

Climate Change: Factor Investing and Portfolio Management

—Evidence from AI-Classified Sentiments

Yehuda Davis¹, Henry He Huang^{1*}, Haiting Sun²

¹Sy Syms School of Business, Yeshiva University, New York, NY, USA

²Ningbo Shufa Private Equity Fund Management Co., Ltd., Ningbo, China

Email: yehuda.davis@yu.edu, *henry.huang@yu.edu, haiting@nyu.edu

How to cite this paper: Davis, Y., Huang, H., & Sun, H. (2026). Climate Change: Factor Investing and Portfolio Management. *Journal of Financial Risk Management*, 15, 335-347.

<https://doi.org/10.4236/jfrm.2026.154019>

Received: September 13, 2026

Accepted: October 13, 2026

Published: October 16, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper examines whether AI-classified climate-change sentiment extracted from financial news contains incremental information for return prediction and portfolio optimization. We construct a novel climate-news sentiment measure using a large language model (LLM) to classify daily climate-related sentiment in Bloomberg and Reuters news articles. We then incorporate this measure into a rolling-window Fama-French five-factor forecasting model to generate predicted returns for a portfolio of ten exchange-traded funds (ETFs) spanning equities, fixed income, commodities, and foreign exchange. Using these predicted returns, we form daily portfolio weights that maximize the Sharpe ratio and compare their performance to both a portfolio constructed without the climate sentiment measure and to an equal-weight baseline portfolio. Over our sample period, the portfolio that incorporates climate-news sentiment achieves substantially higher returns and stronger risk-adjusted performance across multiple metrics. Our results indicate that climate-related news contains incremental information useful for ETF return forecasts and dynamic portfolio allocation. Additionally, we illustrate the ability of LLMs to extract economically meaningful information from financial news articles. Our findings contribute to the literature on climate-related information in financial markets and demonstrate the potential of AI-driven sentiment measures in portfolio management.

Keywords

Climate Risk, ChatGPT, AI Models, Portfolio Optimization

1. Introduction and Literature Review

Climate change has become a prominent factor in investment portfolio decisions.

De Angelis et al. (2023) illustrate how climate change can influence investor decisions and, consequently, investment portfolios. Bioy et al. (2022) examined 860 global mutual funds and exchange-traded funds (ETFs) with a climate mandate, highlighting that the returns of these ETFs are impacted by climate variables. Furthermore, many studies (e.g., Engle, 2014; Engle et al., 2020; Kolle et al., 2022) have focused on refining climate-change measures and their influence on investment portfolio decisions. To advance the integration of climate-related information into investment strategies, we select several ETF tickers across various asset classes, including equities, fixed income products, commodities, and foreign exchange. Prior studies on climate risk in financial markets have focused primarily on equities. Our diverse choice of funds extends the existing literature to non-equity asset funds and lends robustness to our conclusions.

Engle et al. (2020) develop a climate news factor using a mimicking portfolio approach and incorporate it into a multi-factor regression framework to measure stock exposure to climate risk. They use their model to construct a tradable portfolio to hedge against climate risk. To complement their findings, our study evaluates whether the climate risk factor itself is priced as an additional factor within an asset pricing framework. Complementing that work, our study asks a different question: whether AI-classified climate-news sentiment improves time-series return forecasts and resulting portfolio allocations across a broad set of ETFs.

Following the framework of Merton (1987), which addresses the imperfections in financial information, we explore the potential to enhance investment returns by incorporating information about climate change. We posit that climate-related news may contain information about ETF returns that is not fully captured by standard forecasting variables. This builds on the traditional factor investing framework outlined by Fama and French (2015), which we augment with a climate-news sentiment measure in a return-forecasting setting. Our goal is to evaluate whether climate-news sentiment provides incremental predictive information for portfolio construction.

In order to generate our climate-news sentiment measure, we utilize a large language model (LLM), specifically ChatGPT, to analyze the sentiment of daily news data from Bloomberg and Reuters. Using an LLM allows us to incorporate all the news from every trading day—a task that would be very costly and time-consuming to do manually. Recent literature (e.g., Bojić et al., 2025) has demonstrated that LLMs perform sentiment analysis on par with humans, providing strong support for the accuracy of our climate-news sentiment measure.

In addition, our use of an LLM is an important methodological advance. Compared with dictionary-based or supervised natural language processing methods, LLMs capture richer semantic context, leverage more complex reasoning, and approximate human subjective judgment more closely. This is a critical feature because sentiment itself influences investment decisions. LLMs are also easily applicable across languages and data sources, and with continued model improvements, their sentiment classification will continually improve, making this a practical ap-

proach for both current and future research.

Our LLM-based climate-news measure is then integrated into a rolling regression model that we use to calculate daily portfolio weights that maximize the Sharpe ratio. In order to determine the effectiveness of incorporating the climate-news sentiment, we compare the sequence of portfolios generated by our model to portfolios based on the Fama-French model without the climate-news measure and to an equal-weighted baseline portfolio across several performance metrics. Specifically, our climate-informed portfolios outperform the other portfolios in terms of cumulative return, the Sharpe ratio, Sortino ratio, and Calmar ratio, indicating that climate-news sentiment adds useful information for forecasting returns and allocating across diverse asset classes. These performance gains are achieved with the magnitude of maximum drawdown remaining broadly unchanged relative to the model without climate sentiment.

The remainder of the paper is organized as follows: Section 2 introduces the data and provides descriptive statistics and details the use of the LLM for sentiment analysis. Section 3 outlines the construction of our models and their results, and Section 4 concludes.

2. Data and Descriptive Statistics

For our portfolio analysis, we utilize ten ETF tickers across various asset classes, including equities, fixed income products, commodities, and foreign exchange. Specifically, we chose: SPY, IWM, EEM, TLT, LQD, TIP, IYR, GLD, OIH, and FXE. The details of each ETF can be found in **Appendix A**. For each ETF, we download the daily adjusted closing price from Yahoo Finance using the `yfinance` package in Python and calculate daily returns. The cumulative returns for each ETF during our sample can be found in **Figure 1**. The daily Fama-French five factors come from the Fama-French website (including Mkt-RF, SMB, HML, RMW, CMA, and RF)¹.

In order to determine the daily climate-change sentiment, we use a dataset of financial news articles published by Bloomberg and Reuters (Remy & Ding, 2015)². This dataset contains articles from October 20, 2006 through November 26, 2013, so all our analyses are focused on this 7-year period. For each day in our sample, we aggregated all news articles into a single large string joined by spaces. Due to ChatGPT API limitations, this string was truncated at 1 million characters when necessary. Using the OpenAI GPT-5 nano model, we submitted the following prompt (similar to Bojić et al., 2025), consisting of a system message and a user message³:

System message:

You are a single-label classifier for climate-change content. Output EXACTLY

¹https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research

²The news dataset used in this paper was originally obtained from this GitHub repository: <https://github.com/philipperemy/financial-news-dataset>, which was publicly available at the time of download. However, the dataset has since been removed from public access due to copyright concerns.

³For reproducibility, please see further details in **Appendix B**.

one digit from {1,2,3} and nothing else.

Labels:

1 = negative language about climate (e.g., texts conveying casualty, storm, floods, heatwave, wildfire, rising temperature, warming, ice melting, greenhouse gas, harms/risks/losses/setbacks/cost increases/failures/criticism/worsening/dangers/threats, etc.).

2 = Neutral/Off-topic: NOT primarily about climate; OR on-topic but purely factual/balanced/mixed with no clear valence.

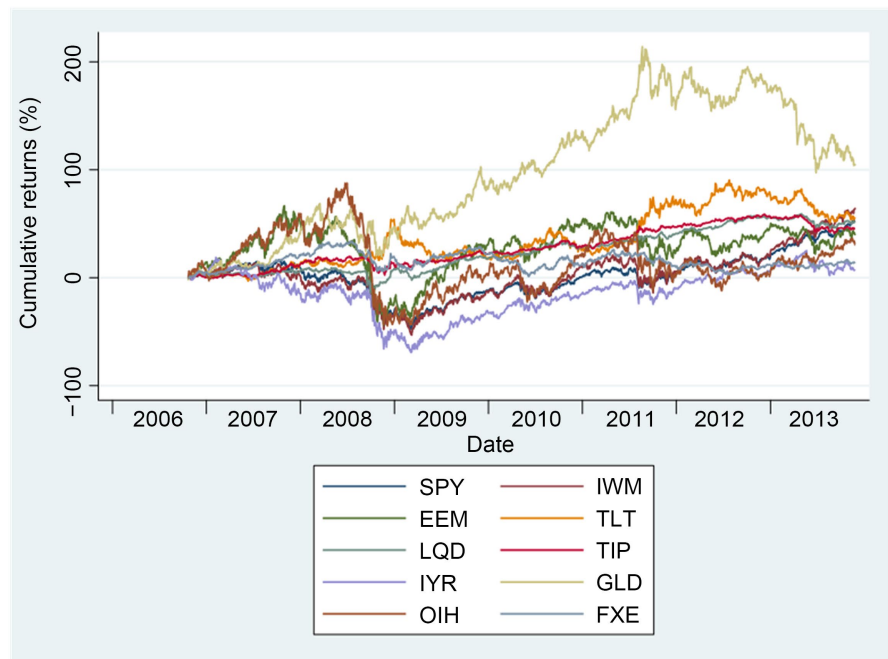
3 = positive language about climate (e.g., texts conveying benefits/progress/success/wins/funding/approvals/cost decreases/improving/opportunities/solutions/success stories/policy wins, etc.).

Guidance: Examples are illustrative only, NOT a keyword checklist.

User message:

Classify this text:

{string of news articles here}



The graph shows the cumulative returns of the ten ETFs in our sample for the period from October 23, 2006 through November 26, 2013. The returns are calculated on a daily basis and expressed as a percentage.

Figure 1. Cumulative ETF returns.

Any day without news was recorded as off-topic. A summary of the daily climate-change sentiment, including a disaggregation of neutral climate-related news and no climate-related news, is presented in **Table 1(A)**. Our complete sample consists of 1788 trading days.

Since it is possible that we are capturing arbitrary classifier noise rather than climate sentiment, we manually validated the sentiment classifications. Specifically,

we randomly selected three months from the sample period (August 2012, February 2012, and January 2011), retaining all 63 trading observations in those months. One author manually coded each day's news sentiment as negative, neutral, or positive without consulting the LLM labels, with the results reported in **Table 1(B)**. The resulting agreement rate between our manual labels and the LLM labels is approximately 89%, macro-F1 is 0.810, and unweighted Cohen's kappa⁴ is 0.669. The kappa falls in the conventional substantial-agreement range (Landis & Koch, 1977), and macro-F1 gives equal weight to the three classes. An always-neutral classifier achieves only 0.285 macro-F1 on this sample. Overall, these results indicate that the classifier provides a reasonably reliable sentiment signal in our setting.

Table 1. (A): Summary of the daily climate-change sentiment; (B): sentiment validation.

(A)		
	Frequency	Percent
Sentiment		
Negative	103	5.76
Climate-related neutral/mixed	758	42.39
Off-topic, including no matched news	863	48.27
Positive	64	3.58
Total	1788	100.00

This table presents the frequency distribution of daily climate-change sentiment labels in the dataset for the period from October 23, 2006 through November 27, 2013. Sentiment is categorized as Positive, Negative, or Neutral, where Neutral includes both neutral climate-related news and no climate-related news. The classification is based on the OpenAI GPT-5-nano model's interpretation of articles published by Bloomberg and Reuters.

(B)				
	LLM	Negative	Neutral/Off-Topic	Positive
Human				
Negative	7	6	0	
Neutral/Off-Topic	0	47	0	
Positive	0	1	2	

This table presents the results of our sentiment validation. The rows represent the human classifications, and the columns represent the LLM classifications.

3. Methodology and Results

In order to assess the ability of the climate-change sentiment to enhance return prediction, we use the following OLS regression model to generate rolling window predictions for each ETF, as described below:

$$r_{i,t} = \alpha_i + \beta_{1,i} NegCC_{t-1} + \beta_{2,i} PosCC_{t-1} + \sum_{j=1}^6 \gamma_{i,j} FF_{j,t-1} + \sum_{j=1}^9 \delta_{i,j} r_{j,t-1} + \epsilon_{i,t} \quad (1)$$

⁴With one annotator, kappa measures human-LLM agreement, not agreement between multiple human coders.

where $r_{i,t}$ is the return of ETF i at time t , and $NegCC_{t-1}$ ($PosCC_{t-1}$) is a binary variable that takes a value of 1 if the lagged climate-change sentiment is negative (positive), and 0 otherwise. $FF_{j,t-1}$ is the lagged j th Fama-French factor from the Fama-French 5-factor model (including the risk-free rate), and for each ETF i , $r_{j,t-1}$ represents the lagged return of each other ETF j . All the explanatory variables are lagged, ensuring that we can generate predicted returns without any lookahead issues. Specifically, for each return date t , the other ETF returns, the Fama-French factors, the risk-free rate, and the news sentiment are all from return date $t-1$. Any news published on date t , even prior to markets opening, is not used as a predictor for date t returns.

Starting from January 25, 2007 (in order to have a pre-window of sufficient length), we construct a rolling-window framework. For each day t , we estimate the regression coefficients using data from day $t-64$ to day $t-2$ ⁵. The estimated coefficients are then used to generate a predicted return for day t . In addition, for comparison, we perform the same procedure without the climate-change sentiment variables.

We now use both sets of predicted returns to construct optimized daily portfolios of the ten ETFs using the SciPy optimization module in Python. Specifically, for each date in our sample, using the predicted returns and the covariance matrix for the prior 63 trading days, we find the vector of weights that maximizes the Sharpe ratio under a long-only constraint (standard in asset management).⁶ For the purpose of comparison, we also calculate a baseline portfolio using equal weights.

To separate the informational value of climate-news sentiment from portfolio-engineering effects, the two optimized strategies (with and without climate sentiment) are implemented with identical portfolio construction rules: the same objective function, long-only constraint set, covariance lookback window, and daily rebalancing schedule. The only difference between these two strategies is whether the climate-sentiment regressors are included in the return-forecasting step.

In order to assess the model's performance, we use the calculated daily weights to calculate portfolio returns and various metrics using i) the baseline (equal-weight) portfolio, ii) the weights from the Fama-French model without the climate-change sentiment measure, and iii) the weights from our model including

⁵This 63-day window corresponds to approximately three trading months, which is commonly used as a training window in investment practice. At the end of this section, we perform a sensitivity analysis to ensure that our results are robust to our training window specification.

⁶Specifically, each weight takes a value between 0 and 1, and the weights all sum to 1. The maximum-Sharpe objective uses forecast returns and the covariance of the preceding 63 daily ETF returns, excluding the forecast date. The risk-free benchmark in this objective is zero; RF separately enters the forecasting regression. Each optimization starts from equal weights of 0.1 and uses `scipy.optimize.minimize` with the equality constraint and `Bounds(0, 1)`. The resulting default solver is SLSQP, with numerical derivatives, `ftol = 1e-6` and `maxiter = 100`. All 3450, 3408 and 3386 optimization calls in the 63-, 84-, and 95-day replications reported successful convergence, respectively, with weights within bounds. The largest weight-sum deviation across these runs was $1.51\text{e-}6$. Normalization leaves the reported return and Sharpe figures unchanged.

the climate-change sentiment measure. The results are reported in **Table 2(A)**. The annualized (cumulative) return for the baseline portfolio over our sample period is about 7.1% (59.7%). The Fama-French model yields an annual (cumulative) return of almost 13.7% (140.6%). Our model incorporating climate-change sentiment surpasses that annual return rate by over 5 percentage points, with an annualized (cumulative) return of about 18.8% (224.6%). A graphical representation of the cumulative returns of each portfolio over our sample period can be found in **Figure 2**. Over almost the entire sample period, the portfolio incorporating climate-change sentiment outperforms the other two portfolios, consistent with the view that the sentiment measure improves the model's forecasting and allocation decisions.

Continuing our portfolio analysis, we look at the Sharpe Ratio for each portfolio. The baseline portfolio has a Sharpe Ratio of 0.51, the Fama-French model yields 0.96, and our climate-change model is a stronger 1.25. The largest peak-to-trough loss (Max Drawdown) is -0.36 , -0.23 , and -0.23 for the baseline, Fama-French, and climate-change model, respectively. For the Sortino Ratio, which is roughly the same as the Sharpe Ratio with a focus only on downside risk, the climate-change model (1.91) once again outperforms the Fama-French model (1.42) and the baseline model (0.73). Lastly, looking at the Calmar Ratio, which is calculated as annualized return divided by the max drawdown, the baseline comes in at 0.19, the Fama-French at 0.59, and our climate-change model at a higher 0.80.

Because **Table 2(A)** compares complete portfolio strategies, we conduct paired inference on the 1725 archived daily portfolio returns. We use 5000 circular moving-block bootstrap replications with the same resampled dates for both strategies, a primary block length of 10 days, and 20-day and 63-day sensitivity checks. The results are reported in **Table 2(B)**. The table reports the differences in the annual return between the model with the climate-change sentiment and the model without the climate-change sentiment, equal-tailed 95% intervals, and matching two-sided, centered bootstrap p -values. The portfolio improvements remain statistically significant at 5% across all three block lengths, including Holm adjustment for the two metrics within each block length. These results provide direct inference for the return and Sharpe improvements reported in **Table 2(A)**.

In order to ensure that our results are robust to different training window specifications, we retest our model using 84-day and 95-day OLS windows in addition to our original 63-day OLS window, and an 84-day Ridge specification with $\alpha = 1$. The results are reported in **Table 3**. For OLS, only the lookback changes, affecting both regression and covariance estimation. Ridge standardizes predictors within each training window and leaves the intercept unpenalized. All specifications use the same 1693 evaluation days: March 13, 2007 to November 27, 2013⁷. The climate-informed strategy retains higher return and Sharpe ratio across all

⁷Note that the start date of this analysis is later than our primary analysis due to the additional trading days required for the longer windows, which explains the small differences between the results for the 63-day OLS window in this table and our main results.

reported specifications. Mean within-ETF coefficient standard deviations for negative/positive sentiment decline from 0.00712/0.00937 (63 days) to 0.00665/0.00875 (95 days), indicating lower coefficient variation with the longer window.

Table 2. (A): Portfolio performance metrics; (b): Portfolio performance—paired inference.

(A)						
	Annualized Return (%)	Cumulative Return (%)	Sharpe Ratio	Max Drawdown	Sortino Ratio	Calmar Ratio
Baseline	7.1	59.7	0.51	−0.36	0.73	0.19
FF without CC factor	13.7	140.6	0.96	−0.23	1.42	0.59
FF with CC factor	18.8	224.6	1.25	−0.23	1.91	0.80

This table provides the portfolio performance metrics for portfolios consisting of a weighted combination of the ten ETFs in our analysis for the period from January 25, 2007 through November 27, 2013. The “Baseline” model uses equal weights. The “FF without CC factor” model uses a rolling Fama-French model to calculate daily weights that maximize the Sharpe Ratio. The “FF with CC factor” model does the same with the addition of positive and negative climate-change sentiment indicator variables to the model. The metrics are defined in **Appendix C**.

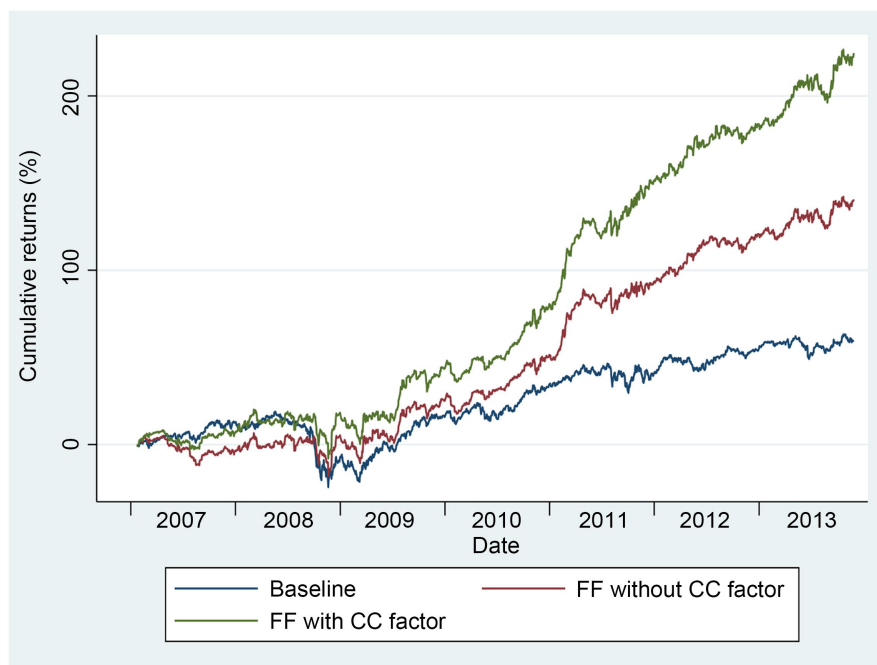
(B)				
Block	Metric	With CC Minus Without CC	95% Interval	<i>p</i> -Value
10	Annual return	5.082	[1.318, 8.772]	0.0116
10	Sharpe	0.296	[0.063, 0.514]	0.0152
20	Annual return	5.082	[1.197, 8.711]	0.0116
20	Sharpe	0.296	[0.056, 0.514]	0.0208
63	Annual return	5.082	[1.334, 8.440]	0.0072
63	Sharpe	0.296	[0.048, 0.494]	0.0192

This table provides the results of paired inferences on the 1725 archived daily portfolio returns. We use 5000 circular moving-block bootstrap replications with the same resampled dates for both strategies, a primary block length of 10 days, and 20-day and 63-day sensitivity checks. “With CC minus Without CC” refers to annual-return differences and is reported in percentage points. The 95% intervals are equal-tailed with matching two-sided, centered bootstrap *p*-values. The metrics are defined in **Appendix C**.

Table 3. Sensitivity analysis.

Specification	Return without CC (%)	Return with CC (%)	Sharpe without CC	Sharpe with CC
OLS 63	13.61	18.19	0.948	1.214
OLS 84	11.67	14.96	0.759	0.934
OLS 95	10.25	13.08	0.702	0.858
Ridge 84	14.81	17.27	0.933	1.062

This table provides a sensitivity analysis of the training window specification in our primary analysis. OLS 63 is our primary model using a 63-day rolling OLS regression for the training window. OLS 84 and OLS 95 are the same, using an 84-day and 95-day window, respectively. Ridge 84 refers to an 84-day Ridge specification with $\alpha = 1$. Ridge standardizes predictors within each training window and leaves the intercept unpenalized. All specifications use the same 1693 evaluation days: March 13, 2007 to November 27, 2013. The metrics are defined in **Appendix C**.



This graph shows the cumulative returns (%) of our three portfolios consisting of a weighted combination of the ten ETFs in our analysis for the period from January 25, 2007 through November 27, 2013. The “Baseline” model uses equal weights. The “FF without CC factor” model uses a rolling Fama-French model to calculate daily weights that maximize the Sharpe Ratio. The “FF with CC factor” model does the same with the addition of positive and negative climate-change sentiment indicator variables to the model.

Figure 2. Cumulative portfolio returns.

4. Conclusion

This study contributes to the growing literature on climate risk and financial markets by introducing an AI-classified measure of climate sentiment and demonstrating its value in return forecasting and portfolio optimization. Using a large language model (LLM) to compute climate-related sentiment from daily financial news and incorporating it into the Fama-French five-factor model, we show that climate-change sentiment contains incremental information that can improve portfolio optimization.

A portfolio of ten diverse ETFs weighted using the model with our climate-change sentiment substantially outperforms both a comparable model that excludes the climate sentiment and an equal-weight baseline portfolio. Over the sample period, our proposed strategy achieves a higher annualized return and improved risk-adjusted performance, as measured by the Sharpe, Sortino, and Calmar ratios. These results suggest that climate-related news is not fully captured by the baseline forecasting variables and can be usefully incorporated into dynamic portfolio allocation.

Our findings highlight the potential of LLMs to extract information from financial news and underscore the usefulness of climate-news sentiment for portfolio decisions across asset classes. We also validate our sentiment classification by man-

ually coding random subsamples, which supports the reliability of the sentiment measure. As climate risks continue to intensify, understanding how climate-related information affects return expectations and investment decisions remains an important area for both researchers and practitioners.

Author Contributions

All the authors make the same contribution to this article.

Data Availability

Most of the data used in this study is available from public sources identified in the study. The news dataset used in this study was originally sourced from <https://github.com/philipperemy/financial-news-dataset>, which was publicly available at the time of data collection. However, it has since been removed from public access due to copyright restrictions. As such, we are unable to redistribute the dataset. Researchers interested in replicating our work may contact the original dataset author or Bloomberg and Reuters directly.

Disclaimer

The views and opinions expressed are those of the authors and do not necessarily reflect the views of Ningbo Shufa Private Equity Fund Management Co., Ltd., its affiliates, or its employees; do not constitute an offer, solicitation of an offer, or any advice or recommendation to purchase any securities or other financial instruments, and may not be construed as such.

Acknowledgements

The authors gratefully acknowledge financial support from the Sy Syms School of Business, Yeshiva University.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- Bioy, H., Wang, B., Stankiewicz, A., & Alladi, A. (2022). *Investing in Times of Climate Change 2022*. Morningstar. https://assets.contentstack.io/v3/assets/blt4eb669caa7dc65b2/blt2317a85c41ab1992/6255cf57fea57c4dece51c53/Investing_in_Times_of_Climate_Change_2022_Final.pdf
- Bojić, L., Zagovora, O., Zelenkauskaitė, A., Vuković, V., Čabarkapa, M., Veseljević Jerković, S. et al. (2025). Comparing Large Language Models and Human Annotators in Latent Content Analysis of Sentiment, Political Leaning, Emotional Intensity and Sarcasm. *Scientific Reports*, 15, Article No. 11477. <https://doi.org/10.1038/s41598-025-96508-3>
- De Angelis, T., Tankov, P., & Zerbib, O. D. (2023). Climate Impact Investing. *Management Science*, 69, 7669–7692. <https://doi.org/10.1287/mnsc.2022.4472>
- Engle, R. F. (2014). *A Financial Approach to Environmental Risk*. Resources.

<https://www.resources.org/archives/a-financial-approach-to-environmental-risk/>

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

<https://doi.org/10.1093/rfs/hhz072>

Fama, E. F., & French, K. R. (2015). A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, 116, 1-22. <https://doi.org/10.1016/j.jfineco.2014.10.010>

Kolle, J., Lohre, H., Radatz, E., & Rother, C. (2022). *Factor Investing in Paris: Managing Climate Change Risk in Portfolio Construction*. *Journal of Investment Management*, Forthcoming. <https://papers.ssrn.com/abstract=4071422>

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33, 159-174. <https://doi.org/10.2307/2529310>

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, 42, 483-510.

<https://doi.org/10.1111/j.1540-6261.1987.tb04565.x>

Remy, P., & Ding, X. (2015). *Financial News Dataset from Bloomberg and Reuters*. GitHub.

<https://github.com/philipperemy/financial-news-dataset>

Appendix A: ETF Descriptions

ETF Ticker	Name	Description
SPY	SPDR S&P 500	U.S. Equities Large Cap
IWM	iShares Russell 2000	U.S. Equities Small Cap
EEM	iShares MSCI Emerging Markets	Global Emerging Markets Equities
TLT	iShares 20+ Years	U.S. Treasury Bonds
LQD	iShares iBoxx \$ Invst Grade Corp Bond	U.S. Liquid Investment-Grade Corporate Bonds
TIP	iShares TIPS Bond	U.S. Treasury Inflation-Protected Securities
IYR	iShares U.S. Real Estate	Real Estate
GLD	SPDR Gold Shares	Gold
OIH	VanEck Vectors Oil Services ETF	Oil
FXE	CurrencyShares Euro ETF	Euro

Appendix B: LLM Classification Protocol

The original implementation requested model = “gpt-5-nano” through the Chat Completions API. Temperature, other sampling parameters, and the output-to-token limit were not explicitly set, so API defaults applied. The original returned snapshot was not saved. For the supplementary neutral/off-topic classification, the API response identified the deployed snapshot as gpt-5-nano-2025-08-07.

The classifier preserved the article order embedded in each archived daily string, with no additional sorting. Missing input was converted to an empty string; other input was converted to text without further cleaning. To manage input length under the model/API context constraints, the code retained the first 1,000,000 characters using `s[:1_000_000]`, without adjusting to article boundaries. This was a character-based implementation cap, not an exact token count. Truncation affected 784 of 1944 daily records (40.33%), including 693 of the 1734 records in the supplementary split (39.97%).

We first classified sentiment in the daily news collections. For the supplementary split, the same model classified only the 1734 original neutral/off-topic records as off-topic (0) or climate-related neutral/mixed (2), retaining the original positive and negative labels. Only afterwards did we apply the original news/trading-calendar padding, one-row lag, and trading-date alignment to obtain the final 1788 observations.

Appendix C: Variable/Metric Definitions

Variable	Definition
r_{it}	Return of asset or ETF i at time t
$NegCC_{t-1}$	Binary variable that takes a value of 1 if the lagged climate-change sentiment derived from LLM analysis of news is negative, and 0 otherwise

Continued

$PosCC_{t-1}$	Binary variable that takes a value of 1 if the lagged climate-change sentiment derived from LLM analysis of news is positive, and 0 otherwise
$FF_{j,t-1}$	Lagged j th Fama-French factor from the Fama-French 5-factor model
Metric	Definition
Sharpe Ratio	Risk-adjusted return: (Portfolio Return – Risk-Free Rate)/Volatility
Max Drawdown	Largest peak-to-trough loss in portfolio value
Sortino Ratio	Downside risk-adjusted return: focuses on downside deviation
Calmar Ratio	Annualized return divided by maximum drawdown