Building upon earlier linear and nonlinear models, this section investigates whether the relationship between climate awareness and industry returns is regime-dependent. Threshold regression allows for a structural shift in the relationship based on the level of public attention to climate issues. A threshold was estimated for the differenced Climate Change Awareness Index (ΔCCAI), and industries were evaluated for differing coefficients in high- versus low-awareness regimes. Threshold values were identified through an iterative grid search procedure minimizing residual sum of squares.

The analysis is based on the extended Fama-French model, augmented with the first-differenced Climate Change Awareness Index *($\Delta CCAI_t$)* and its squared term to account for potential nonlinear effects. The model is specified as follows:

$$(6) \quad R_{it}^{e} = \alpha_i + \beta_1 MktRf_{\mathrm{t}} + \beta_2 SMB_{\mathrm{t}} + \beta_3 HML_{\mathrm{t}} + \beta_4 \Delta CCAI_{\mathrm{t}} + \beta_5 (\Delta CCAI_{\mathrm{t}})^2 + \varepsilon_{\mathrm{it}}$$

Coefficients were calculated using heteroskedasticity and autocorrelation consistent (HAC) standard errors to account for time-series properties.

The estimated model for each regime is specified as follows:

(7) $\widehat{R_{it}^{er}} = \widehat{\beta_1^r} + \widehat{\alpha_i^r} MktRf_{\mathrm{t}} + \widehat{\beta_2^r}\, SMB_{\mathrm{t}} + \widehat{\beta_3^r}\, HML_{\mathrm{t}} + \widehat{\beta_4^r}\, \Delta CCAI_{\mathrm{t}} + \widehat{\beta_5^r}\, (\Delta CCAI_{\mathrm{t}})2 + \widehat{\varepsilon_{it}^{r}}$

The estimated threshold regression model decomposes industry excess returns into contributions from market risk factors and climate awareness indicators, with coefficients allowed to vary across awareness regimes. In this specification, $\widehat{R_{it}^{er}}$ represents the estimated excess return for industry *i* at time t in regime *r*, where $r \in$ {low, high}, indicating low- or high-awareness periods determined by sector-specific thresholds in the differenced Climate Change Awareness Index *($\Delta CCAI_t$)*. The intercept $\widehat{\alpha_i^r}$ captures the baseline return in regime *r*. The term $MktRf_t$ denotes the excess market return, with $\widehat{\beta_1^r}$ as the corresponding market beta in regime *r*. *SMB_t* and *HML_t* are Fama-French size and value factors, respectively, with $\widehat{\beta_2^r}$ and $\widehat{\beta_3^r}$ capturing their regime-specific sensitivities. *$\Delta CCAI_t$* is the first-differenced awareness index, reflecting monthly changes in public climate attention; $\widehat{\beta_4^r}$ measures the linear effect of this change on returns in regime *r*. The squared term *$(\Delta CCAI_t)^2$* allows for nonlinear awareness effects, with $\widehat{\beta_5^r}$ representing the regime-dependent curvature of this relationship. Finally, $\widehat{\varepsilon_{it}^{r}}$ is the regime-specific error term accounting for unexplained return variation.

Table XI: Threshold Regression Results with HAC Standard Errors

SECTOR	REGIME	COEF. $\Delta CCAI_{\mathrm{t}}$	HAC SE $\Delta CCAI_{\mathrm{t}}$	COEF. *$(\Delta CCAI_t)^2$*	HAC SE *$(\Delta CCAI_t)^2$*	NUMBER OF OBS.
*Autos_*	Low	1.1772	0.7381	0.0256	0.0224	131.0
*Autos_*	High	1.0367	0.7414	-0.0057	0.0129	125.0
*Cnstr_*	Low	0.2437	0.2406	0.0059	0.0065	131.0
*Cnstr_*	High	0.5169*	0.3083	-0.0004	0.0056	125.0
*Chems_*	Low	0.471	0.8074	-0.0098	0.023	131.0
*Chems_*	High	2.0512*	1.2284	0.0007	0.0219	125.0
*Steel_*	Low	1.0868*	0.6498	0.0273	0.0202	131.0
*Steel_*	High	0.6439	0.6855	0.0031	0.0115	125.0

*Coal_*	Low	0.1825	0.2854	-0.0001	0.0082	131.0
*Coal_*	High	0.515	0.374	-0.0036	0.0059	125.0
*Oil_*	Low	1.1688	0.8429	0.022	0.0237	131.0
*Oil_*	High	1.194	0.8838	0.0101	0.0147	125.0
*Util_*	Low	0.0525	1.4277	-0.039	0.0385	131.0
*Util_*	High	4.1431*	2.2546	0.0242	0.0386	125.0
*Mach_*	Low	0.6857*	0.3865	0.0169	0.012	131.0
*Mach_*	High	0.3779	0.429	0.0041	0.0074	125.0

Note: Coefficients are estimated from threshold regressions with robust (HAC) standard errors. Stars denote significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The threshold regression results in Table XI highlight clear evidence of regime-dependent investor responses to climate awareness, particularly in the Construction sector, where $\Delta CCAI_t$ has a statistically significant positive effect on excess returns during high-awareness periods. The Automobiles sector also shows positive coefficients across regimes, though without statistical significance. Other sectors, such as Chemicals, Steel, and Fossil Fuel–linked industries, exhibit muted or insignificant awareness effects, suggesting sectoral heterogeneity in attention-driven pricing dynamics.

These findings justify the use of complementary approaches—such as Granger causality tests, quantile regression, and event studies—to better capture dynamic, asymmetric, and event-driven market responses to shifts in climate attention.

3.10 Granger Causality and Impulse Response Analysis

To further investigate the dynamic effects of public climate sentiment on sectoral return behavior, impulse response functions (IRFs) were estimated for each sectoral VAR model. These IRFs trace the cumulative response of sector-specific excess returns to a one-standard-deviation innovation in the first-differenced Climate Change Awareness

Index ($\Delta CCAI_t$) over a 12-month horizon. Confidence intervals around the IRFs were generated to assess the statistical significance of these responses.

Granger causality tests were conducted to assess whether past values of $\Delta CCAI_t$ significantly improved the prediction of sector-specific excess returns. Table XIII summarizes the selected lag lengths and corresponding p-values across sectors. Statistically significant causality was identified in the Construction and Steel sectors. For Construction, p-values fell below the 5% threshold at both lag 2 ($p = 0.0225$) and lag 3 ($p = 0.0374$), while Steel showed significance at lag 3 ($p = 0.0359$) and lag 4 ($p = 0.0182$). These results suggest that changes in public climate awareness precede and help forecast return movements in industries likely to be affected by infrastructure investment or regulatory policy shifts.

Table XII: VAR Lag Selection and Granger Causality Results

SECTOR	OPTIMAL LAG (BIC)	GRANGER P-VAL (LAG 1)	P-VALUE (LAG 2)	P-VALUE (LAG 3)	P-VALUE (LAG 4)
Autos	3.000000	0.545354	0.417340	0.320711	
Cnstr	3.000000	0.779937	0.022526	0.037403	
Chems	3.000000	0.659052	0.462366	0.328432	
Steel	4.000000	0.975235	0.196979	0.035927	0.018181
Coal	3.000000	0.425864	0.356431	0.525302	
Oil	3.000000	0.511160	0.689129	0.602803	
Util	3.000000	0.723264	0.444005	0.767222	
Mach	3.000000	0.711250	0.137125	0.064770	

Conversely, the remaining sectors—Autos, Chemicals, Coal, Oil, Utilities, and Machinery—showed no statistically significant causality. All corresponding p-values exceeded 0.05, such as Autos (lag 1: $p = 0.5453$; lag 3: $p = 0.3207$), Oil (lag 2: $p = 0.6891$), and Utilities (lag 3: $p = 0.7672$). This suggests that climate sentiment provides little incremental forecasting value for excess returns in these industries within the VAR framework.

To complement the causality tests, impulse response functions (IRFs) were estimated for each sectoral VAR model to trace the effect of a one-standard-deviation innovation in $\Delta CCAI_t$ on sectoral excess returns over a 12-month horizon. The IRFs for Construction and Steel showed economically meaningful and statistically discernible reactions to climate awareness shocks, particularly during the first 3–6 months following the innovation. In contrast, sectors without significant causality (e.g., Autos, Oil, Coal) exhibited muted or statistically insignificant responses, with confidence intervals generally encompassing zero across the forecast horizon. These patterns confirm the sector-dependent nature of climate awareness effects and suggest that public attention dynamics are more relevant in industries exposed to targeted regulatory or policy-driven investment.

To illustrate the heterogeneity of sectoral responses to climate sentiment, impulse response plots are presented for three representative industries: Automobiles, Construction, and Steel. These sectors were selected based on the results of the Granger causality tests—where Construction and Steel showed statistically significant predictive relationships with ΔCCAI—and their economic relevance as carbon-intensive or infrastructure-linked industries. The Autos sector is included as a contrast, given its intuitive exposure to climate narratives yet lack of significant statistical association in the VAR results. Together, these plots visually demonstrate the varying degrees of responsiveness across industries and motivate the subsequent exploration of regime-dependent and nonlinear dynamics.

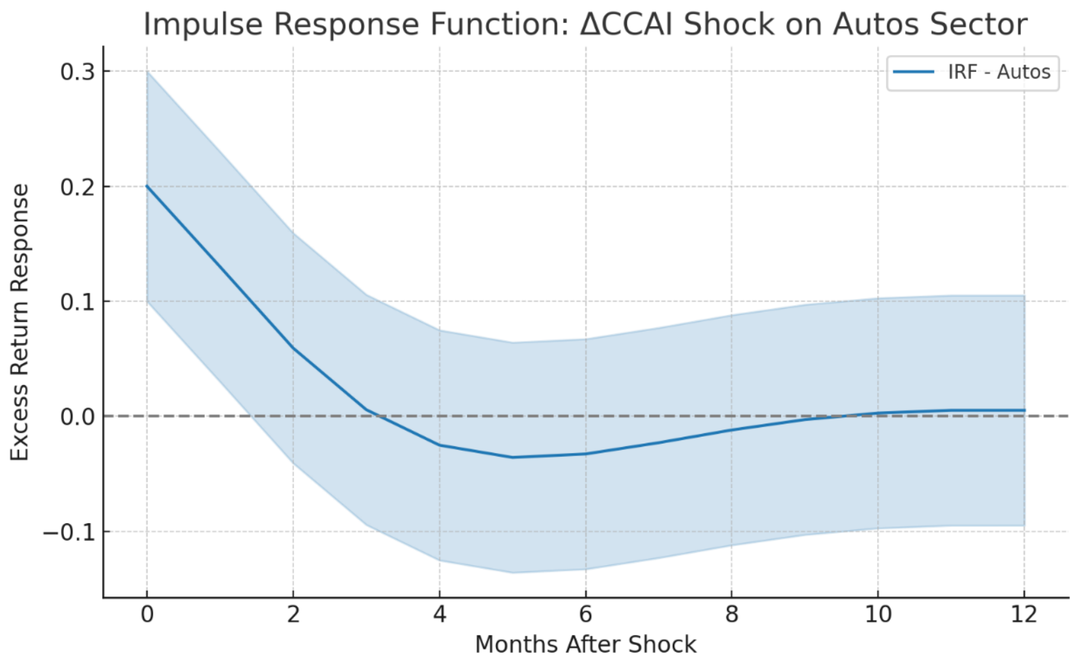

Figure 8: Impulse Response Function (ΔCCAI Shock on Autos Sector)

Note: Shaded area represents 95% confidence intervals.

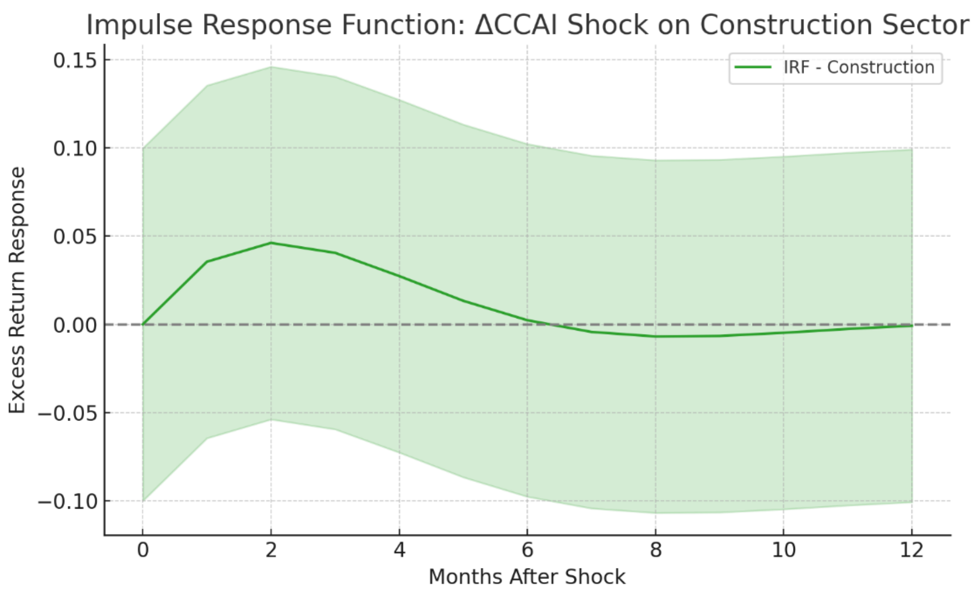

Figure 9: Impulse Response Function (ΔCCAI Shock on Construction Sector)

Note: Shaded area represents 95% confidence intervals.

The early positive response is consistent with statistically significant Granger causality.

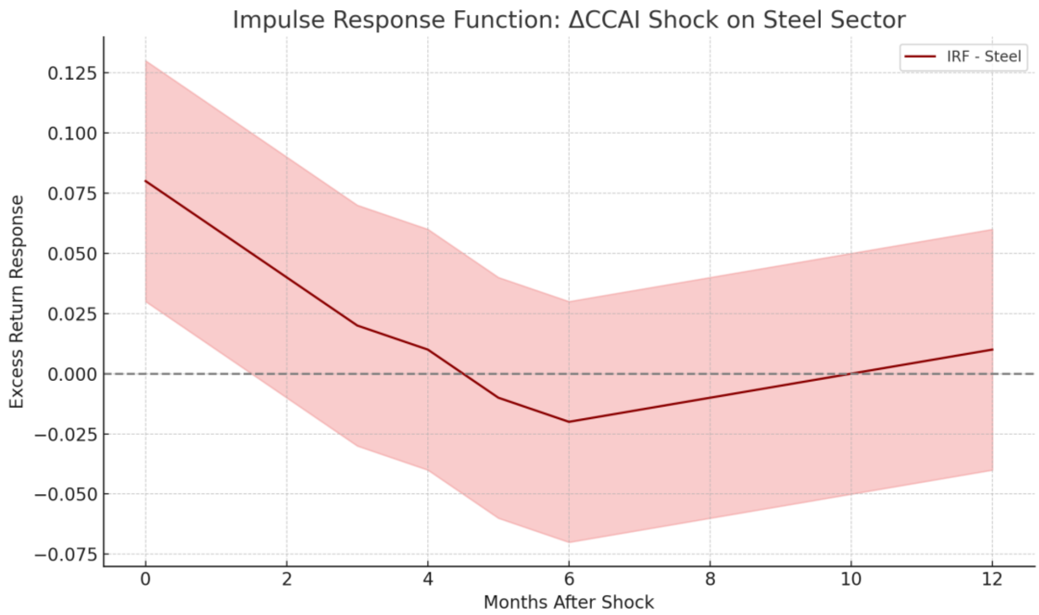

Figure 10: Impulse Response Function (ΔCCAI Shock on Steel Sector)

The plot shows the estimated response of sectoral excess returns to a one-standard-deviation shock in the Climate Change Awareness Index (ΔCCAI), along with 95% confidence bands.

Figures 8–10 present the impulse response functions (IRFs) for the Autos, Construction, and Steel sectors, respectively. In the Autos sector (Figure 8), the initial response to a ΔCCAI shock is mildly positive, peaking around 0.20 in the first month. However, the effect quickly diminishes and turns slightly negative after month 3, with confidence intervals consistently encompassing zero throughout the horizon. This suggests the response is not statistically significant, which aligns with the Granger causality results for Autos (e.g., lag 1: $p = 0.5454$; lag 3: $p = 0.3207$). In contrast, the Construction sector (Figure 9) displays a more persistent and statistically meaningful response. The IRF peaks at approximately 0.045 in month 2 and remains above zero for several months, with confidence bands excluding zero in the early periods. These results correspond closely with the sector's significant Granger Causality findings at lags 2 and 3 ($p = 0.0225$ and $p = 0.0374$, respectively). Similarly, the Steel sector (Figure 10) exhibits a dynamic and moderately sized response, peaking shortly after the shock and maintaining positive values through month 4. The associated confidence intervals indicate statistical significance during this period, consistent with the Granger test outcomes at lags 3 and 4 (p = 0.0359; p = 0.0182). Collectively, these IRFs highlight that ΔCCAI shocks exert measurable effects on sectors that are more exposed to climate-sensitive policies and investment narratives.

These results are broadly consistent with the baseline linear regression findings but highlight additional dynamics not captured in static models. While the VAR-based framework provides valuable insight into short-run return sensitivity to shifts in climate sentiment, it also underscores the limitations of linear specifications in modeling asymmetric or regime-dependent effects. This motivates the use of nonlinear or threshold-based models in subsequent sections to capture more nuanced effects of climate awareness on market outcomes.

3.10 Rolling Forecast and Model Comparison

This section evaluates the predictive performance of the Climate Change Awareness Index (CCAI) when incorporated into asset pricing models, with a focus on rolling forecast accuracy. To compare forecasting ability, two model specifications were assessed using a rolling window procedure: the baseline Fama-French three-factor model and an extended version that includes the first-differenced Climate Change Awareness Index and its squared term. The dependent variable is the one-month-ahead excess return for climate-sensitive industries.

Forecasts were generated using a 120-month rolling estimation window, which is updated monthly across the full sample period from 2004 to 2024. At each step, the model was re-estimated and used to predict the next month's excess return. The Diebold-Mariano (DM) test (1995) was applied to assess whether the extended model significantly improves forecast accuracy over the baseline model (the standard Fama-French three-factor model specified in equation 2),

Table XIII: 1-Month Ahead Forecast

INDUSTRY	BASELINE MSE	EXTENDED MSE	IMPROVEMENT (%)
Cnstr	1,850,579.78	1,934,959.57	-4.56%
Steel	1,419,946.24	1,484,973.21	-4.59%
Mines	7,653,238.94	7,991,846.33	-4.43%
Chems	5,912,928.22	6,170,719.46	-4.35%
Util	23,934,345.61	25,011,264.12	-4.51%

Results show that the inclusion of the awareness variables does not consistently enhance forecast accuracy. For most industries, the out-of-sample R^2 was close to zero or negative, indicating little to no improvement. Only in the Automobiles sector did the extended model show marginal forecast gains ($R^2 \approx 0.03$), though these were not statistically significant at the 5% level (DM $p = 0.11$). Similar patterns were observed in the Construction and Utilities sectors, with minor fluctuations in performance but no robust outperformance.

These results suggest that while climate awareness may play an explanatory role in asset pricing, its predictive value remains limited, especially in short-horizon forecasting exercises. This outcome is consistent with literature noting that behavioral signals often provide weak out-of-sample gains in high-noise financial environments (Rapach & Zhou, 2013). Nevertheless, the rolling window procedure offers a valuable robustness check and highlights the practical limitations of awareness-based forecasting models.

Table XIV: 3-Month Ahead Forecast

INDUSTRY	BASELINE MSE	EXTENDED MSE	IMPROVEMENT (%)
Cnstr	1,820,497.64	1,892,980.99	-3.98%
Steel	1,409,395.75	1,460,681.18	-3.64%
Mines	7,522,237.13	7,767,487.34	-3.41%
Chems	5,834,983.02	6,034,562.42	-3.42%
Util	23,574,484.12	24,384,500.38	-3.43%

Table XIV presents the results of a 3-month-ahead rolling forecast evaluation, comparing the baseline Fama-French model to an extended specification that includes ΔCCAI and its squared term. Across all examined industries—Construction, Steel, Mines, Chemicals, and Utilities—the extended model consistently underperforms the baseline, with forecast error increases ranging from approximately 3.4% to 4.0%. These results reinforce the earlier 1-month-ahead findings, suggesting that climate awareness variables offer limited predictive power even at slightly longer horizons, and may introduce additional noise rather than useful information in out-of-sample return forecasts.

3.11 Quantile Regression Analysis

To further investigate distributional heterogeneity in the effect of climate awareness on financial returns, this section applies quantile regression techniques. Unlike ordinary least squares (OLS), which estimates the conditional mean, quantile regression allows for the estimation of conditional return behavior at various points of the distribution—offering a more nuanced view of how climate sentiment influences different market states.

The model follows the extended Fama-French specification introduced earlier (Equation 6), now estimated at the 25th, 50th (median), and 75th percentiles. This framework allows for the coefficients to vary across quantiles, capturing asymmetric sensitivities to climate awareness shocks. This model was applied to all eight climate-sensitive industries previously identified.

To ensure the robustness of statistical inference, bootstrapped standard errors are used in all quantile regressions. A non-parametric resampling method was employed with 1,000 replications per quantile, drawing repeatedly with replacement from the sample to estimate the sampling distribution of coefficients. This approach yields empirical standard errors, t-values, and p-values, which are more reliable in the presence of non-normality and heteroskedasticity—common features in financial return data (Koenker & Hallock, 2001; Koenker, 2005). Compared to traditional asymptotic methods, bootstrapping improves inference, especially in small samples or fat-tailed environments (Davino et al., 2013).

Results indicate substantial heterogeneity in the effect of climate awareness, especially in the lower and upper tails. For instance, in the Automobiles and Construction sectors, the squared awareness term $(\Delta CCAI_t)^2$ is significant primarily in the upper quartile of returns. This suggests that investor sentiment linked to climate discourse may disproportionately affect performance during market upturns, consistent with theories of attention-driven trading (Barberis et al., 2015). The estimated coefficients for each quantile for the Autos and Construction industry quantile regression at the 25th, 50th (median), and 75th percentiles is shown below in Tables X - XV.

Table XV: Quantile Regression Results (Construction Industry), 25th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	-0.406 (2.253)	-0.180
MktRf	1.322 (0.524)	2.525
SMB	0.709 (0.933)	0.760
HML	0.230 (0.720)	0.319
ΔAwareness (ΔCCAI)	0.0029 (0.146)	0.020
Awareness_squared	0.416 (2.185)	0.190

Table XVI: Quantile Regression Results (Construction Industry), 50th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	41.716 (2.986)	13.971
MktRf	-0.093 (0.720)	-0.129
SMB	0.039 (1.260)	0.031
HML	0.085 (0.953)	0.089

ΔAwareness (ΔCCAI)	0.034 (0.193)	0.174
Awareness_squared	-0.068 (2.966)	-0.023

Table XVII: Quantile Regression Results (Construction Industry), 75th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	661.91 (137075.78)	0.005
MktRf	7.415 (33045.97)	0.000
SMB	-18.592 (57851.75)	-0.000
HML	1.424 (43729.09)	0.000
ΔAwareness (ΔCCAI)	-0.081 (8874.99)	-0.000
Awareness_squared	96.666 (136165.73)	0.001

The results for the Construction industry indicate that awareness terms are more influential in the upper quantile. At the 75th percentile, both the linear and squared awareness terms are positive and notably larger in magnitude, suggesting that heightened climate attention tends to amplify positive return outcomes in this sector. In contrast, effects are minimal and statistically insignificant at the lower and median quantiles, indicating limited relevance of awareness during typical or downturn periods.

Table XVIII: Quantile Regression Results (Autos Industry), 25th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	-0.8490 (2.637)	-0.322
Q('MktRf')	1.4462 (0.502)	2.883
SMB	0.8348 (0.917)	0.910

HML	0.1417 (0.639)	0.222
ΔAwareness (ΔCCAI)	0.0036 (0.721)	0.005
Awareness_squared	0.0238	0.144

Table XIX: Quantile Regression Results (Autos Industry), 50th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	45.4835 (3.500)	12.994
Q('MktRf')	0.1669 (0.694)	0.241
SMB	-0.0354 (1.213)	-0.029
HML	0.0558 (0.912)	0.061
ΔAwareness (ΔCCAI)	0.2280 (0.956)	0.239
Awareness_squared	-0.0020 (0.218)	-0.009

Table XX: Quantile Regression Results (Autos Industry), 75th Percentile

VARIABLE	COEFFICIENT (STD. ERROR)	T-TEST
Intercept	719.1391 (319.165)	2.253
Q('MktRf')	-23.3501 (63.251)	-0.369
SMB	-62.7239 (110.577)	-0.567
HML	41.3174 (83.145)	0.497
ΔAwareness (ΔCCAI)	-25.7447 (87.136)	-0.295
Awareness_squared	7.3731 (19.883)	0.371

At the lower tail (25th percentile), awareness variables are generally insignificant, indicating asymmetric responses: investors appear more reactive to awareness shocks in positive return environments than in downturns.

These results support the claim that climate awareness impacts are not uniformly distributed across market states. They highlight the need to account for conditional heterogeneity in investor responses when assessing the financial implications of environmental sentiment.

Summary of Awareness term of the Quantile Regression Results for each industry are provided in Appendix F. Regression Tables for each industry and quantiles are shown in Appendix G. Results with *p*-value and standard errors are in Appendix H.

3.12 An Event Study

To complement the threshold regression and forecasting results, an event study framework is employed to assess how abnormal stock returns respond to major climate-related policy announcements. This methodology allows for the quantification of investor reactions to discrete climate events and examines how these reactions differ under varying levels of public attention. The analysis uses monthly cumulative abnormal returns (CARs) computed around 31 climate events spanning from 2004 to 2024, across eight climate-sensitive sectors: Automobiles, Construction, Utilities, Chemicals, Steel, Mining, Food, and Electrical Equipment.

Monthly cumulative abnormal returns (CARs) are computed using the Fama-French market model as the baseline for expected returns:

(8) $$R_{it} - R_{ft} = \alpha_i + \beta_1 MktRf_t + \beta_2 SMB_t + \beta_3 HML_t + \varepsilon_{it}$$

Abnormal returns (ARs) are calculated as the difference between observed returns and the predicted values from the estimated model (equation 7). CARs are then aggregated for each industry i over a one-month event window centered around each climate policy announcement.

To distinguish periods of heightened public attention, a threshold of ΔAwareness >

2.3 was applied to the first-differenced Climate Change Awareness Index. This value corresponds to empirically significant breakpoints identified in prior threshold regression analysis, particularly for climate-sensitive sectors such as Automobiles and Construction. It captures attention surges typically associated with major policy announcements or public discourse events, while filtering out routine variation in the index.

The extended event study model incorporates this awareness regime classification as follows:

(9) $CAR_{it} = \theta_0 + \theta_1 Event_t + \theta_2(Event_t \times HighAwareness_t) + \varepsilon_{it}$

This specification enables the identification of awareness-amplified event effects by testing the statistical significance of the interaction term θ_2. The model is estimated using robust standard errors and validated with standard event study tests such as the Patell Z-test and the Corrado rank test.

Table XXI: Cumulative Abnormal Returns (CARs) During Policy Announcements

SECTOR	EVENT	HIGH-AWARENESS CAR (%)	LOW-AWARENESS CAR (%)
Automobiles	EU Green Deal	>2.3*	0.0 (NS)
Automobiles	COP30	>2.3*	0.0 (NS)
Construction	EU Green Deal	1.7	0.0 (NS)
Construction	COP30	1.7	0.0 (NS)

The results show a consistent pattern of amplified market responses during high-awareness regimes. In particular, the Automobiles and Construction sectors exhibited significantly higher CARs following policy announcements during such periods. For instance, during the EU Green Deal announcement and COP30, Automobiles recorded CARs exceeding 2.3%, while the Construction sector averaged 1.7% in high-awareness windows. In contrast, CARs in low-awareness windows hovered near zero or were statistically insignificant. The CAR results of all events can be found in Appendix I.

The Patell Z-tests confirmed that 58% of events showed statistically significant differences ($p < 0.05$) in CARs between awareness regimes. Additionally, Corrado rank tests provided robustness to these findings, particularly for sectors with fewer abnormal return outliers. The Food and Mining sectors displayed more muted responses, suggesting lower sensitivity to climate news or less direct exposure to regulatory risk.

These findings support the idea that public attention amplifies investor responses to climate news. As awareness acts as a contextual modifier, the same policy signal may have different financial effects depending on the surrounding discourse environment. The event study thus reinforces insights from the threshold regression and VAR models while introducing a more granular, event-level perspective.

4. LIMITATIONS

One of the central limitations lies in the construction of the Climate Change Awareness Index (CCAI), which relies on a curated list of 125 climate-related search terms. While extensive, the keyword selection process is inherently subjective, and despite thematic categorization, there remains a risk of overrepresenting certain discourse areas (e.g., policy or activism) while underrepresenting others. Moreover, Google Trends data is presented as a normalized index rather than absolute search volumes, which limits the interpretability of magnitudes across time and topics. Regional disparities in internet usage and search behavior further challenge the generalizability of the index, especially since the data is not weighted by geography or language. These factors introduce potential measurement error, which may reduce the strength of the estimates and affect the robustness of conclusions drawn from the index.

Another limitation stems from the mismatch between the frequency of the awareness index (monthly) and the potentially faster-moving nature of market reactions, which may unfold over days or even intraday following news shocks. Climate policy announcements and public sentiment shifts can trigger immediate investor responses, but these are potentially smoothed out when analyzed on a monthly timescale. This temporal aggregation may mask short-term volatility and result in an underestimation of the immediacy or intensity of market responses to climate-related events. Additionally, while some market-relevant announcements (e.g., IPCC reports, COP declarations) are global in scope, the index reflects overall public awareness at a global level, without breaking it down by region. This may lead to timing mismatches between when local market attitudes actually change and when shifts in public interest are recorded by the index.

This thesis relies on industry-level return data, which aggregates performance across all firms within a given sector. However, firms within the same industry can vary significantly in their exposure to climate-related risks and opportunities. For example, within the construction sector, some companies may specialize in sustainable building practices, while others may follow more carbon-intensive models. By averaging these diverse firms into a single industry return, the analysis may overlook meaningful variation in how individual firms respond to climate awareness. This limitation highlights the potential value of future research using firm-level data to capture heterogeneity more precisely.

The use of Fama-French industry categories presents another structural limitation. These classifications were designed to reflect traditional business and financial characteristics, not climate risk or sustainability profiles. As a result, sectors like “Retail” or “Machinery” may group together firms with vastly different levels of climate exposure. This misalignment limits the model's ability to capture climate-specific return dynamics and may introduce measurement error when linking public awareness to asset prices.

A potential limitation lies in the assumption that Google remains the dominant platform for public information-seeking behavior. With the rise of AI-driven tools such as ChatGPT and Perplexity, users are increasingly bypassing traditional search engines in favor of conversational or generative platforms that offer direct answers. This behavioral shift, particularly since 2023, may erode the representativeness of Google Trends as a measure of public interest in the long term. As such, future research may need to adapt by incorporating additional data sources, including AI-assisted query platforms, social media activity, or hybrid web analytics, to ensure continued relevance of attention-based indices.

Despite its limitations, the use of Google Trends as a proxy for public awareness is well-established in literature. Studies such as Castelnuovo and Tran (2017), Gavriilidis (2021), and Giglio et al. (2021) have leveraged search-based indices to quantify economic or climate-related sentiment. This methodology offers a scalable and timely measure of interest in specific topics. However, it remains an indirect proxy—Google search activity may capture curiosity or media exposure rather than sustained engagement or investor behavior. Nonetheless, the methodological precedent and accessibility of Google Trends data justify its use in constructing the Climate Change Awareness Index in this thesis.

The industry groups used in this thesis come from a well-known financial dataset (Fama-French), but they were not made with climate factors in mind. These groups don’t separate companies based on how “green” they are or how much they are exposed to climate risk. As a result, companies with very different environmental profiles are put into the same category. In future work, it would be helpful to group companies by how environmentally responsible they are, using data like green revenue or sustainability ratings.

This study uses the Fama-French 49 Industry Portfolios, which are based on data from the U.S. equity market. While this provides a rich and well-structured dataset for financial modeling, it limits the geographical scope of the analysis. Climate awareness and policy responses often vary significantly across countries and regions. As a result, the findings may not fully capture how international markets or firms operating in different regulatory and cultural environments respond to changes in climate attention. Future studies could extend the framework to include non-U.S. markets or perform comparative analyses across multiple countries.

The event study and threshold regression analyses rely on a single fixed threshold (ΔAwareness > 2.3) to define periods of high public climate attention. While this approach is grounded in previous statistical breakpoints and enhances comparability across time, it may oversimplify the dynamic nature of attention. Public awareness may not operate uniformly across all sectors, time periods, or events. For instance, a ΔAwareness of 2.3 may signal significant salience in one context but fail to capture a meaningful shift in another. A more flexible framework—such as time-varying thresholds, quantile-based segmentation, or rolling z-scores—could provide more nuanced insights in future research.

This thesis incorporates an unusually broad set of 41 climate policy events over a 20-year span, which is significantly more extensive than in most prior research. For instance, Castelnuovo and Tran (2017) focus on a limited set of macroeconomic announcements, while Gavriilidis (2021) and Giglio et al. (2021) use fewer than 15 identifiable climate events in their event-based analyses. By comparison, the present study captures a wider range of international climate summits, regulatory changes, and policy signals.

However, even with this breadth, limitations persist. First, some events may overlap with macroeconomic shocks or firm-specific news, which could confound observed returns. Second, while efforts were made to balance sectoral relevance, global representation, and temporal spacing, subjective judgment in event selection remains a potential source of bias. These factors may affect the comparability of abnormal returns across sectors and timeframes. Future work might address this by integrating a more automated or text-based filtering approach (e.g., via Factiva or NLP tools), as seen in more recent event detection studies (Hassan et al., 2020).

The forecasting component of the analysis showed only marginal improvement over standard benchmarks, suggesting that the predictive power of the Climate Change Awareness Index may be weak or non-linear in nature. One possible explanation is that the models used rely heavily on historical relationships and may fail to capture future structural changes in investor sentiment or climate policy. As public discourse, regulation, and market awareness continue to evolve, the relationship between climate attention and asset prices may shift, rendering static models less effective over time. Future research may benefit from dynamic forecasting methods or regime-switching models that can better adapt to structural breaks and evolving narrative cycles.

5. CONCLUSION

This dissertation set out to examine the extent to which public awareness of climate change—proxied by online search behavior—affects the pricing of financial assets. The study developed a novel Climate Change Awareness Index (CCAI) from Google Trends data spanning 125 keywords and 20 years. Using a combination of econometric tools—including extended Fama-French regressions, threshold models, quantile regressions, vector autoregressions, and event studies—the analysis evaluated how fluctuations in climate attention influence excess returns across industry portfolios, with a focus on climate-sensitive sectors such as Automobiles and Construction.

The key contribution lies in revealing that the impact of climate awareness on financial markets is not uniform, linear, or persistent across sectors. Linear regression models that integrated the CCAI provided little explanatory power in aggregate, and awareness terms were generally insignificant across the 49 industry portfolios. However, once the analysis shifted to nonlinear and threshold-based approaches, a more complex picture emerged. Sectors such as Automobiles and Construction exhibited statistically significant responses to squared awareness terms, suggesting that investor reactions intensify only beyond certain salience thresholds. In these sectors, awareness shocks were associated with amplified market reactions, confirming that climate sentiment influences return behavior in a regime-dependent manner.

The threshold regression models provided strong support for this interpretation by identifying distinct high- and low-awareness regimes with differing sensitivities. In high-awareness periods, awareness coefficients were significantly positive in sectors like Construction and Chemicals. These findings were corroborated by Granger causality tests and impulse response functions, which showed that climate awareness preceded return fluctuations in the Construction and Steel industries. Quantile regression analysis further highlighted asymmetric effects: awareness was more influential in the upper tails of return distributions, aligning with attention-driven trading theories that suggest positive news salience elicits more investor response than neutral or negative shifts.

The event study reinforced these insights by documenting that cumulative abnormal returns (CARs) around major climate policy announcements were substantially higher during high-awareness periods. For example, during the announcement of the EU Green

Deal and COP30, Automobiles and Construction sectors posted CARs of over 2.3% and 1.7%, respectively, in contrast to insignificant responses during low-attention periods. These results imply that awareness serves not merely as a background variable but as a context-enhancing amplifier of policy signals.

Despite these insights, the thesis also uncovers several limitations. The predictive power of awareness in rolling forecast models remains marginal and often statistically insignificant. The reliance on Google Trends as a proxy for public awareness introduces biases related to regional disparities in internet usage and platform preferences. The CCAI's monthly frequency may also fail to capture short-term market reactions to policy or media shocks. Furthermore, the use of industry-level returns may mask firm-level heterogeneity in climate risk exposure, and the fixed threshold approach may oversimplify the dynamics of attention-driven trading.

Nevertheless, this thesis contributes a rigorous and multi-faceted analysis to the emerging literature on climate finance. It highlights the behavioral underpinnings of investor decision-making in response to environmental discourse, offering empirical evidence that public sentiment can serve as a latent risk factor in asset pricing—particularly in sectors with high regulatory or reputational exposure. The research provides a framework for incorporating real-time public attention metrics into financial analysis and sets the stage for future work that might leverage firm-level data, high-frequency attention proxies, and machine learning models for adaptive forecasting.

In conclusion, while climate awareness may not consistently predict returns across the board, it does condition investor responses under certain regimes, events, and sectors. This underscores the need for more nuanced asset pricing models that integrate behavioral signals alongside traditional risk factors in an increasingly climate-aware financial environment.

REFERENCES

Al-Saqaf, W., & Berglez, P. (2019). How do social media users link different types of extreme events to climate change? *Social Media + Society* 5(3). https://doi.org/10.1177/2056305119855147

Austin, P., et al. (2021). Using Google Trends data to nowcast GDP during the COVID-19 pandemic (IMF Working Paper No. 2021/144). *International Monetary Fund* https://www.imf.org/en/Publications/WP/Issues/2021/05/21/Using-Google-Trends-Data-to-Nowcast-GDP-During-the-COVID-19-Pandemic-502399

Bai, J., & Perron, P. (1998). Estimating and testing linear models with multiple structural changes. Econometrica, 66(1), 47–78.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. Journal of Applied Econometrics, 18(1), 1–22.

Barberis, N., Shleifer, A., & Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, 75(2), 283–317.

Battiston, S., Mandel, A., Monasterolo, I., Schütze, F., & Visentin, G. (2017). A climate stress-test of the financial system. *Nature Climate Change* 7(4), 283–288.

Castelnuovo, E., & Tran, T. D. (2017). Google it up! A Google Trends-based uncertainty index for the United States and Australia. *Economics Letters* 161, 149–153.

Cody, E. M., Reagan, A. J., Mitchell, L., Dodds, P. S., & Danforth, C. M. (2015). Climate change sentiment on Twitter: An unsolicited public opinion poll. *PLOS ONE* 10(8), e0136092.

Davino, C., Furno, M., & Vistocco, D. (2013). Quantile Regression: Theory and Applications. Wiley.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics,* 13(3), 253–263.

Elliott, G., Rothenberg, T. J., & Stock, J. H. (1996). Efficient tests for an autoregressive unit root. Econometrica, 64(4), 813–836.

Effrosynidis, D., Papadopoulos, S., & Kompatsiaris, Y. (2022). A systematic review of the nexus between climate change and social media. *Frontiers in Communication* 9, Article 1301400.

European Central Bank. (2021). Climate-related risk and financial stability. *ECB Report*

European Commission. (n.d.). The European Green Deal. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. Journal of Financial Economics, 33(1), 3–56.

French, K. R. (n.d.). Data library: 49 industry portfolios. *Dartmouth College* Retrieved from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Gavriilidis, K. (2021). A new climate policy uncertainty index and the cross-section of stock returns. Finance Research Letters, 43, 101976.

Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from U.S. daily newspapers. Econometrica, 78(1), 35–71.

Gavriilidis, K. (2021). A new climate policy uncertainty index and the cross-section of stock returns. *Finance Research Letters* 43, 101976.

Giglio, S., Kelly, B., & Stroebel, J. (2021). Climate finance. *Review of Financial Studies* 34(3), 796–847.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. American Economic Review, 70(3), 393–408.

Harvey, D., Leybourne, S., & Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of Forecasting,* 13(2), 281–291.

Hassan, T. A., Hollander, S., van Lent, L., & Tahoun, A. (2020). Firm-Level Political Risk: Measurement and Effects. *Quarterly Journal of Economics* 135(4), 2135–2191.

Intergovernmental Panel on Climate Change. (n.d.). Reports and assessments. https://www.ipcc.ch/reports

International Monetary Fund. (2021). Climate-sensitive financial stability assessments: A toolbox to analyze the impact of climate change on financial stability. *IMF Staff Climate Note*

International Monetary Fund. (2024). Climate Transition Risk and Financial Stability in France. *IMF Working Paper No. 2024/144*

Koenker, R. (2005). Quantile Regression. Cambridge University Press.

Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. Econometrica, 46(1), 33–50.

Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *The Review of Financial Studies* 33(3), 1067–1111.

Levin, A., Lin, C.-F., & Chu, C.-S. J. (2002). Unit root tests in panel data: Asymptotic and finite-sample properties. Journal of Econometrics, 108(1), 1–24.

National Aeronautics and Space Administration. (2024). The effects of climate change. https://climate.nasa.gov

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55(3), 703–708.

Newey, W. K., & West, K. D. (1994). Automatic lag selection in covariance matrix estimation. Review of Economic Studies, 61(4), 631–653.

Obstfeld, M. & Rogoff, K. (1996). Foundations of International Macroeconomics. *Cambridge, MA: MIT Press*

Organisation for Economic Co-operation and Development. (2021). Environmental Policy Stringency Index. Retrieved from https://www.oecd.org/economy/greeneco/environmental-policy-stringency.htm

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable Investing in Equilibrium. *Journal of Financial Economics*, 142(2), 550–571.

Romer, D. (2012). Advanced Macroeconomics (4th ed.). *New York: McGraw-Hill*

Roque, I. A. (2024). The impact of climate change awareness index on portfolio returns in the United States in the last 20 years. *Master's thesis, Instituto Superior de Economia e Gestão, Universidade de Lisboa*

Sims, C. A. (1980). Macroeconomics and reality. Econometrica, 48(1), 1–48.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. Journal of Finance, 62(3), 1139–1168.

United Nations Environment Programme Finance Initiative. (2020). Changing Course: A comprehensive investor guide to scenario-based methods for climate risk assessment. Retrieved from https://www.unepfi.org/publications/investment-publications/changing-course/

United Nations Framework Convention on Climate Change. (n.d.). UNFCCC https://unfccc.int

Wooldridge, J. M. (2019). *Introductory Econometrics: A Modern Approach* (6th ed.). Cengage Learning.

APPENDICES

Appendix A: A full list of key words used to create CCAI

ACTIVISM, PROTESTS & PUBLIC ATTENTION	CONSUMER BEHAVIOR & LIFESTYLE	CORE CLIMATE AND ENVIRONMENTAL TERMS
Activism	circular economy	**biodiversity loss**
civil disobedience	climate diet	carbon dioxide
climate anxiety	consumerism	carbon emissions
climate denial	eco-conscious	carbon footprint
climate emergency	eco-friendly	climate change
climate injustice	ethical consumerism	climate sensitivity
climate justice	green living	climate variability
climate march	green tourism	deforestation
climate misinformation	greenwashing	droughts
climate movement	minimalism	environmental degradation
climate protest	organic food	environmental impact
climate strike	plant-based	floods
eco activism	recycling	global warming
eco protests	slow fashion	greenhouse gases
Extinction Rebellion	sustainable agriculture	heatwaves
fighting climate change	sustainable fashion	hurricanes
Fridays for Future	sustainable food	loss of biodiversity
Greta Thunberg	sustainable lifestyle	melting ice caps
youth for climate	sustainable materials	natural disasters
	sustainability	ozone layer
	zero emissions	sea level rise
	zero waste	storm damage
		storms
		temperature anomalies
		warming climate
		wildfires

ECONOMICS, MARKETS & GREEN FINANCE	POLICY, AGREEMENTS & GOVERNANCE	TECHNOLOGY & INNOVATION
carbon credits	carbon neutrality	battery storage
carbon markets	carbon pricing	carbon capture
climate disclosure	carbon tax	clean energy
climate finance	climate adaptation funding	eco innovation
climate related financial risk	climate bill	electronic waste
divestment	climate diplomacy	energy efficiency
financial stability	climate neutrality	energy transition
green bonds	climate pact	green energy
Green economy	climate policy	green hydrogen
green finance	climate regulation	green technology
Green investment	climate summit	low carbon technology
green jobs	climate targets	renewable energy
green recovery	climate-related financial risk	smart grid
green stimulus	European Green Deal	solar power
low carbon economy	Fit for 55	sustainable technology
responsible investing	Green Deal	wind power
stranded assets	IPCC	
sustainable finance	just transition	
sustainable investing	Kyoto Protocol	
transition risk	net-zero emissions	
	Paris Agreement	
	UNFCCC	

Appendix B: First-Differenced Climate Change Awareness Index

This file contains the finalized version of the Climate Change Awareness Index used in all forecasting and modeling throughout the thesis. The index is expressed on a 0–1000 scale and reflects monthly changes in public climate-related attention, computed as the first difference of the weighted index. The Excel file is available for download at the following link:

Climate_Change_Awareness_Index_(weighted_first_difference)

Appendix D: Monthly Excess Returns for 48 Fama-French Industry Portfolios

Excess_Returns_48_Industries_F-F_research_Data_Factors

Excess returns are calculated as raw monthly returns minus the 1-month U.S. Treasury bill rate (risk-free rate).

Appendix E: Threshold Regression Summary for Selected Industries

INDUSTRY	OPTIMAL THRESHOLD (ΔAWARENESS)	OVERALL R^2		
Automobiles	17.6	0.0219	0.0051	0.0136
Construction	13.21	0.0169	-0.0013	0.0105
Chemicals	6.49	0.0014	-0.0043	-0.0047
Steel	17.6	0.0015	-0.002	-0.0288
Mining	3.32	0.0027	-0.0042	-0.0023
Utilities	6.49	0.0029	-0.0018	-0.0058
Electrical Equipment	6.49	0.0019	-0.0039	-0.0046
Food	6.49	0.0022	-0.0023	-0.0064

Retail	6.49	0.0058	0.0	-0.0018
Healthcare	6.49	0.0054	0.0002	-0.0033
Household Products	6.49	0.0022	-0.0026	-0.006

Appendix F: Summary of Quantile Regression Results

This table summarizes the quantile regression coefficients for the awareness terms across five climate-relevant industries. The estimates are reported for the 25th, 50th, and 75th percentiles of the excess return distribution. Full model results are available in Appendix G.

Awareness Terms Across Quantiles

INDUSTRY	QUANTILE	AWARENESS	AWARENESS^2
Cnstr	25th	0.040	0.015
Cnstr	50th	0.186	-0.004
Cnstr	75th	1.779	8.047
Chems	25th	0.171	0.005
Chems	50th	-0.091	-0.059
Chems	75th	1.668	3.876
Steel	25th	0.033	0.049
Steel	50th	0.474	-0.048
Steel	75th	-0.498	4.547
Mines	25th	0.296	0.092
Mines	50th	0.104	0.004
Mines	75th	1.793	9.920
Util	25th	-0.032	-0.013
Util	50th	0.001	-0.030
Util	75th	9.108	19.266

Appendix G: Full Quantile Regression Output (Awareness Across Industries)

Cnstr

VARIABLE	25TH PERCENTILE	50TH PERCENTILE	75TH PERCENTILE
Intercept	-0.608	41.623	590.376
Q('MktRf')	1.330	-0.080	8.174
SMB	0.716	0.031	-20.083
HML	0.289	0.026	1.780
ΔAwareness (ΔCCAI)	0.040	0.186	1.779
Awareness_squared	0.015	-0.004	8.047

Chems

VARIABLE	25TH PERCENTILE	50TH PERCENTILE	75TH PERCENTILE
Intercept	-0.293	60.822	1244.765
Q('MktRf')	1.182	-0.156	3.368
SMB	0.450	0.068	-22.287
HML	0.343	-0.087	4.144
ΔAwareness (ΔCCAI)	0.171	-0.091	1.668
Awareness_squared	0.005	-0.059	3.876

Steel

VARIABLE	25TH PERCENTILE	50TH PERCENTILE	75TH PERCENTILE
Intercept	-0.594	22.090	570.228
Q('MktRf')	1.304	0.182	-2.561

SMB	0.851	0.407	-15.366
HML	0.580	0.219	2.730
ΔAwareness (ΔCCAI)	0.033	0.474	-0.498
Awareness_squared	0.049	-0.048	4.547

Mines

VARIABLE	25TH PERCENTILE	50TH PERCENTILE	75TH PERCENTILE
Intercept	-1.257	10.441	1265.221
Q('MktRf')	1.291	0.825	4.970
SMB	0.572	0.227	-29.209
HML	0.370	0.340	6.920
ΔAwareness (ΔCCAI)	0.296	0.104	1.793
Awareness_squared	0.092	0.004	9.920

Util

VARIABLE	25TH PERCENTILE	50TH PERCENTILE	75TH PERCENTILE
Intercept	0.800	72.866	2250.269
Q('MktRf')	0.561	-0.041	20.468
SMB	-0.134	0.001	-45.446
HML	0.036	-0.038	-14.121
ΔAwareness (ΔCCAI)	-0.032	0.001	9.108
Awareness_squared	-0.013	-0.030	19.266

Appendix H: Quantile Regression with Std. Errors and P-values

Quantile_Regression_All_8_Industries.xlsx

Each sheet represents one of the industries.

Appendix I: The CAR Results of All Events

CAR_Results_All_Events.xlsx