From Climate Talk to Stock Shock: Modeling Policy Headline Sentiment and Index Fund Performance

Aditya Jain, Allison Lin, and Valerie Rogel
Columbia University
August 2023

Author Note

aj3128@columbia.edu (Jain); yl5212@columbia.edu (Lin);
vr2509@columbia.edu (Rogel)

Abstract

Significant shifts have occurred in U.S. domestic climate policy since the ratification of the 2015 Paris Climate Accords. The efficient market hypothesis expresses that market prices are driven by investor response to major shifts in news events and stories. Despite the magnitude of these climate policy shifts, there is a lack of research on investor reactions to such policy changes. In our study, we address this research gap by attempting to determine and model the relationship between the sentiment of climate policy news headlines and the financial metrics of major U.S. index funds. Initially, we employ the Valence Aware Dictionary and sEntiment Reasoner (VADER) — a natural language processing (NLP) technique — to derive sentiment scores from 14,000 relevant climate policy headlines spanning January 2014 to June 2023. Following this, we apply Granger causality tests and compare the effectiveness of univariate autoregressive forecasting models to bivariate machine learning regression models. The latter incorporates climate policy news sentiment as a predictive feature, allowing us to assess the value of this sentiment in forecasting index fund metrics. Our findings indicate that consumer discretionary, consumer staples, and information technology funds are most effectively predicted using climate policy news sentiment. Through this study, we aim to illuminate the relationship between climate policy news sentiment and shifts in investor behavior.

From Climate Talk to Stock Shock: Modeling Policy Headline Sentiment and Index Fund Performance

Introduction

Since 2015, the year the United States adopted the Paris Agreement, climate change has increasingly influenced institutional and personal investment decision making. A growing strand of research into finance and investor behavior focuses on how investors respond to climate risks, particularly in extreme weather scenarios, by aiming to build portfolios less vulnerable to climate impacts. While the paper "Investor Rewards to Climate Responsibility" (Ramelli et al., 2021) analyzes stock price responses to discrete and significant U.S. political shifts - such as Donald Trump's election in 2016, Scott Pruitt's appointment to the Environmental Protection Agency (EPA) in 2017, and Joseph Biden's election in 2020 - there have been limited efforts to conduct research focused on these responses over a larger period of time.

Since 2014, there have been numerous policies implemented with the goal of impacting emissions levels in different ways, including the Inflation Reduction Act (2022), executive action on the Keystone XL Pipeline (in 2015, 2017, and 2021), and on the Paris Climate Accords (in 2016, 2017, and 2021). The entire legislative lifetime of these initiatives generates numerous headlines as they are debated, tabled, negotiated, and criticized. As the United States continues to develop new policies aligned with its emissions reductions goals, further investigation is necessary to determine how investors respond to news about policy changes in a rapidly evolving policy landscape.

Our objective is to model how investors respond to news about changes in climate policy by executing a Natural Language Processing (NLP) sentiment analysis of relevant climate policy headlines. We aim to build a time series of climate policy sentiment, and analyze its trends in tandem with trends in a general collection of stocks (such as the S&P 500) and industry-specific indexes to develop a model of investor behavior in response to climate policy events.

Literature Review

Our assumption that investor decisions are often driven by responses to major shifts and news events is motivated by Ardia et al., 2020 in their paper "Climate Change Concerns and the Performance of Green vs. Brown Stocks," who test a theoretical framework (initially proposed by Pástor et al., 2022) linking unexpected shifts in sustainability preferences to asset price changes. Engle et al., 2020, in their paper "Hedging Climate News," further explain that investor behavior is centered around hedging for climate risk, even in spite of short-run costs. They posit that "events that plausibly contain ... information about changes in climate risk" invite newspaper attention ultimately influential to investor sentiment, drawing their eventual sentiment analysis from the Wall Street Journal (WSJ) and other media outlets consumed by market participants.

To measure the extent of climate change discussions in the news media, Engle et al., 2020 utilize a two-pronged approach to measure climate change discussion in the news media. First, they utilize word-cloud and term frequency-inverse document frequency (tf-idf) techniques to observe the correlation between monthly Wall Street Journal (WSJ) text content and a fixed climate change vocabulary derived from United Nations and U.S. Environmental Protection Agency (EPA) sources. Next, they utilize sentiment analysis to gauge the intensity of negative sentiment in a given month. By examining firms' Environmental, Social, Governance (ESG) metrics, their paper was able to provide insight an effective strategy to hedge against climate risk. Similarly, our paper will utilize NLP techniques to shed light on the relationship between climate change news and investor response.

Ardia et al., 2020 utilized daily published news articles as a novel proxy for unexpected changes in climate concerns, constructing a daily 'Media Climate Change Concerns' (MCCC) index. They found a positive relationship between increases in unexpected media climate change concerns and the performance of a 'green-minus-brown' portfolio; this portfolio was classified based on companies' ASSET4/Refinitiv greenhouse gas (GHG) emissions data. Their findings suggest green firms (in which GHG emissions necessary to generate one million dollars in revenue falls below the 25th percentile among the S&P 500) stocks will outperform brown firms (in which GHG emissions to generate one million dollars in revenue are above the 75th percentile) stocks when climate change concerns unexpectedly increase.

Bhavsar et al., 2022 emphasize the efficient market hypothesis, which expresses that market prices do not rely solely on historical performance, but instead respond to new information such as news stories. Using the Valence Aware Dictionary and Sentiment Reasoner (VADER), a lexicon and rule-based sentiment analysis tool developed by Hutto and Gilbert, 2014 in the Natural Language Toolkit (NLTK) module of Python, they found that there was a weak correlation between sentiment and stocks' closing price; this may be due to the fact that the collected daily news headlines were not differentiated by topic, such as climate disaster or business. In our methodology, we might also make use of VADER as an accepted sentiment detector.

Vicari and Gaspari, 2021 investigated forecasting market sentiment based on news headlines using long short-term memory, an improved version of a recurrent neural network. They trained their model by preprocessing news headlines from an eight-year Kaggle data set through removing stopwords, numbers, and special elements, and vectorizing only the 3000 most frequent words through a tokenizer. They found that trading strategies based on sentiment may be particularly prone to the "herd behavior" assuming that these algorithms must train on data sets containing lots of news that may turn out to be extremely similar. The researchers also note that eliminating the assumption that positive financial sentiment is the same as "positive words" in a headline reduces the accuracy of a deep learning model. In our methodology, we maintain this assumption because few language models already exist that can effectively differentiate between positive financial or climate statements and positive words within headlines.

Data Collection and Processing

Financial and Climate Policy Data Collection

In our study, we collect both financial and headline data. We accessed the closing price and a ten-day moving volatility of eleven different industry-based index funds as well as the S&P 500 from Vanguard as shown in Table 1 from January 1, 2014 until May 28, 2023 using Bloomberg Terminal. We selected industry index funds because there is only one closing price per trading day on which they are bought and sold.

Table 1
Selected index funds' ticker symbols and their corresponding industry

Index Fund Ticker Symbol	Industry
VSPVX	S&P 500
VUIAX	Utilities
VGHCX	Healthcare
VFAIX	Financials
VITAX	Technology
VGSLX	Real Estate
VINAX	Industrials
VTCAX	Communications
VCSAX	Consumer Staples
VENAX	Energy
VCDAX	Consumer Discretionary

We also created a list of keywords which pertain to climate policy using three categories of terms: the actual word “climate,” a policy-related term (eg. "deal" or "agreement"), and a climate-related term (eg. "coal" or "fossil fuels") (8). Using logical operators so that all three categories are included in the headline, we obtained news articles from LexisNexis for each month spanning from January 1, 2014 to June 16, 2023. LexisNexis is an archive site which enables users to download the headline, summary, date published, and source of a news article. The headlines are from The New York Times, The Financial Times Online, the Associated Press, and CNN.com, which have relatively high volume of circulation. We assume that the headlines are representative of all climate policy news for each particular day even though their explicit perspectives may differ. However this may introduce some bias as investors may get their financial news from less traditional sources such as social media, or more politically radical sources.

Climate Policy Headline Relevancy Processing

Upon reviewing the LexisNexis data, we observed that out of the 78,000 headlines gathered, not all were pertinent to climate policy, a limitation stemming from our search terms. To address this, we manually-labeled a subset of these headlines (n = 1899) to determine their relevance to climate policy. Using 89% of the expert-labeled data (n = 1688) as our training set and the remaining 11% (n = 211) as our test set, we deployed a binary classification model through the AutoTrain platform on Hugging Face.

The model achieved an accuracy of 90.54% and a precision of 68.42%. For binary classifiers, the quality is often represented by the area under the curve (AUC). Our model's AUC was .9317. Typically, an AUC value exceeding .9 is indicative of a high-quality binary classification model.

Applying the model on the set of all 78,000 headlines yielded our final dataset of 14,334 relevant headlines.

Climate Policy Headline Sentiment Analysis

In order to ascribe a sentiment to the headlines, we used VADER (Valence Aware Dictionary and sEntiment Reasoner), which is typically used for determining sentiment in social media. We concatenated each day's headline in Python in order to create a longer string which would be more optimal for determining sentiment, since VADER was initially designed for Tweets, a longer string of text than a single headline. On average we had 14334 headlines divided by 2936 days for an average of 4.88 headlines per day. The resulting Strings were of varying lengths, but still generally comparable to the Twitter limit of 280 characters per Tweet for most of the time period from 2014-2023. VADER creates three scores which represent the proportion of text that VADER classifies as positive, negative, and neutral. VADER also returns a compound score between -1 and 1, where -1 is the most negative and 1 is the most positive. We used the compound score to represent daily sentiment.

Table 2

Jan 2014—Sentiment Analysis on Concatenated Headlines

Date	Concatenated Headline(s)	VADER Score
1/3/2014	Kerry Making Pact on Climate A Top Priority	0.2023
1/7/2014	Big four EU economies seek tougher cuts in gas emissions	-0.128
1/8/2014	EU makes carbon pollution more expensive	0
1/10/2014	Scientists back David Cameron on weather link to global warming	0.1531
1/13/2014	EU considers scrapping 2030 binding renewables targets	0
1/14/2014	Compelling case for global deal on climate, says UN	0.2263
1/15/2014	Falling clean energy investment threatens UN climate goals. Under Investor Pressure, Utility to Study Emissions. Shale gas is no silver bullet for EU energy market	-0.1779
1/16/2014	EPA denies politics delayed pollution rules	-0.5719
1/17/2014	Sluggish Economy Prompts Europe to Reconsider Its Intentions on Climate Change. U.N. report: Don't delay on climate change. U.N. Says Lag in Confronting Climate Woes Will Be Costly. UN warns against delayed action on global warming. Kerry: No rush to decide on Keystone XL pipeline	-0.8732
1/19/2014	European Commission resists calls for carbon 'central bank'	0
1/20/2014	Rewrite energy policy and re-industrialise Europe. China's exports linked to western U.S. air pollution. Cost of EU energy widens over trade partners	0.4939
1/21/2014	Germany cautions on impact of renewables. High energy prices hold Europe back	0.2732
1/22/2014	Q&A: EU's 2030 climate and energy targets. New EU climate targets 'weaken renewable energy goals'. At Davos, push for clean energy as climate weapon. EU sets out new climate change goals. EU relaxes renewables target, reaps criticism. EU expected to trim ambitions on environment. Judges say Arctic offshore lease sale was flawed	0.3818
1/23/2014	European Union Proposes Easing of Climate Rules. New drive to cut green trade barriers	-0.0258
1/24/2014	Cameron, Bono link poverty, climate at AP debate	-0.5106
1/28/2014	Schäuble warns green policies are harming German economy. Obama to unilaterally raise minimum wage for federal contractors	-0.6124
1/29/2014	Obama's 'all of the above' energy policy offends everyone. Energy price gap with the US to hurt Europe for 'at least 20 years'. New North Sea oil will cost the climate. Obama calls for new incentives for cleaner fuel	-0.0516
1/30/2014	INSIDE WASHINGTON: Greens hit Obama on energy plan. Foundations Band Together to Get Rid of Fossil-Fuel Investments	0.2732
1/31/2014	Henry Waxman, Key Democrat and Force for Health Care Law, Is to Retire. U.S. report on Keystone indicates little climate impact. Australia approves dredging near Great Barrier Reef. Ex-NY mayor Bloomberg named UN climate envoy	0.8591

Financial Data Processing

We treated the headline sentiment, price, and volatility of each index fund over time as different time series. To conduct time series analysis, each time series must be stationary with a constant mean. Since the compound score for the climate sentiment time series has already been normalized, we only have to detrend and deseasonalize the financial data. To factor out macroeconomic conditions and identify any potential sub-trends, we linearly detrended the financial data, then de-seasonalized the data by subtracting the seasonal mean from each index fund closing price (fig. 2-1 and 2-2).

Here is the sample before vs after of detrending and deseasonalizing the VCSAX index fund's price volatility.

Figure 1

Fig. 2-1: Before vs After Linear De-trending; Fig. 2-2 Before vs After De-seasonalizing

detrend7.png	detrendp4.png

Methodology

After the financial and headline data were processed, we proceeded with the time series analysis. To model the impact of climate policy news sentiment on funds prices and volatility, we explored two distinct approaches: the Granger causality test and the comparison of univariate, autoregressive models to bivariate machine learning (ML) models. We divided our analysis into two different time frames: ten-year and yearly. We made this choice to examine whether causality or predictive modeling techniques would be improved over a short time frame, possibly due to the fact that perspectives on media may shift less over one year than ten years.

Granger Causality Test: While no test can evaluate true "causality," this test instead assesses "temporal precedence," the assumption that one variable is useful in forecasting the other (Bressler and Seth, 2011). The test is primarily used for linear relationships, and is based on the idea of fitting two autoregressive models on stationary time series data:

$$Y(t) = c + \beta_0 Y_{t-1} + \beta_1 Y_{t-2} + \beta_p Y_{t-p} + \epsilon_t \tag{1}$$

$$Y(t) = c + \beta_0 Y_{t-1} + \beta_1 Y_{t-2} + \beta_p Y_{t-p} + \epsilon_t + \gamma_0 \chi_m + \gamma_1 \chi_m - 1 + ...\gamma_p \chi_n + +\epsilon_t \tag{2}$$

as proven by Profillidis and Botzoris, 2019 where

$Y_t, Y_{t-1}, ...Y_{t-p}$	represents the full time series of Y
$X_m, X_{m-1}, ...X_n$	represents the full, lagged time series of X beginning at time m until time n
$\beta_0, \beta_1, ...\beta_p$ and $\gamma_0, \gamma_1, ...\gamma_n$	represent the parameters of the autoregressive model for variables Y and X
c	represents the constant term of the autoregressive model
ϵ_t	represents the error/residuals of the autoregressive model

If the ϵ_t term of the second model is statistically significantly smaller than the ϵ_t term of the first model, there is said to be an improvement in predicting Y *due to* X (i.e. X 'Granger-causes' Y). Using the F-statistic between the two models, we assess the following:

H_0; $p \geq .05$ (null hypothesis)	variable X does not Granger-cause Y
H_a; $p < .05$ (alternative hypothesis)	variable X Granger-causes Y

Univariate Autoregression vs Bivariate Machine Learning (ML)

We also compared the effectiveness of an autoregressive model, which only takes into account past prices, to six ML techniques, which take into account past prices as well as sentiment scores (Christoffersen, 2012; Karlsson, 2013.

In their 2023 paper, Mohsin and Jamaani, 2023 describe how ML methodologies' "capacity to simulate complicated properties such as non-linearity and volatility is highly valued" alongside traditional autoregressive forecasting models. We experimented with five different ML techniques:

First, the **Support Vector Machine (SVM)** is a supervised learning algorithm that constructs hyperplanes and conducts optimization problems to separate objects of different class labels (Roy et al., 2015). SVMs can be adapted for regression tasks, known as support vector regression (SVR), where the aim is to fit linear or non-linear models for relationship between variables (Reza Keyvanpour and Shirzad, 2022). Next, **Random Forest (RF) Regressor** is a type of ensemble technique that is used in classification and regression problems. For regression tasks, this means that multiple decision trees are trained on different subsets of the sample data, and a final prediction is made by averaging each individual decision tree (Muschelli et al., 2014).

Then, **Gradient Boosting Regressor (GBR)** is an ensemble technique. However, instead of training multiple decision trees at once, trees are sequentially constructed in an iterative manner to improve on past performance and generate a final prediction. Also, **Extreme Gradient Boosting (XGBoost)** is a robust and widely-used algorithm for regression. Like GBR, XGBoost sequentially and iteratively constructs decision trees to improve model performance. However, the model also employs several key techniques to avoid overfitting and enhance effectiveness. (Ekanayake et al., 2023, Xu et al., 2023). Lastly, **Light Gradient Boosting Machine (Light GBM)** was developed to improve model speed and further address overfitting through several techniques (Thongthammachart et al., 2022). We took an ensemble approach where we assessed the predictive ability of a simple average between the five models.

Data Analaysis

Our volatility data contains some missing days, namely weekends and holidays. Additionally, climate policy headlines are not available for every day over approximately the past ten years. Therefore, while our data spans 3435 days, we have 2935 days with headlines and 2372 days with volatility data. We only examined days with both headline and index fund data available.

Granger Causality Test

To determine if climate policy news sentiment has temporal precedence over index fund price and volatility, we ran Granger causality tests, considering a maximum lag of five days in two time frames: 10-year and year. We chose to look at two time frames

Time Frame	**Null Hypothesis:** ***for the S&P 500 and ten index funds analyzed***
Ten-year (110 tests)	H_0: on a 10-year time frame from January 2014 - Jun 2023, climate policy news sentiment does not have temporal precedence over fund price and volatility
Yearly (1100 tests)	H_0: on annual time frames from January 2014 - June 2023, climate policy news sentiment does not have temporal precedence over fund price and volatility.

A time series is considered to show Granger causality if at least one of its combinations shows statistically significant evidence of Granger causality at any lag value.

5400 individual tests were conducted with p-values ranging from 0.0003 to 0.999:

Time Frame	Rejecting H_0
Ten-year (110 tests)	4 out of 22 (18.18%) time series showed Granger causality over the entire decade
Yearly (1100 tests)	19 out of 22 (86.36%) time series showed Granger causality in at least one year. 31 out of 220 combinations (14.09%) showed evidence of Granger causality.

Table 3

Granger Causality 10-Year Volatility Test Results

Lag	VCDAX	VCSAX	VENAX	VFAIX	VGSLX	VGHCX	VITAX	VSPVX	VTCAX	VUIAX	VINAX
#1 P-Value	0.6852	0.0764	0.3189	0.9382	0.9016	0.6449	0.4329	0.9935	0.6687	0.517	0.8184
#1 Significant	no	no	no	no	no	no	no	no	no	no	no
#2 P-Value	0.8955	0.2309	0.555	0.8246	0.6599	0.7139	0.7502	0.6914	0.8598	0.7812	0.8403
#2 Significant	no	no	no	no	no	no	no	no	no	no	no
#3 P-Value	0.8166	0.0071	0.7515	0.499	0.0668	0.4079	0.2884	0.6585	0.6602	0.2461	0.9686
#3 Significant	no	yes	no	no	no	no	no	no	no	no	no
#4 P-Value	0.8574	0.0138	0.896	0.5184	0.0616	0.557	0.392	0.6975	0.283	0.322	0.9714
#4 Significant	no	yes	no	no	no	no	no	no	no	no	no
#5 P-Value	0.6859	0.0222	0.3478	0.5929	0.0505	0.2509	0.4502	0.626	0.2495	0.2837	0.7407
#5 Significant	no	yes	no	no	no	no	no	no	no	no	no

Table 4

Granger Causality 10-Year Price Test Results

Lag	VUIAX	VGHCX	VFAIX	VSPVX	VITAX	VGSLX	VINAX	VTCAX	VCSAX	VENAX	VCDAX
#1 P-Value	0.1041	0.2367	0.1026	0.2215	0.0061	0.6387	0.2009	0.064	0.0317	0.0878	0.0291
#1 Significant	no	no	no	no	yes	no	no	no	yes	no	yes
#2 P-Value	0.2706	0.2267	0.2857	0.4958	0.0241	0.8926	0.4308	0.1873	0.0672	0.196	0.0933
#2 Significant	no	no	no	no	yes	no	no	no	yes	no	no
#3 P-Value	0.3926	0.1222	0.2592	0.4628	0.0767	0.9196	0.5904	0.3621	0.051	0.1922	0.2167
#3 Significant	no	no	no	no	yes	no	no	no	yes	no	no
#4 P-Value	0.5692	0.1666	0.207	0.4498	0.1529	0.7488	0.5909	0.4518	0.1186	0.219	0.3344
#4 Significant	no	no	no	no	yes	no	no	no	yes	no	no
#5 P-Value	0.5762	0.1068	0.302	0.45	0.2489	0.8539	0.643	0.5893	0.0118	0.2987	0.4551
#5 Significant	no	no	no	no	yes	no	no	no	yes	no	no

Table 5

Year-Quarter-Fund Price combinations for which there is Granger Causality

Year	VCDAX	VCSAX	VENAX	VFAIX	VGSLX	VGHCX	VITAX	VSPVX	VTCAX	VUIAX	VINAX
2014	No	No	No	No	No	No	No	No	No	No	No
2015	No	No	No	No	No	Yes	No	No	No	Yes	No
2016	Yes	No	No	No	No	No	No	No	No	Yes	No
2017	No	No	Yes	No	No	No	No	No	No	No	No
2018	No	No	No	No	No	No	No	No	No	No	No
2019	No	No	No	No	No	No	Yes	Yes	No	No	No
2020	Yes	No	Yes	No	No	No	Yes	No	No	No	No
2021	No	No	No	No	No	No	No	No	No	No	No
2022	No	No	No	No	No	Yes	No	No	No	No	No
2023	No	No	No	No	Yes	No	No	No	No	No	No

Table 6

Year-Quarter-Fund Volatility combinations for which there is Granger Causality

Year	VCDAX	VCSAX	VENAX	VFAIX	VGSLX	VGHCX	VITAX	VSPVX	VTCAX	VUIAX	VINAX
2014	No	No	Yes	No	Yes	No	No	No	No	No	No
2015	No	No	No	No	No	No	No	No	No	No	No
2016	No	No	No	No	No	No	No	No	No	No	No
2017	No	No	No	No	No	No	No	No	No	Yes	No
2018	No	Yes	No	No	Yes	No	No	Yes	Yes	No	No
2019	No	No	No	Yes	No	No	No	No	No	No	No
2020	No	No	No	No	No	No	No	No	No	No	No
2021	No	No	No	Yes	No	No	No	No	No	No	No
2022	No	No	No	No	No	No	No	No	No	No	No
2023	Yes	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes

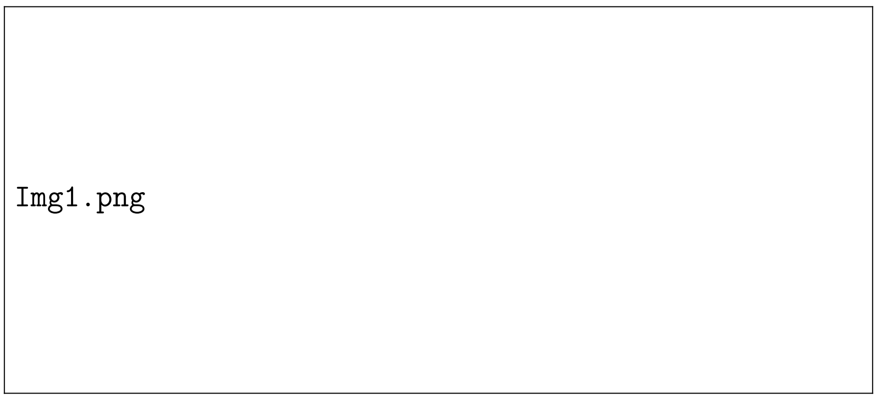

Figure 2

Mean absolute percentage error across autoregression and ML techniques for index fund volatility

Figure 3

Mean absolute percentage error across autoregression and ML techniques for index fund price

The mean absolute percent errors for both categories (price and volatility) indicate overall low accuracy by each of the techniques. However, for specific funds during specific time frames, we see improved accuracy. For example, the only yearly instances when machine learning methods outperformed autoregression were during 2023. This occurred twice among fund-year-techniques for index fund price and thirteen times for index fund volatility. The outperforming machine learning models were predominantly XGBoost (eight out of thirteen instances for volatility and two out of two instances for price), but GBM and RF also outperformed autoregression for four and one fund during 2023, respectively.

Disregarding performance relative to autoregression, the machine learning technique with the highest performance (lowest MAPE) on a yearly level was XGBoost. The six fund-year-technique combinations with the highest performance for volatility used XGBoost and the eight fund-year-technique combinations with the highest performance for price were also XGBoost.

Table 7

Comparing mean absolute percentage error rate between autoregression and machine learning techniques (SVR)for index fund price and volatility over 10 years

Volatility	Autoregressive MAPE	ML SVR Model MAPE	Price	Autoregressive MAPE	ML SVR Model MAPE
VCDAX	2.79482518	1.884728411	VCDAX	1.899636947	1.290444085
VCSAX	1.500764728	1.332623482	VCSAX	4.036113628	1.829663946
VENAX	1.404409573	1.370830368	VENAX	1.752577423	1.161190294
VFAIX	1.303364222	1.380001391	VFAIX	1.632842048	1.290144705
VGSLX	2.229505407	1.565339304	VGSLX	3.18283123	1.330611586
VGHCX	2.272729407	2.447401826	VGHCX	1.290031539	1.118670719
VITAX	1.654836197	1.630861196	VITAX	2.326662209	1.24214393
VSPVX	2.495081613	1.49699782	VSPVX	5.2273169	2.453597382
VTCAX	4.186263456	1.526308515	VTCAX	7.169936925	6.812388764
VUIAX	1.411882279	1.671035171	VUIAX	1.997958171	1.153352809
VINAX	3.552011519	2.42932703	VINAX	1.62061737	1.319052665

Discussion

Overlaps in Different Approaches

To determine the industries (and index funds) most significantly impacted by climate policy headline sentiment, we first examined overlaps in the results of the Granger causality and selected ML models in the overall, ten-year time frame.

According to the Granger causality test, **VCSAX** had a statistically significant relationship with climate policy news sentiment for both price and volatility, while **VITAX** and **VCDAX** had a statistically significant relationship with climate policy news sentiment for price.

Using the MAPE rates to compare both the autoregressive and ML models, we found that the ML approach generated lower error rates than the autoregressive approach on a ten year span for VCDAX, VCSAX, VENAX, VGSLX, VITAX, VSPVX, VTCAX, and VINAX for **volatility** and VCDAX, VCSAX, VENAX, VGSLX, VGHCX, VITAX, VSPVX, VTCAX, VUIAX, and VINAX for **price**. For both volatility and price, there was a complete overlap (volatility: VCSAX; price: VCDAX, VCSAX, VITAX) in the index funds for which statistically significant Granger causality was found and the bivariate ML model outperformed the univariate autoregressive model. This suggests that, on a ten year span, these are the index funds for which accounting for climate sentiment actually improves your index fund prediction accuracy.

We conducted a similar analysis of ML and autoregressive mean absolute percentage error rates on a **yearly** time frame. We selected the SVR model for this analysis, as it performed with the lowest error for volatility, and second-lowest for price. A Chi-square test indicated that ML model SVR did not perform statistically significantly better on a yearly level for funds which showed Granger causality across a decade than funds which did not show Granger causality.

It is possible that the VADER sentiment scoring process may have contributed to the machine learning models' heightened prediction error relative to the autoregressive method for many of the index funds. Despite being accurate most of the time, we found many instances where VADER scored headlines with specific climate-related jargon incorrectly. This is especially noticeable on days with only one headline, since VADER was originally intended to analyze Twitter posts, which tend to be longer than headlines. Moreover, headlines tend to use more polished, formal language than tweets, which may cause VADER to categorize most headlines as neutral or near-neutral, even if they may seem quite negative or positive. The most likely conclusion regarding the efficacy of machine learning models and index funds is that headline sentiment generally does not enhance the ability to predict price and volatility.

As demonstrated earlier in Figure 1, there are not headlines every day, and the headlines that exist are not evenly spaced temporally. While we did not take into account days missing either index fund data or headline data, we do not believe our results were significantly impacted. There is no index fund data for weekends, but laws are generally not passed on weekends. Therefore, any major climate policy events likely take place during the week. Calculating a weekend index fund value would require additional assumptions.

It is important to note that existing market conditions may be responsible for the fact that we did not find Granger causality for most index funds excepting VCSAX, VCDAX, and VITAX. The rise of automation and algorithmic trading has caused fewer manual trades proportional to the total number of trades across financial markets. This may mean that investors who make trading decisions based upon climate policy may not 'move the needle' on index fund price or volatility because their impact is outweighed by algorithmic trading. We were particularly surprised, however, that VENAX (which has a portfolio composition including Exxon Mobil, Chevron, and ConocoPhillips), seemed unaffected by climate policy sentiment.

Conclusion

Ultimately, our findings suggest that the consumer discretionary (Amazon, Nike, Disney, etc), consumer staples (P&G, Costco, Walmart, etc) and information technology (Apple, Microsoft, Amazon, etc) sectors have the most significant relationship with climate policy headline sentiment. Conversely, sectors such as energy, real estate, and healthcare are not significantly impacted by climate policy headlines. When machine learning methods such as random forest regression and gradient boosting are used, they outperform autoregression on a ten-year scale but tend to demonstrate worse performance than autoregressive methods on a yearly level when the mean absolute percentage errors are compared. Support vector machine was the highest performing machine learning method on a ten-year scale, but XGBoost, widely considered to be a robust ML model, performed better on a yearly scale.

Through analyzing climate policy sentiment patterns over the past ten years and detrending financial metrics, our research importantly suggests that the sentiment behind climate policy headlines does not meaningfully impact most index funds. The funds which most notably do not show impact include VENAX (energy) and VUIAX (utilities). This may reflect these funds' unique resilience to shifting media attitudes with regards to climate policy developments.

Further Investigation

We see the opportunity to continue this strand of research by optimizing a few critical elements of our methodology and adding slight modifications to our research question. Our approach to classifying the "positivity" of various headlines was subject to various inaccuracies, since VADER is not specific to climate-related text. As such, the accuracy of climate-related sentiment in our dataset could be better represented using a model trained on climate-related text. Fine-tuning ClimateBERT (Bingler et al., 2023), a language processing model developed to track the strength of corporate ESG commitments, to climate policy news headlines could provide a promising opportunity for further investigation.

It may also be promising to leverage ESG in our investigation of the predictive ability of climate policy news sentiment on financial metrics. It is possible that firms with high ESG rankings may be more impacted by climate policy news, and incorporating a particular security or firm's ESG ranking as a predictive feature alongside climate policy news sentiment could be a fruitful avenue for future investigation.

Acknowledgements

The authors acknowledge the 2023 Columbia University Summer Undergraduate Research Experiences in Mathematical Modeling for providing the foundational platform for this investigation. Further recognition is extended to the Columbia Department of Mathematics for its consistent support during the research endeavor.

Appreciation is also directed towards Professor George Dragomir, the program director, for his structured guidance. Additional thanks are reserved for Casey Ivanovich, student mentor and PhD Candidate within the Department of Earth and Environmental Sciences, for her constructive feedback and mentorship throughout the course of this study.

Finally, we extend our gratitude to our fellow CSUREMM students for their constructive support and collaborative spirit during this research journey.

Appendix

Table 8

This list of climate search terms entered into LexisNexis news article database was combined with the term "climate" to create three search categories total. This ensured that the articles would always relate to climate in some way and excluded articles that might discuss, for example, "power" in a non-climate context.

Item	
Term Type	Term
Climate	hydrogen OR power plants OR methane OR natural gas OR oil OR solid waste landfills OR “cap-and-trade” OR clean energy OR biofuel OR renewable energy OR fuel efficiency OR “cap and trade” OR ZEV OR nuclear energy OR energy storage OR sea level rise OR flooding OR fuel economy OR restoration OR resilience OR carbon sequestration OR public lands OR nitrogen oxides OR energy efficiency OR carbon reduction technologies OR coal OR plant OR electricity OR carbon dioxide OR power OR utility OR gas OR energy OR fossil fuel OR wind OR nuclear power OR pipeline OR greenhouse gas emission OR emission OR carbon
Regulations	regulation OR law OR bill OR act OR order OR policy OR framework OR authorization OR mitigation OR initiative OR plan OR agreement OR talk OR deal OR meeting OR summit OR action OR fund OR funding OR development OR budget OR target OR goal OR government OR legislation OR cap OR vote OR lawmaker OR measure OR program OR legal action OR state OR administration OR standard OR rule OR regulation OR agency OR plan OR court OR decision OR case OR carbon credits market OR market OR scheme OR credit OR permit OR tax OR carbon tax OR scheme OR investment OR invest

References

Ardia, D., Bluteau, K., Boudt, K., & Inghelbrecht, K. (2020). Climate change concerns and the performance of green versus brown stocks. *Available at SSRN 3717722*.

Bhavsar, H., Jivani, A., Amesara, S., Shah, S., Gindani, P., & Patel, S. (2022). Stock price prediction using sentiment analysis on news headlines. In *Ict with intelligent applications: Proceedings of ictis 2022, volume 1* (pp. 25–34). Springer.

Bingler, J., Kraus, M., Leippold, M., & Webersinke, N. (2023). *How cheap talk in climate disclosures relates to climate initiatives, corporate emissions, and reputation risk* (Working paper). Available at SSRN 4000708.

Bressler, S. L., & Seth, A. K. (2011). Wiener–granger causality: A well established methodology. *Neuroimage*, *58*(2), 323–329.

Christoffersen, P. F. (2012). 3 - a primer on financial time series analysis. In P. F. Christoffersen (Ed.), *Elements of financial risk management (second edition)* (Second Edition, pp. 39–64). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-374448-7.00003-8

Ekanayake, I., Palitha, S., Gamage, S., Meddage, D., Wijesooriya, K., & Mohotti, D. (2023). Predicting adhesion strength of micropatterned surfaces using gradient boosting models and explainable artificial intelligence visualizations. *Materials Today Communications*, *36*, 106545. https://doi.org/https://doi.org/10.1016/j.mtcomm.2023.106545

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies*, *33*(3), 1184–1216.

Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the international AAAI conference on web and social media*, *8*(1), 216–225.

Karlsson, S. (2013). Chapter 15 - forecasting with bayesian vector autoregression. In G. Elliott & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 791–897). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-62731-5.00015-4

Mohsin, M., & Jamaani, F. (2023). Green finance and the socio-politico-economic factors' impact on the future oil prices: Evidence from machine learning. *Resources Policy*, *85*, 103780. https://doi.org/https://doi.org/10.1016/j.resourpol.2023.103780

Muschelli, J., Betz, J., & Varadhan, R. (2014). Chapter 7 - binomial regression in r. In M. B. Rao & C. Rao (Eds.), *Computational statistics with r* (pp. 257–308). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-444-63431-3.00007-3

Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2022). Dissecting green returns. *Journal of Financial Economics*, *146*(2), 403–424.

Profillidis, V., & Botzoris, G. (2019). Chapter 7 - econometric, gravity, and the 4-step methods. In V. Profillidis & G. Botzoris (Eds.), *Modeling of transport demand* (pp. 271–351). Elsevier. https://doi.org/https://doi.org/10.1016/B978-0-12-811513-8.00007-8

Ramelli, S., Wagner, A. F., Zeckhauser, R. J., & Ziegler, A. (2021). Investor rewards to climate responsibility: Stock-price responses to the opposite shocks of the 2016 and 2020 us elections. *The Review of Corporate Finance Studies*, *10*(4), 748–787.

Reza Keyvanpour, M., & Shirzad, M. B. (2022). Chapter 14 - machine learning techniques for agricultural image recognition. In M. A. Khan, R. Khan, & M. A. Ansari (Eds.), *Application of machine learning in agriculture* (pp. 283–305). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-323-90550-3.00011-4

Roy, K., Kar, S., & Das, R. N. (2015). Chapter 6 - selected statistical methods in qsar. In K. Roy, S. Kar, & R. N. Das (Eds.), *Understanding the basics of qsar for applications in pharmaceutical sciences and risk assessment* (pp. 191–229). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-801505-6.00006-5

Thongthammachart, T., Araki, S., Shimadera, H., Matsuo, T., & Kondo, A. (2022). Incorporating light gradient boosting machine to land use regression model for estimating no2 and pm2.5 levels in kansai region, japan. *Environmental Modelling Software*, *155*, 105447. https://doi.org/https://doi.org/10.1016/j.envsoft.2022.105447

Vicari, M., & Gaspari, M. (2021). Analysis of news sentiments using natural language processing and deep learning. *AI & society*, *36*, 931–937.

Xu, N., Wang, Z., Dai, Y., Li, Q., Zhu, W., Wang, R., & Finkelman, R. B. (2023). Prediction of higher heating value of coal based on gradient boosting regression tree model. *International Journal of Coal Geology*, *274*, 104293. https://doi.org/https://doi.org/10.1016/j.coal.2023.104293