Media Climate Change Concern and Stock Returns

Liya Chu Jun Tu Luying Wang

Abstract

This study constructs a comprehensive media climate change concern index based on the media coverages about climate change across a wide range of media formats from print text (newspapers) to voice (radios), and further to video (televisions). We find that the monthly change in the comprehensive index negatively predicts the aggregate stock market returns, both in and out-of-sample, and can deliver sizable economic gains for mean-variance investors in asset allocation. The predictive power of our comprehensive index holds after controlling for various previously studied market return predictors, and is robust to using many alternative approaches. Our evidence illustrates the strong impact of climate change concern on the aggregate stock market.

Keywords: Return Predictability, Climate Change, Media, Sustainable Finance, Environment

---

We thank seminar participants at Central University of Finance and Economics and Singapore Management University Lee Kong Chian School of Business for their insightful and constructive comments.
Chu: Assistant Professor of Finance, Business School, East China University of Science and Technology, 200237, China
Tu: Associate Professor of Finance, Lee Kong Chian School of Business, Singapore Management University, 178899, Singapore
Wang: PhD Candidate in Finance, Lee Kong Chian School of Business, Singapore Management University, 178899, Singapore

1. Introduction

Climate change seems to be 'the defining issue of our time' given its important role on confronting the escalating threats from rising sea levels, natural disasters, etc. Moreover, besides those physical threats, climate change matters from other aspects, such as triggering regulatory shocks including carbon taxes and potential opportunities for green technologies. Given its broad implications in both the natural and social world, climate change has become one central topic in media.

Through disseminating information to a wide range of audience, such as investors, consumers, firm managers, and policymakers, the widespread mass media can have substantial influence on raising public concern on climate change and global warming, and the importance of a transit to a low-carbon economy. This may be one reason why a fast-growing number of managers and investors are adopting the idea of sustainable investing associated with climate change and more broadly with environmental, social, and governance (ESG). Anecdotal evidence shows that by the year of 2020, about one third of the total US assets under professional management could be based on sustainable investing strategies.

Recent growing literature investigating the effects of climate change on financial market mainly focuses on the heterogeneity in exposure to climate change across different types of firms. For example, Pastor et al. (2021) and Ardia et al. (2021) examine the impacts of unexpected change in climate change concern on green versus brown stocks respectively. However, whether the change in climate change concern can affect the aggregate stock market is largely left unexamined.[1] Considering its potential strong influence on the financial market through the sizable sustainable investing, it seems important to examine whether and how the concern of climate change reflected in and probably also reinforced by media coverage can affect the aggregate stock market returns.

In this study, we first construct a comprehensive media climate change concern index based on the media coverage from a wide range of sources including newspaper, radio, television, and wire service. We then investigate the forecasting power of the change in the comprehensive

[1] For instance, the relatively positive (negative) impacts of unexpected change in climate change concern on green (brown) firms may cancel out when aggregated into market level.

media climate change concern index ($\Delta CMCCC$) on the aggregate stock market returns. Engle et al. (2020) construct time-series indices using newspaper articles on climate related risk, but the primary goal of their study is to build portfolios to hedge climate change risk. Ardia et al. (2021) propose a media climate change concern index, but their purpose is to empirically test the prediction of Pastor et al. (2021) that green firms outperform brown firms when climate change concerns increase unexpectedly. Our paper is distinct from them and contributes to the existing studies at least in two aspects. First, we aim at the role of the $\Delta CMCCC$ index in explaining time-series return predictability on the aggregate stock market, which impacts many fundamental areas of finance from portfolio theory to capital budgeting and is one of the central issues in finance as emphasized by Cochrane (2008), whilst the existing studies mainly focus on the impact of cross-sectional stock returns. Second, we consider a broad range of different forms of media sources, from print text (newspapers) to voice (radios), and further to video (televisions), when constructing the $\Delta CMCCC$ index. In contrast, other studies limit their examinations on one form of media, that is, the print text type of newspapers.

Motivated by recent studies demonstrating the importance of visual channel from news photos and audio channel from music songs in predicting market returns via conveying sentiment efficiently (e.g., Obaid and Pukthuanthong 2021; Edmans et al. 2021), we conduct a more comprehensive analysis of climate change concern incorporating visual and audio content, and explore the role of a variety of mass media communication channels such as televisions and radios[2]. Specifically, we consider 15 individual media climate change concern measures based on five newspapers, seven televisions, one radio, and two wire services, respectively. The 15 individual media outlets we select have more readers, viewers, and listeners than many others and are among the most influential ones in the United States. We find that most of the changes in individual media climate change concern measures are negatively related with subsequent stock market return. Approximately half of them can significantly predict next month market return in-sample, and three of them have significant power out-of-sample. More specifically, changes in individual media climate change concern measures based on media coverage with videos and sounds (e.g., television and radio) tend to predict market return better than changes in individual concern measures based on media coverage with printed text and pictures (e.g., newspaper). Although changes in some individual media climate change concern measures,

[2] Some newspaper articles may include a handful of static images, but they should mostly not as vivid as dynamic videos from TVs.

especially those based on television and radio coverage, exhibit certain time-series forecasting power in one-month horizon, they fail to perform well in longer horizons. Hence, the overall forecasting performance of changes in individual media climate change concern measures is limited. This can be expected as individual concern measure relying on only one specific media source might serve as noisy proxy for aggregate media climate change concern, resulting in unstable forecasting performance.

Next, we propose a comprehensive index by aggregating multiple media information from 15 individual media climate change concern measures. In contrast to prior studies on climate change news which covers newspapers only, our comprehensive index is novel in that it is able to capture audio and visual (from both static images and dynamic videos) sensations about climate change concern. Our primary aggregation method is the partial least squares (PLS) method in Kelly and Pruitt (2013, 2015). PLS is an efficient method to obtain aggregated climate change concern from various individual measures. We find that change in comprehensive media climate change concern (i.e., $\Delta CMCCC^{PLS}$) significantly predicts market returns up to 12 months. A one standard-deviation increase in $\Delta CMCCC^{PLS}$ is associated with a 0.68% decrease in the next one-month market return and a 6.36% decrease in the next 12-month market return over the sample period from 2002 to 2021. The in- and out-of-sample R^2s are 2.53% and 2.86% at the one-month horizon and 16.26% and 10.07% at the 12-month horizon. We also use equal-weight and volatility-weight approaches to aggregate information from multiple media sources, and find similar albeit slightly weaker results: a one standard-deviation increase in $\Delta CMCCC^{Equ}$ ($\Delta CMCCC^{Vol}$) is associated with a 0.57% (0.56%) decrease in the next one-month market return and a 5.97% (5.86%) decrease in the next 12-month market return. As additional robustness checks, we consider alternative method such as forecast combination, and other machine learning methods such as Elastic net and Lasso to aggregate individual media climate change concern measures. We find that all of them consistently generate significant out-of-sample R^2s, with the magnitudes being slightly smaller at the one-month horizon and larger at the 12-month. For example, using forecast combination method, the out-of-sample R^2 is 1.83% at the one-month horizon and 15.11% at the 12-month horizon, respectively; both are significant at the 5% level. These results uncover genuine predictability of change in comprehensive media climate change concern on aggregate market returns.

Furthermore, we compare the predictive power of $\Delta CMCCC^{PLS}$ with common return predictors, including macroeconomic variables used by Goyal and Welch (2008), various sentiment measures (i.e., investor sentiment index in Baker and Wurgler 2006, aligned investor sentiment index in Huang et al. 2014, news sentiment measures in Calomiris and Mamaysky 2019), and alternative climate related concern measures (i.e., natural disaster and environmental concern measures in Bybee et al. 2020, climate change news indices in Engle et al. 2020, media climate change concern index in Ardia et al. 2021 and Pastor et al. 2021). We find that $\Delta CMCCC^{PLS}$ maintains significant predictability after controlling for them. The results suggest that monthly change in comprehensive media climate change concern contains substantial explanatory power for the stock market, which is beyond the economic fundamental variables, various sentiment indices and alternative climate related concern measures.

In addition to the superior predictability of $\Delta CMCCC^{PLS}$, we also investigate whether it can yield sizable economic gains. We show that change in comprehensive media climate change concern can lead to sizable investment gains for a mean-variance investor from an asset allocation perspective as well. The annualized certainty equivalent return (CER) gains of $\Delta CMCCC^{PLS}$ are 5.40% and 3.44% under no transaction cost and 50bps transaction costs, respectively, when the investor with a risk aversion coefficient of three allocates investments between the market and risk-free rate. Moreover, investment portfolios based on $\Delta CMCCC^{PLS}$ can reach annualized Sharpe ratios ranging from 1.01 to 1.20 for investors with different levels of risk aversion coefficients. The asset allocation results are robust to predictors constructed based on alternative aggregating methods, i.e. $\Delta CMCCC^{Equ}$ and $\Delta CMCCC^{Vol}$.

Finally, we provide additional insights to help us better understand the economic driving force through which $\Delta CMCCC^{PLS}$ relates to market return. We investigate whether the predictability of $\Delta CMCCC^{PLS}$ comes primarily from time variation in cash flows or discount rates. Specifically, in order to disentangle the source of predictability for $\Delta CMCCC^{PLS}$, we adopt the Campbell (1991) and Campbell and Ammer (1993) vector auto-regression (VAR) analysis and the information contained in widely used economic predictors to decompose total stock market return into three components: the expected return, discount rate news, and cash flow news. We find that the ability of $\Delta CMCCC^{PLS}$ to predict future stock market return results predominantly from its ability to predict future cash flow news. When media climate change concern increases, both consumers' demands for green products and investors' demands to hold green assets increase, however, brown firms' costs upon climate-concern shocks such as tighter environmental regulations might be larger than green firms' benefits, which generates a negative net cash flow effect and predicts a low return. This is especially the case when there is a delay in incorporating this effect into price as illustrated below.

As discussed in Pastor et al. (2021), green stocks have lower expected return than brown stocks, whereas green stocks can outperform brown stocks due to shocks of environmental concerns which might shift tastes of investors and consumers towards green direction. Hence, we examine the differential impacts of shocks to comprehensive media climate change concerns on green-versus-brown stock returns, respectively. We find that an increase in comprehensive media climate change concern shock appears to affect both green and brown stocks with a delay. Moreover, the delayed reaction tends to be stronger for brown stocks. Therefore, we observe that the return spread of green-minus-brown portfolio positively reacts to comprehensive media climate change concern shock with a delay. Our result confirms empirical findings in the literature which document that stock prices are slow to incorporate climate news, and brown stocks response to climate news more slowly than green stocks (e.g., Pastor et al. 2021). Our result in the predictability of the market return across time is consistent with the overall evidence of the delay for the stock markets to react to the media climate change concern shock.

Our study fits into the burgeoning literature on "climate finance". Early research within the field begins from the interactions between climate change and the economy via macro-finance models. It is only recently that researchers have started to test the implications of climate change on financial markets. Many studies have explored the cross-sectional impacts of climate risks across a large number of asset classes including equities, fixed income securities, and real estate (see Giglio et al. 2020 for a review). Bolton and Kacperczyk (2020 a, b) claim that firms with high carbon emissions are perceived as riskier and investors demand compensation for investing in these firms. Hsu et al. (2020) show a similar return spread pattern between high- and low- pollution firms. Baker et al. (2018) find that green bonds trade at lower yields than bonds without green designation. Murfin and Spiegel (2019) explore whether house prices reflect differential sea level rise risk. Ilhan et al. (2020) document that high carbon intensities associated with higher tail risk are priced in the option market. Engle et al. (2020) aim at building portfolios to hedge climate change risk. Ardia et al. (2021) empirically test the prediction of Pastor et al. (2021) that green firms outperform brown firms when climate change concerns increase unexpectedly. Our paper is distinct from them and contributes to the existing studies by investigating the impact of media concern on climate change on time-series return predictability for the aggregate stock market.

This paper is also related to recent research exploiting attention to and/or concern on climate change and its impact on the financial market. Choi et al. (2020) document that when the local temperature is abnormally high, retail investors tend to pay more attention on climate change, and they are more likely to sell carbon-intensive firms' stocks. Giglio et al. (2020) construct a measure of attention to climate risk in the housing markets. Engle et al (2020) propose measures about attention/concern to climate change using climate news in Wall Street Journal. Ardia et al. (2021) and Pastor et al. (2021) both consider climate change/environmental concerns based on newspaper reports. Our paper complements this strand of literature by exploiting a more extensive range of various forms of media coverages, from print text, audio, to video, in driving aggregate climate change concern and is the first to examine how it helps us understand and predict the aggregate stock market return over time.

The remainder of the paper is organized as follows. Section 2 describes the data and shows the limited predictive power of 15 individual media climate change concern measures over one- to 12-month horizons. Section 3 proposes a PLS comprehensive media climate change concern index and shows its significant forecasting power of market return both in-sample and out-of-sample. Section 4 studies the economic source of the predictability. Section 5 concludes.

2. Forecasting power of individual media climate change concern measures

In this section, we show that the individual media climate change concern measures are lack of the capability to predict stock market return both in-sample and out-of-sample over one- to 12-month horizons persistently.

2.1. Individual Media Climate Change Concern measures

We measure individual media climate change concern (i.e., MCCC) based on media coverage data from 15 different sources. Specifically, we employ the monthly number of news coverages about climate change from five newspapers, seven televisions, one radio between January 2000 and September 2021, and two wire services between January 2004 and September 2021 in U.S. The data are collected by accessing archives through the Lexis Nexis, ProQuest and Factiva databases and these sources are selected based on their circulations over time. The five newspapers include Washington Post, Wall Street Journal, New York Times, USA Today and Los Angeles Times; the seven televisions include ABC, CBS, CNN, FOX, MSNBC, NBC and PBS; one radio is the National Public Radio (US); two wire services include Associated Press and United Press International.

Following Ardia et al. (2021)' s method, we first take the square root of this number since they mentioned that "One concerning article about climate change may increase concerns, but 20 concerning articles are unlikely to increase concerns 20 times more". Then, as it usually takes a long time to generate concern about climate change among people, we use 24 months moving average to capture source-specific concern level about climate change as

$$MCCC_{i,t} = (1/24)\sum_{t-23}^{t}\sqrt{N_{i,t}} \;\; , i = 1,2, \dots 15 \qquad (1)$$

where $N_{i,t}$ is the number of news articles published about climate change on month t by source i. $MCCC_{i,t}$ represents individual media climate change concern from source i.[3] Moreover, in line with many studies (e.g., Engle et al., (2020 RFS), Pastor et al., (2021 JFE)) we assume that no news is good news for climate change.

Finally, we follow Pastor et al., (2021) to define changes in climate concern as $\Delta MCCC_{i,t}$= $MCCC_{i,t}$-$MCCC_{i,t-1}$ as an unanticipated shock to stock market. In doing so, the sample period for five newspapers, seven televisions, one radio spans from January 2002 to September 2021, and sample period for two wire services spans from January 2006 to September 2021. [4]

Table 1 reports the median, quartile (25% and 75%) distributions, skewness, and first-order autocorrelation coefficient of the 15 individual $\Delta MCCC_{i,t}$ proxies. All concern variables are standardized to have mean of 0 and variance of 1. As shown in the table, for unanticipated changes about climate change, the values of median vary from -0.18 for PBS to 0.08 for Wall Street Journal. $\Delta MCCC_{CNN}$ has the largest 75% quartile and the smallest 25% quartile among all variables. Moreover, most of the measures are positive skewed. The positive first-order autocorrelation coefficients indicate some persistent for these measures. Table 2 shows the pairwise correlations between individual media climate change concern measures. We observe

[3] The results are similar when we replace the moving average of 24 months by the moving average of alternative horizons, such as 18 months or 30 months.
[4] The results are similar when we use AR(1) error series as unexpected climate change concern, we report the results in the Appendix table A1.

that all the individual concern measures are positively correlated, and the coefficients range from 0.25 to 0.85. This indicates that all these 15 individual concern measures contain not only some common component capturing investors' concern about climate change and/or other common activities (e.g., climate gate around the year of 2010) but also sizable noises specific to individual measures.

[Insert Table1 and Table 2 about here]

2.2. Forecasting power of individual $\Delta MCCC_{i,t}$ proxies

We explore the forecasting power of media climate change concern on stock market return based on the following predictive regression:

$$R_{t+1} = \alpha + \beta \times X_t + \varepsilon_{t+1} \tag{2}$$

where R_{t+1} is the log excess return of the S&P 500 index in month $t+1$. When forecasting market return in h months, we denote the cumulative market return as $R_{t,t+h} = \sum_{j=1}^{h} R_{t+j}$, where $h = 1,3,6,$ and 12. X_t is one of the 15 media climate change concern measures $\Delta MCCC_{i,t}$.

The in-sample predictive ability of $\Delta MCCC_{i,t}$ is tested by estimating the regression (2). Specifically, if the estimate of β is significantly different from zero in regression (2) or the in-sample R^2 is significantly large than zero, it means that $\Delta MCCC_{i,t}$ is informative for predicting market return. A time-varying expected stock return model applies. We use Newey and West (1987) standard error to compute the *t*-statistic corresponding to $\hat{\beta}$.

Besides in-sample test, we also include the out-of-sample analysis as Goyal and Welch (2008), among others, have argued that out-of-sample tests can be more relevant for estimating genuine return predictability in the practice. Basically, the out-of-sample forecast of next one-month market return is computed as

$$\hat{R}_{t+1} = \hat{\alpha}_t + \hat{\beta}_t \times X_t \tag{3}$$

where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the OLS estimates of regression (2) based on data from the start of the sample period through month t. We recursively estimate regression (2) and repeatedly construct the monthly out-of-sample forecasts based on Equation (3) for the following periods, until we get to the end of the sample period.

To assess the out-of-sample performance, we use the out-of-sample R^2 statistic in Campbell and Thompson (2008) and define it as

$$R_{OS}^2 = 1 - \frac{\sum_{t=M+1}^{T}\left(R_t - \hat{R}_t\right)^2}{\sum_{t=M+1}^{T}(R_t - \bar{R}_t)^2} \tag{4}$$

where M is the in-sample training period and $T - M$ is the out-of-sample forecasting period. $\hat{R}_t$ denotes the estimated market return based on Equation (3), and $\bar{R}_t$ is the historical mean of market return, both of which are estimated based on data up to month $t-1$. If $\Delta MCCC_{i,t}$ is a valid predictor of market return, its mean squared forecast error (MSFE) should be lower than MSFE of historical mean which indicates a positive R_{os}^2. To evaluate whether the predictive regression forecast generates a statistically significant improvement in MSFE, we use the *MSFE-adjusted* statistic proposed by Clark and West (2007) to test the null hypothesis that the historical mean MSFE equals to or is less than that of the predictive regression forecast against the alternative hypothesis, which claims that the historical mean MSFE is larger compared to that of the predictive regression forecast, corresponding to H_0: $R_{OS}^2 \leq 0$ against H_A: $R_{OS}^2 > 0$.

[Insert Table 3 about here]

The table 3 presents the regression slope β(%), Newey–West *t*-value, in-sample R^2(%), and out-of-sample R_{os}^2(%) of predicting market returns with individual media climate change concern measures over one- to twelve-month horizon. The in-sample period for first 13 individual measures spans from 2002:01 to 2021:09 and for last two variables spans from 2006:01 to 2021:09. Throughout this paper, the out-of-sample forecast period is from 2009:01 to 2021:09. Statistical significance for R_{os}^2 is based on the *p*-value of the Clark and West (2007) MSFE-adjusted statistic for testing H_0: $R_{os}^2 \leq 0$ against H_A: $R_{os}^2 > 0$. In the left panel, it reports the results of forecasting next one-month market return with individual media climate change concern measures. Except for $\Delta MCCC_{CBS}$, all the other individual measures have a negative predictive sign and seven of them reveal significant in-sample predictive power at the 5% level. Moreover, among these seven measures, three of them have positive and significant out-of-sample R^2_{os}. The middle and right panel of table 3 reports the similar results when extending the prediction horizon to three and twelve months respectively. The in-sample regression slopes are mostly negative while only limited number of individual measures to perform well both in- and out- of sample.

When taking all three panels into consideration together, Table 3 shows that while some of individual media climate change concern measures may have certain time-series forecasting power for certain horizons, such as one-month horizon, none of them could keep its forecasting power across multiple horizons from one-, three- to 12-month. For instance, for one-month horizon, $\Delta MCCC_{NBC}$ exhibits the highest out-of-sample R^2_{os} of 4.04%, and it also generates the highest in-sample R^2 of 3.14%. Moreover, $\Delta MCCC_{MSNBC}$ and $\Delta MCCC_{radio}$ also exhibit the significant in-sample coefficients and out-of-sample R^2_{os}. However, all these three individual measures, performing well at one-month horizon, lose their forecasting power out-of-sample at three-month and twelve-month horizons.

3. Change in Comprehensive Media Climate Change Concern index

Existing studies often examine media climate change concern/risk from one specific media source (e.g., Wall Street Journal) or from one media channel (e.g., newspapers). Although any specific media source or media channel may serve as a proxy for the underlying media climate change concern, each individual proxy could be noisy, which may result in non-stable performances as mentioned above. In this study, we extract the common component of the 15 individual measures, which are from 4 different media sources, to create the change in Comprehensive Media Climate Change Concern (i.e., $\Delta CMCCC$) index, which is expected to be a better proxy than those potentially too noisy individual measures.

3.1. Forecasting Model

The predictive regression exploring the forecasting power of comprehensive media climate change concern measures on stock market return is as follows:

$$R_{t+1} = a + b \times \Delta CMCCC_t^* + \varepsilon_{t+1} \tag{5}$$

where R_{t+1} is realized excess stock market return in month $t+1$, $\Delta CMCCC_t^*$ is the unobservable change in comprehensive media climate change concern measure in month t, and ε_{t+1} denotes a noise term that is unforecastable and irrelevant to $\Delta CMCCC_t^*$.

Then, we assume a linear factor structure for the individual media climate change concern measures. Let $\Delta MCCC_t = (\Delta MCCC_{1,t}, \ldots, \Delta MCCC_{N,t})'$ denote an $N \times 1$ vector of individual predictors which refer to N individual media climate change concern measures at time t; N represents the number of individual measures. The model of $\Delta MCCC_{i,t}$ $(i = 1, \ldots N)$ is given by the following factor structure:

$$\Delta MCCC_{i,t} = \eta_{i,0} + \eta_{i,1} \times \Delta CMCCC_t^* + \eta_{i,2} \times Error_t + e_{i,t} \quad (6)$$

where $\Delta CMCCC_t^*$ is the true but unobservable change in comprehensive media climate change concern in model (5), $\eta_{i,1}$ is the slope coefficient that summarizes the sensitivity of individual measures $\Delta MCCC_{i,t}$ to the movement of the unobservable change in comprehensive media climate change concern measure $\Delta CMCCC_t^*$, $Error_t$ represents the common approximation error component of all the proxies that are unrelated to stock returns, and $e_{i,t}$ is the idiosyncratic disturbance term only relevant to measure i.

To identify the critical role of change in comprehensive media climate change concern in the stock market, we aim to efficiently estimate $\Delta CMCCC_t^*$. One essential issue here is to impose the factor structure (6) on individual measures to estimate $\Delta CMCCC_t^*$, and in the meanwhile, to remove the idiosyncratic noise $e_{i,t}$ and the common approximation error $Error_t$ from the estimation process. To do so, we employ one common aggregation method: Partial Least Square (PLS). To avoid the estimation errors and show the robustness of our index measure, we further incorporate two alternative aggregation approaches: equal-weight and volatility-weight.

3.2. Partial Least Square (PLS)

The PLS approach extracts $\Delta CMCCC_t^*$ from the individual media climate change concern measures based on its covariance with future stock market returns. It applies a linear combination of the individual media climate change concern measures to predict returns and consists of three steps.

The first step is a time-series regression of each individual media climate change concern at month t on the realized subsequent excess stock return (as a proxy for expected excess returns), r_{t+1},

$$\Delta MCCC_{i,t-1} = \pi_0 + \pi_i R_t + u_{i,t-1} \tag{7}$$

where $\Delta MCCC_{i,t-1}$ is the changes in media climate concern from source i. The coefficient of π_i in the first-step time-series regression (7) captures the sensitivity of the individual media climate change concern measure $\Delta MCCC_{i,t}$ to the unobservable comprehensive media climate change concern $\Delta CMCCC_t^*$ instrumented by future stock market return R_t. Because the future stock market return R_t is driven by $\Delta CMCCC_t^*$, as presented in model (5), individual media climate change concern measures are associated with the predictable component of stock returns and have no association with the unforecastable errors. Therefore, the coefficient π_i approximately represents how each individual measures depends on the unobservable comprehensive media climate change concern $\Delta CMCCC_t^*$.

In the second step, we run a cross-sectional regression of $\Delta MCCC_{i,t}$ on $\hat{\pi}_i$ for each month t:

$$\Delta MCCC_{i,t} = c_t + \Delta CMCCC_t^{PLS} \hat{\pi}_i + v_{i,t} \tag{8}$$

where $\hat{\pi}_i$ is the regression loading in regression (7) and $\Delta CMCCC_t^{PLS}$, the regression slope, is the PLS media climate change concern at time t. In the regression (8), the first step loading estimated is the independent variable while $\Delta CMCCC_t^{PLS}$ refers to the regression slope needs to be estimated.

The factor nature of a joint system consisting of Equations (5) and (6) has been exploited through PLS to infer the relevant comprehensive climate change concern index $\Delta CMCCC_t^{PLS}$. The $\Delta CMCCC_t^{PLS}$ can be consistently estimated when the true factor loading π_i were known by simply applying the cross-sectional regressions of $\Delta MCCC_{i,t}$ on π_i month by month. However, given that π_i is unknown, an approximate estimation of how $\Delta MCCC_{i,t}$ relies on $\Delta CMCCC_t^{PLS}$ is provided by the first-stage regression slopes. In other words, the dimension reduction is disciplined by the PLS method to extract $\Delta CMCCC_t^*$ which is related with prediction via future market returns. Meanwhile, the idiosyncratic and common components, such as $Error_t$ and $\varepsilon_{i,t}$ which are irrelevant with the prediction, are eliminated based on the PLS method.

As discussed in Kelly and Pruitt (2015), since the individual proxies could be measured with noise, an error-in-variables form might be taken by the first-stage regression while an estimate is produced in the second stage for a unique but unknown rotation of the latent factor $\Delta CMCCC_t^*$. Nevertheless, because the common component of individual proxies spans the relevant factor space, the predictive regression of realized returns on the estimated PLS factor can forecast expected returns driven by the latent factor consistently.

In the empirical implementation, we use the full sample data from January 2002 to September 2021 to estimate the PLS concern index. Specifically, in the time-series regression (4), we estimate the loadings (π_i) for $\Delta MCCC_{WP}$, $\Delta MCCC_{WSJ}$, $\Delta MCCC_{NYT}$, $\Delta MCCC_{USAT}$, $\Delta MCCC_{LAT}$, $\Delta MCCC_{ABC}$, $\Delta MCCC_{CBS}$, $\Delta MCCC_{CNN}$, $\Delta MCCC_{FOX}$, $\Delta MCCC_{MSNBC}$, $\Delta MCCC_{NBC}$, $\Delta MCCC_{PBS}$ and $\Delta MCCC_{radio}$ from January 2002 to September 2021, $\Delta MCCC_{AP}$ and $\Delta MCCC_{UPI}$ from January 2006 to September 2021. In the second step, we run the cross-sectional regression (5) for each time t from January 2002 to September 2021 and estimate the $\Delta CMCCC_t^{PLS}$ based on the available loadings π_i for each month. Consequently, we obtain monthly PLS-based comprehensive $\Delta CMCCC_t^{PLS}$ from January 2002 to September 2021.

In the third step, we run the following predictive regression to investigate its in-sample return predictability:

$$R_{t+1} = \alpha + \beta \times \Delta CMCCC_t^{PLS} + \varepsilon_{t+1} \tag{9}$$

For out-of-sample analysis, the standard approach is to repeat these three steps by truncating the unknown observations at month $t+1$. Specifically, suppose we need to forecast return at month $t+1$, we could only rely on the information known through month t. In the first step, the latest return that we can use on the right-hand side of regression (7) is R_t and the last observation of individual concern measures on the left-hand side is $\Delta MCCC_{i,t-1}$. In the second step, we run the cross-sectional regressions for months 1 through t. In the third step, the latest return on the left-hand side of regression (9) entering the predictive regression is R_t and the forecast for R_{t+1} is $\hat{\alpha}_t + \hat{\beta}_t \times \Delta CMCCC_t^{PLS}$, where $\hat{\alpha}_t$ and $\hat{\beta}_t$ are the estimates using information up to month t. In summary, for out-of-sample forecasting, we construct all inputs to the forecast using data observed no later than month t. Moreover, we impose an economic restriction in predicting stock returns, that the expected risk premium must be positive to be consistent with theory of Campbell and Thompson (2008) and Pettenuzzo et al. (2014).

For comparison, we further consider two alternative aggregation media climate change concern measures as benchmarks: $\Delta CMCCC_t^{Equ}$ and $\Delta CMCCC_t^{Vol}$. The equal-weighted index simply treats 15 individual $\Delta MCCC_{i,t}$ (scaled by standard deviation) equally and the volatility-weight index uses the reciprocal of the variance of each $\Delta MCCC_{i,t}$ as the combination weight.

[Insert Figure 1 about here]

Figure 1 displayed the time series of the three media climate change concern index, $\Delta CMCCC_t^{PLS}$, from January 2002 to September 2021. We observe that the concern index is time varying for our sample from January 2002 to September 2021. In general, it spikes during salient climate events, such as 2009 Copenhagen UN Climate change Conference on December 2009, and adoption of Paris Agreement in 2015. This evidence is consistent with the findings (e.g., Figure 2) in Engle et al. (2020) who have documented the increase of climate change news risk during some salient climate events based on news content from Wall Street Journal.

3.3. Forecasting performance

This section explores the in- and out-of-sample forecasting performance of three comprehensive media climate change concern indexes, which include $\Delta CMCCC_t^{PLS}$, $\Delta CMCCC_t^{Equ}$ and $\Delta CMCCC_t^{Vol}$.

[Insert Table 4 about here]

Table 4 presents the regression slope β(%), Newey–West *t*-value, in-sample R^2(%), and out-of-sample R_{os}^2(%) of predicting market returns with individual media climate change concern measures over one- to twelve-month horizon. Panel A of table 4 reports the forecasting results for $\Delta CMCCC^{PLS}$. Specifically, we observe that $\Delta CMCCC^{PLS}$ significantly and negatively predicts market excess returns both in-sample and out-of-sample and its predictability persists up to one year. For instance, at one-month horizon, the β estimate is -0.68% with a t-statistic of -2.99 based on Newey-West standard error. Since we standardize all predictors to have zero mean and unit variance for analysis, our result for the monthly horizon implies that a one-standard deviation increase in media climate change concern leads to a 0.68% decrease in next one-month market return with the $\Delta CMCCC^{PLS}$ index. For longer horizons, the β estimate keeps negative and significant with t-statistic shrinks to -2.07 in magnitude at the one-year horizon.

When it comes to out-of-sample performance, the R_{os}^2 with PLS comprehensive media climate change concern index is 2.86% and significant at the 5% level for monthly horizon. And it increases to 10.07% at the annual horizon. Such R_{os}^2 is also economically sizable since in Campbell and Thompson (2008) paper, they have shown that a monthly out-of-sample R_{os}^2 of 0.5% can generate significant economic value. Thus, the out-of-sample R_{os}^2 with PLS comprehensive media climate change concern which are much larger than 0.5% indicate the substantial economic significance.

Panel B and Panel C of table 4 reports the similar results of forecasting market return with $\Delta CMCCC_t^{Equ}$ and $\Delta CMCCC_t^{Vol}$, respectively. Both results are consistent with the PLS aggregation index. For media climate change concern by equal-weighted combination method, one-standard deviation increase in media climate change concern leads to a 0.57% decrease in next one-month market return with the $\Delta CMCCC^{Equ}$ index. It also delivers economically sizable out-of-sample R_{os}^2 of 2.19% at monthly horizon. Moreover, the magnitudes of in-sample R^2 and out-of-sample R_{os}^2 are both significantly amplified over longer horizon. For media climate change concern by volatility-weighted combination method, the in-sample β estimate is -0.56% with a t-statistic of -2.56 based on Newey-West standard error at one-month horizon. In addition, both in-sample R^2 and out-of-sample R_{os}^2 are significant from one- to 12-month horizons.

In general, the three measures, namely, the PLS, the volatility-weighted and the equal weighted aggregation indices all perform well both in- and -out- of sample. Overall, the empirical evidence has proven that the forecasting power of comprehensive media climate change concern on market return over one- to 12-month horizon.

3.4. Comparison with Economic Predictors

In this subsection, we further compare the forecasting power of PLS comprehensive media climate change concern $\Delta CMCCC^{PLS}$ with economic predictors and examine whether its forecasting power on market return remains significant after controlling for extant economic predictors. We consider the 14 economic predictors in Goyal and Welch (2008).[5] First, we run a univariate regression on a single economic predictor as:

$$R_{t+1} = \alpha + \psi Z_t^k + \varepsilon_{t+1} \tag{10}$$

where R_{t+1} denotes the monthly excess market return (%) and Z_t^k is one of the 14 economic predictors described in Appendix A. We also consider the first principal component of these 14 economic predictors as the independent variable.

[Insert Table 5 about here]

Panel A of Table 5 present the in-sample regression results for Equation (10). Among 14 economic predictors, only Long-term bond yield (LTY) exhibit significant predictive ability for market return at the 5% or better significance level during January 2002 to September 2021. Although its in-sample R^2 is 3.18% which is larger than $\Delta CMCCC^{PLS}$ (2.53%), its out-of-sample R_{OS}^2 (%) 1.06% is much smaller than $\Delta CMCCC^{PLS}$ (2.86%). Thus, $\Delta CMCCC^{PLS}$ outperforms 14 economic predictors in forecasting the monthly market returns.

Then we run a bivariate regression to explore the forecasting power of $\Delta CMCCC^{PLS}$ after controlling for one of the economic predictors as:

$$R_{t+1} = \alpha + \psi Z_t^k + \beta \Delta CMCCC_t^{PLS} + \varepsilon_{t+1} \tag{11}$$

Panel B of Table 5 present the in-sample regression results for Equation (11). After controlling economic variables, $\Delta CMCCC^{PLS}$ remain negative and statistically significant during the sample period. For example, when controlling for dividend-price ratio, the regression slope of $\Delta CMCCC^{PLS}$ slightly decreases to -0.64 in absolute value and is significant at the 1% level. In addition, the magnitude of β estimates of $\Delta CMCCC^{PLS}$ remains around -0.60%, which is almost

[5] The data is available from Amit Goyal's website, http://www.hec.unil.ch/agoyal/.

the same as without controlling for the economic variables. In the last row of panel B, we report the estimates results after controlling for the first principal component of these 14 economic variables. The regression slope of $\Delta CMCCC^{PLS}$ is -0.67% and significant at 1% level. Overall, the PLS comprehensive media climate change concern generates significant predictive ability for the stock market return beyond economic predictors.

3.5. Comparison with Sentiment and Climate-related concern variables

In this subsection, we compare the forecasting power of PLS comprehensive media climate change concern $\Delta CMCCC^{PLS}$ with sentiment measures and alternative media climate change concern measures.

Specifically, for sentiment measures, we employ four indices: Baker and Wurgler (2006) investor sentiment index ($SENT^{BW}$) which has been constructed based on five sentiment proxies; aligned investor sentiment index ($SENT^{PLS}$) by Huang et al. (2015) ; average article sentiment ($SENT^{news}$) and sentiment on market topic ($SENT^{mkt\_topic}$) from Calomiris and Mamaysky (2019) which are constructed based on Thomson Reuters News Analytics (TRNA) data[6]. For climate-related concern, we select two measures from Bybee et al., (2020) which are constructed based on a topic modelling of Wall Street Journal articles content: *Natural Disaster Attention* and *Environmental Attention*. [7] Existing research also rely on the media coverage data to construct climate-related concern or risk, although most of them are using data from one specific channel, i.e., newspapers. Thus, we also consider four alternative measures in existing literatures: the first two variables are from Engle, Giglio, Kelly, Lee and Stroebel (2020) which are climate change news index based on Wall Street Journal and CH Negative Climate Change News Index based on massive media coverage data provided by data analytics vendor Crimson Hexagon. We follow their approach to use AR (1) innovations of these two indices in predictive regression and we denote them as EGKLS^{wsj} and EGKLS$^{\text{chneg}}$.[8] They are

[6] The data of Baker and Wurgler (2006) investor sentiment index is available from Jeffrey Wurgler's website, http://people.stern.nyu.edu/jwurgler/. The data of aligned investor sentiment index by Huang et al. (2015) is available from Guofu Zhou's website, http://apps.olin.wustl.edu/faculty/zhou/. The data of average article sentiment and sentiment on market topic from Calomiris and Mamaysky (2019) are available from Harry Mamaysky's website, https://sites.google.com/view/hmamaysky/home?authuser=0 .

[7] The data of two climate-related attention measures from Bybee et al., (2020) is available at http://www.structureofnews.com/ .

[8] The data is available at Stefano Giglio's website, https://sites.google.com/view/stefanogiglio/ .

available from February 1984 to June 2017 and from July 2008 to May 2018, respectively. The next one alternative measure is Media Climate Change Concern constructed by Ardia, Bluteau, Boudt and Inghelbrecht (2021) based on eight U.S. major newspapers.[9] We follow their method to use the AR (1) innovations of their overall index in predictive regression and we denote it as ABBI^{overall} which is available from February 2003 to June 2018. The last alternative is constructed by Pastor, Stambaugh and Taylor (2021) based on Ardia et al., (2021)' Media Climate Change Concern overall index. We follow their method to calculate the level of climate concern based on a memory decay model and then take change as final predictor. We denote it as PST^{decay_change} which is available from February 2006 to June 2018.

We first analyze the forecasting power of these sentiment and climate-related concern predictors by running the univariate regression with a single predictor among them as the independent variable. We next analyze the incremental forecasting power of PLS comprehensive media climate change concern after controlling for sentiment or climate-related attention as following:

$$R_{t+1} = \alpha + \psi D_t^k + \beta \Delta CMCCC^{PLS} + \varepsilon_{t+1} \tag{12}$$

where R_{t+1} denotes the monthly excess market return (%), D_t^k is one of the sentiment measures or climate-related concern measures.

[Insert Table 6 about here]

Panel A of table 6 present the in-sample regression results for sentiment predictors. The sample period for all the variables starts from January 2002 and ends at the data available time. According to the second column in panel A, only $SENT^{BW}$ exhibits a significant predictive power during the sample period. After controlling for $SENT^{BW}$, the regression slope on $CMCCC^{PLS}$ is still negative and significant at 10% level. Panel B of table 6 reports the in-sample regression results for climate-related concern predictors. According to the univariate regression results, none of them have the predictive ability for market return. Overall, $\Delta CMCCC^{PLS}$ remains negative and statistically significant when augmented by other predictors.

[9] The data is available at https://sentometrics-research.com/post/climate-change/ .

These results illustrate that $\Delta CMCCC^{PLS}$ contains sizeable complementary forecasting information beyond other sentiment or climate-related attention predictors.

3.6. Asset Allocation Analysis

In this subsection, we explore the economic value of the forecasting stock market returns with comprehensive media climate change concern measures from investment perspective. Following Kandel and Stambaugh (1996), Campbell and Thompson (2008), and Ferreira and Santa-Clara (2011), we calculate the certainty equivalent return (CER) gain and Sharp ratio for a mean-variance investor who optimally allocates her wealth between the stock market and one-month T-bill based on the out-of-sample forecast. The higher the CER gain and Sharpe ratio, the larger the risk-rewarded returns when using $\Delta CMCCC$ index.

At the beginning of each month, the mean-variance investor allocates a proportion of w_t to the stock market to maximize her next one-month expected utility.

$$U(R_p) = E(R_p) - \frac{\gamma}{2} Var(R_p) \tag{14}$$

where R_p is the return of the investor's portfolio, $E(R_p)$ and $Var(R_p)$ represent the mean and the variance of the market returns and γ is investor's risk aversion.

Let R_{t+1} and $R_{f,t+1}$ represent the market return and T-bill rate. The portfolio return of the investor at the end of each month is

$$R_{p,t+1} = w_t R_{t+1} + R_{f,t+1} \tag{15}$$

where $R_{f,t+1}$ is known at time t. With a simple calculation, the optimal portfolio weight is

$$w_t = \frac{1}{\gamma} \frac{\hat{R}_{t+1}}{\hat{\sigma}_{t+1}^2} \tag{16}$$

where $\hat{R}_{t+1}$ and $\hat{\sigma}_{t+1}^2$ are investor's estimates on the mean and variance of market returns based on information up to time t.

Then, the CER of portfolio is computed as

$$CER = \hat{\mu}_p - \frac{\gamma}{2}\hat{\sigma}_p^2 \quad (17)$$

where $\hat{\mu}_p$ and $\hat{\sigma}_p^2$ represent the mean and variance of investor's portfolio throughout the out-of-sample evaluation period. The CER can be interpreted as the compensation to investor for having the stock market. The CER gain is the difference between the CER for the investor using the predictive regression based on $\Delta CMCCC$ index and based on historical return mean. We multiply this difference by 1200 to get the annualized percentage CER gain which can be interpreted as the annual spend that an investor would be willing to pay to have access to the predictive regression forecast instead of the historical mean forecast. To be consistent, we also calculate the annualized Sharpe ratio of investor's portfolio.

[Insert Table 7 about here]

Table 8 reports the portfolio gains of a mean-variance investor trading on our comprehensive media climate change concern indices using PLS, equal weight and volatility weight aggregation methods. We consider risk aversion coefficients of 1, 3, and 5, respectively. Additionally, we also incorporate the case of 50bps transaction costs to check the robustness of our asset allocation results. We follow Campbell and Thompson (2008) to use a ten-year moving window of past monthly returns to estimate the variance of market returns, and constraints w_t to lie between 0 and 1.5 to exclude extreme cases.

In panel A, when risk aversion $\gamma = 1$ and there is no transaction cost, the annualized CER gain by using the PLS (equal weight, volatility weight) media climate change concern index is 4.39% (5.10%, 5.29%), suggesting that investing with the PLS (equal weight, volatility weight) index forecast can generate a 4.39% (5.10%, 5.29%) greater risk-adjusted return relative to the historical return mean. The corresponding annualized Sharp ratio is 1.20 (1.20, 1.22) which is much higher than the Sharp ratio 0.95% based on historical return mean forecast in our sample period. When there is a transaction cost of 50 basis points, the CER gain by using the PLS (equal weight, volatility weight) is 2.86% (3.71%, 4.03%), which is still economically sizeable. The corresponding Sharpe ratio is 1.09 (1.10, 1.12). Panel B and Panel C show the similar results when the investor risk-aversion $\gamma = 3$ or $\gamma = 5$. For example, the net-of-transactions-costs

CER gains for $\Delta CMCCC^{PLS}$ ($\Delta CMCCC^{Equ}$, $\Delta CMCCC^{Vol}$) is 3.55% (2.84%, 2.85%) when risk-aversion $\gamma = 5$. And it decreases into 1.78% (1.27%, 1.28%) with a transaction cost of 50 basis points which is still considered as economically sizeable. Overall, the comprehensive media climate change concern using PLS, equal-weight and volatility-weight aggregation methods could generate substantial economic value for a mean-variance investor.

3.8. Alternative econometric methods

In the prior sections, we have shown that the PLS comprehensive media climate change concern could significantly predict market returns over one- to 12-month horizon. In this subsection, we further examine whether the result is robust to alternative econometric approaches. Specifically, we consider three alternative methods: Simple Combination (Rapach et al.,2010), elastic net (Kozak et al., 2020) and Lasso.

[Insert Table 8 about here]

The results have been shown in Table 9. For these three alternative methods, the out-of-sample R^2_{os}(%) s are all significant at the one- to 12-month horizons, and also economic sizable using Campbell and Thompson (2008) threshold of 5%. For example, with the elastic net method, the R^2_{os} ranges from 1.83% to 15.11% for one- to 12 months horizon, and statistically significant. Such results confirm that the predictability of the PLS comprehensive media climate change concern on the market returns. However, the out-of-sample performance of these three alternative methods all underperform the PLS results. This finding is consistent with Kelly and Pruitt (2015)'s argument that the PLS forecast is asymptotically consistent and will generate the minimum MSFE so long as the consistency condition is satisfied.

4. Economic Explanation

4.1. Cash flow and discount rate predictability

The section examines the economic underpinnings of the comprehensive media climate change concern's predictive power, whether it is from the discount rate channel or the cash flow channel or both. Specifically, we measure the news component using the VAR methodology which is developed by Campbell (1991) and Campbell and Ammer (1993).

Based on Campbell (1991), the total market return can be decomposed into three parts:

$$R_{t+1} = E_t(R_{t+1}) + \eta_{t+1}^{CF} - \eta_{t+1}^{DR} \tag{18}$$

where $E_t(R_{t+1})$ is the expected return, η_{t+1}^{CF} is the cash flow news and η_{t+1}^{DR} is the discount rate news. Following Cochrane (2011), when running the three components of (18) on the comprehensive media climate change concern index,

$$E_t(R_{t+1}) = \alpha + \beta_E \Delta CMCCC_t^{PLS} + \varepsilon_{t+1}^{E} \tag{19}$$

$$\eta_{t+1}^{CF} = \beta_{CF} \Delta CMCCC_t^{PLS} + \varepsilon_{t+1}^{CF} \tag{20}$$

$$\eta_{t+1}^{DR} = \beta_{DR} \Delta CMCCC_t^{PLS} + \varepsilon_{t+1}^{DR} \tag{21}$$

one can obtain:

$$\beta = \beta_E + \beta_{CF} - \beta_{DR} \tag{22}$$

where β is the regression slope in Equation (9). Then, by comparing the estimated regression slopes through Equation (19) to (21), we can verify the extent to which the comprehensive media climate change concern's capacity to forecast the total market returns relates with its capacity to forecast the latter two components in Equation (18).

Table 9 reports the estimation results of $\hat{\beta}_{CF}$ and $\hat{\beta}_{DR}$. The cash flow news and discount rate news are estimated based on individual VARs comprised of the S&P 500 log return, log dividend-price ratio, and one of the 13 popular predictors from Goyal and Welch (2008). We follow Engsted, Pedersen, and Tanggaard (2012) to incorporate the log dividend-price ratio all the time since they argue that dividend-price ratio is important for cash flow and discount rate components estimation properly in VAR decomposition method. Additionally, we further include the first principal components extracted from the 14 economic predictors, the dividend-price ratio and the total market return in the VAR decomposition model and the results have been shown in the last row of Table 9.

[Insert Table 9 about here]

For the sample period from January 2002 to September 2021, nearly all the $\hat{\beta}_{CF}$ estimates are significant although it becomes weak when the return decomposition based on Net equity expansion (NTIS) or Treasury bill rate (TBL). However, for discount rate news, the comprehensive media climate change concern is insignificantly related to it. According to Cochrane (2011), a component that forecasts the market return must estimate its discount rate news, its cash flow news or both. Our results in Table 9 confirms this argument that the ability of comprehensive media climate change concern in predicting market returns tend to operate through cash flow channels.

4.2. The cross-section predictability for green and brown stocks

To further understand the forecasting source of our comprehensive media climate change concern, we adopt Pastor et al., (2021)'s method to test the cross-section predictability of $\Delta CMCCC$ on green and brown stocks, respectively.

Specifically, we employ the environmental scores data from Morningstar Sustainalytics during August 2009 to October 2018 to measure firm-level greenness and we merge it with the return data from CRSP with CRSP share codes of 10 or 11. Then we sort the firm based on its greenness score at the end of month t-1 and obtain the green (brown) stocks portfolio with the greenness score in the top (bottom) third of all firms on month t. We focus on the value-weighted return of green (brown) stock portfolio and obtain the GMB (green-minus-brown) value-weighted return by taking the difference between green and brown portfolios.

Table 10 reports the estimation results of cross-section predictability on green, brown and GMB portfolio returns. For both green and brown portfolios, our comprehensive media climate change concern tends to negatively affect its return with one-month delay, especially for brown portfolios (the t-value is around 1.50). For GMB (green-minus-brown) portfolio, we find that our comprehensive media climate change concern significantly positively predicts it with one-month delay which is consistent with the existing literature that stock prices are underreacting to climate change risk and/or concern (e.g., Hong et al., (2019), Pastor et al., (2021)). Moreover, the delayed reaction seems mainly driven by brown stocks rather than green stocks. This is also consistent Pastor et al., (2021), which points out that brown stocks are usually smaller than green stocks and hence have more underreaction to the shock than green stocks.

Such underreaction to climate change concern seems supporting the cash flow channel to be the main driver of our time-series predictability. That is, when climate change concern increases, the future cash flows of firms, especially those brown firms, could be negatively affected due to many reasons, such as tighter environmental regulations. However, there could be some underreaction from investors due to market frictions (e.g., limited investor attentions). When the negative impact on the cash flow incorporates into stock prices gradually, the delay in price reaction then leads to return predictability.

[Insert Table 10 about here]

5. Conclusion

In this paper, we study the time-series pricing effects of media climate change concern on the aggregate stock market. We propose a novel comprehensive media climate change concern index by measuring the overall media concern of climate change based on the coverage from various types of media sources including newspaper, television, radio, and wire service. We find that change in comprehensive media climate change concern predicts subsequent stock market return negatively and significantly, and this pattern holds at multiple horizons from one-month to one-year. In contrast, individual media climate change concern measures based on single media outlet have limited return predictability. The predictive power of our comprehensive concern index is still present after controlling for common market return predictors. Moreover, the predictability exists out-of-sample and delivers sizable economic value for mean-variance investors in asset allocation. Although our main index is extracted by using the partial least squares, the results are similar by using alternative machine learning methods.

In contrast to prior studies on the media concern about climate change that either focuses on Wall Street Journal or covers several newspapers, our comprehensive index is novel in that it can capture audio and visual sensations about climate change concern. In addition, our study focuses on the impact of media climate change concern on the aggregate stock market, which is largely ignored by the existing studies that mainly focus on cross-section predictability of climate change concerns.

Reference

Ardia, D., Bluteau, K., Boudt, K., and Inghelbrecht, K., 2021. Climate change concerns and the performance of green versus brown stocks. Working paper.

Baker, M., and Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. Journal of Finance 61, 1645–1680.

Baker, M., Bergstresser, D., Serafeim, G. and Wurgler, J., 2018. Financing the response to climate change: The pricing and ownership of US green bonds . Working paper.

Bolton, P. and Kacperczyk, M.T., 2020. Carbon premium around the world. Working paper.

Bolton, P., and Kacperczyk, M., 2020. Do investors care about carbon risk?. Journal of Financial Economics, forthcoming.

Bybee, L., Kelly, B.T., Manela, A. and Xiu, D., 2020.The Structure of Economic News,Working paper

Calomiris, C.W. and Mamaysky, H., 2019. How news and its context drive risk and returns around the world. Journal of Financial Economics, 133(2), pp.299-336.

Campbell, J. Y., 1991. A variance decomposition for stock returns. Economic Journal 101, 157–179.

Campbell, J., and Thompson, S., 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? Review of Financial Studies 21, 1509–1531.

Campbell, J.Y., and Ammer, J., 1993. What moves the stock and bond markets? a variance decomposition for long-term asset returns. Journal of Finance 48, 3–37.

Choi, D., Gao, Z. and Jiang, W., 2020. Attention to global warming. The Review of Financial Studies, 33(3), pp.1112-1145.

Clark, T. E., and West, K. D., 2007. Approximately normal tests for equal predictive accuracy in nested models. Journal of Econometrics 138, 291–311.

Cochrane, J. H., 2008. The dog that did not bark: A defense of return predictability. Review of Financial Studies 21, 1533–1575.

Cochrane, J. H., 2011. Presidential address: Discount rates. Journal of Finance 66, 1047–1108.

Edmans, A., Fernandez-Perez, A., Garel, A., and Indriawan, I., 2021. Music sentiment and stock returns around the world. Journal of Financial Economics, forthcoming.

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

Engsted, T., Pedersen, T. Q., and Tanggaard, C., 2012. Pitfalls in VAR based return decompositions: A clarification. Journal of Banking & Finance 36, 1255–1265.

Ferreira, M. A., and Santa-Clara, P., 2011. Forecasting stock market returns: The sum of the parts is more than the whole. Journal of Financial Economics 100, 514–537.

Giglio, S., Kelly, B. and Stroebel, J., 2020. Climate finance. Working paper.

Giglio, S., Maggiori, M., Rao, K., Stroebel, J. and Weber, A., 2021. Climate change and long-run discount rates: Evidence from real estate. The Review of Financial Studies, 34(8), pp.3527-3571.

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

Gu, S., Kelly, B., and Xiu, D., 2020. Empirical asset pricing via machine learning. Review of Financial Studies 33, 2223–2273.

Hong, H., Karolyi, G.A. and Scheinkman, J.A., 2020. Climate finance. The Review of Financial Studies, 33(3), pp.1011-1023.

Hong, H., Li, F.W. and Xu, J., 2019. Climate risks and market efficiency. Journal of econometrics, 208(1), pp.265-281.

Hsu, P.H., Li, K. and Tsou, C.Y., 2020. The pollution premium. Working paper.

Huang, D., Jiang, F., Tu, J., and Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. Review of Financial Studies 28, 791–837.

Ilhan, E., Krueger, P., Sautner, Z. and Starks, L.T., 2020. Climate risk disclosure and institutional investors. Working Paper, (19-66).

Kandel, S., and Stambaugh, R., 1996. On the predictability of stock returns: An asset allocation perspective. Journal of Finance 51, 385–424.

Kelly, B., and Pruitt, S., 2013. Market expectations in the cross-section of present values. Journal of Finance 68, 1721–1756.

Kelly, B., and Pruitt, S., 2015. The three-pass regression filter: A new approach to forecasting using many predictors. Journal of Econometrics 186, 294–316.

Kozak, S., Nagel, S., and Santosh, S., 2020. Shrinking the cross-section. Journal of Financial Economics, 135(2), 271-292.

Mele, A., 2007. Asymmetric stock market volatility and the cyclical behavior of expected returns. Journal of financial economics, 86(2), pp.446-478.

Murfin, J. and Spiegel, M., 2020. Is the risk of sea level rise capitalized in residential real estate?. The Review of Financial Studies, 33(3), pp.1217-1255.

Newey, W., and West, K., 1987. A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Obaid, K. and Pukthuanthong, K., 2021. A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. Journal of Financial Economics.

Pastor, L., Stambaugh, R. F., and Taylor, L. A. ,2021. Dissecting Green Returns. Working paper.

Pastor, L., Stambaugh, R. F., and Taylor, L. A.,2020. Sustainable investing in equilibrium. Journal of Financial Economics.

Pettenuzzo, D., Timmermann, A., and Valkanov, R., 2014. Forecasting stock returns under economic constraints. Journal of Financial Economics 114, 517–553.

Rapach, D. E., Strauss, J. K., and Zhou, G., 2013. International stock return predictability: What is the role of the United States? Journal of Finance 68, 1633–1662.

Rapach, D., Ringgenberg, M., and Zhou, G., 2016. Short interest and aggregate stock returns. Journal of Financial Economics 121, 46–65.

Rapach, D.E., Strauss, J.K., and Zhou, G., 2010. Out-of-sample equity premium prediction: combination forecasts and links to the real economy. Review of Financial Studies 23, 821–862.

Figure 1. Time series of Comprehensive Media Climate Change Concern index

This figure plots the time series dynamics of Comprehensive Media Climate Change Concern (i.e.,$\Delta CMCCC$) index constructed using PLS aggregation methods. The index is standardized. The sample period is 2002:01-2021:09.

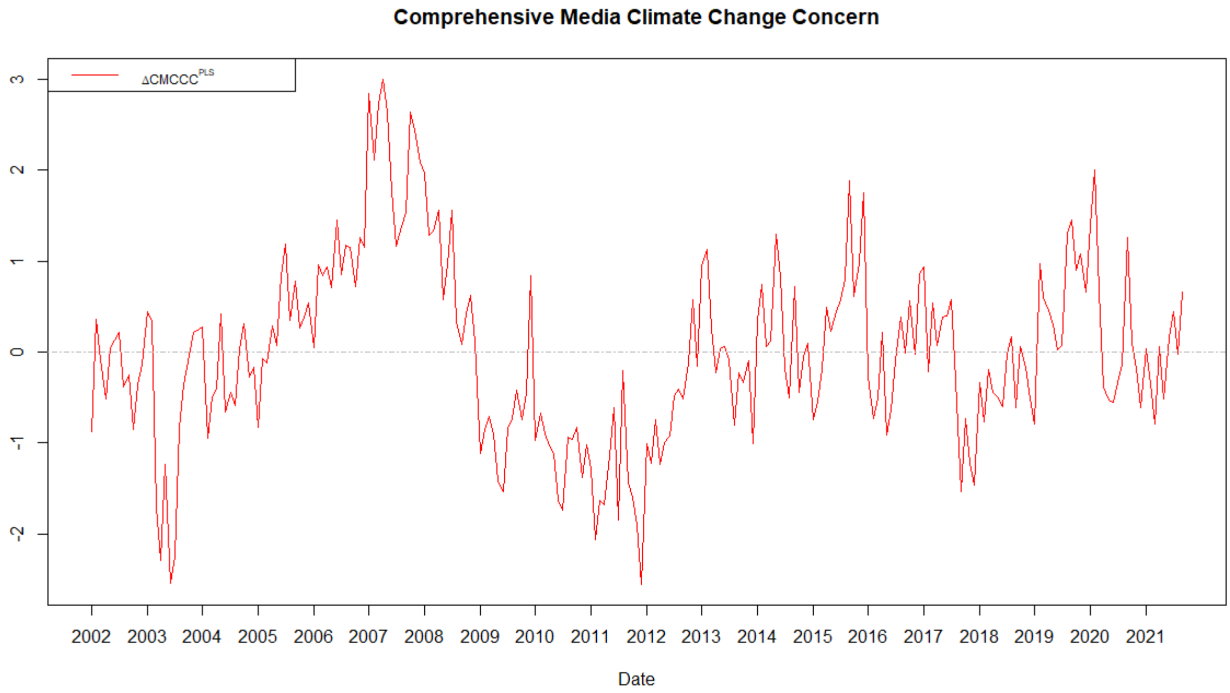

Table 1: Summary statistics of individual media climate change concern

This table reports the median, 25% and 75% quartiles, skewness, and first-order autocorrelation coefficient ($\rho(1)$) of 15 individual media climate change concern used in this paper. The first 5 individual concern measures are from newspapers: Washington Post, Wall Street Journal, New York Times, USA Today and Los Angeles Times; the next 7 individual concern measures are from televisions: ABC, CBS, CNN, FOX, MSNBC, NBC and PBS and another one is from National Public Radio (US). All the 13 variables are available between January 2002 and September 2021. The last 2 individual concern measures are from wire services: Associated Press and United Press International and they are available from January 2006 and September 2021. All individual media climate change concern variables are standardized to have mean of 0 and variance of 1.

	sample period	Q25	median	Q75	skew	ρ (1)
$\Delta MCCC_{WP}$	200201-202109	-0.62	-0.06	0.64	0.23	0.70
$\Delta MCCC_{WSJ}$	200201-202109	-0.68	0.08	0.66	-0.10	0.56
$\Delta MCCC_{NYT}$	200201-202109	-0.63	-0.10	0.41	0.88	0.76
$\Delta MCCC_{USAT}$	200201-202109	-0.64	0.00	0.60	0.33	0.53
$\Delta MCCC_{LAT}$	200201-202109	-0.69	0.06	0.54	0.22	0.70
$\Delta MCCC_{ABC}$	200201-202109	-0.50	-0.04	0.56	0.16	0.56
$\Delta MCCC_{CBS}$	200201-202109	-0.70	-0.08	0.65	0.17	0.57
$\Delta MCCC_{CNN}$	200201-202109	-0.73	-0.02	0.70	0.02	0.61
$\Delta MCCC_{FOX}$	200201-202109	-0.67	0.01	0.63	0.18	0.63
$\Delta MCCC_{MSNBC}$	200201-202109	-0.60	-0.03	0.66	0.18	0.56
$\Delta MCCC_{NBC}$	200201-202109	-0.66	-0.12	0.62	0.44	0.53
$\Delta MCCC_{PBS}$	200201-202109	-0.68	-0.18	0.56	0.11	0.30
$\Delta MCCC_{radio}$	200201-202109	-0.61	-0.01	0.58	0.02	0.64
$\Delta MCCC_{AP}$	200601-202109	-0.71	-0.06	0.45	0.56	0.87
$\Delta MCCC_{UPI}$	200601-202109	-0.62	-0.01	0.67	-0.06	0.64

Table 2: Correlation of individual media climate change concern change concern

This table shows the pairwise correlation of the 15 individual media climate change concern measures: $\Delta MCCC_{WP}$,$\Delta MCCC_{WSJ}$, $\Delta MCCC_{NYT}$, $\Delta MCCC_{USAT}$, $\Delta MCCC_{LAT}$, $\Delta MCCC_{ABC}$, $\Delta MCCC_{CBS}$, $\Delta MCCC_{CNN}$, $\Delta MCCC_{FOX}$, $\Delta MCCC_{MSNBC}$,$\Delta MCCC_{NBC}$,$\Delta MCCC_{PBS}$ and $\Delta MCCC_{radio}$ from January 2002 to September 2021, $\Delta MCCC_{AP}$ and $\Delta MCCC_{UPI}$from January 2006 to September 2021. All concern variables are standardized to have mean of 0 and variance of 1.

	$\Delta MCCC_{W}$	$\Delta MCCC_{WS}$	$\Delta MCCC_{NY}$	$\Delta MCCC_{US}$	$\Delta MCCC_{LA}$	$\Delta MCCC_{AB}$	$\Delta MCCC_{CB}$	$\Delta MCCC_{CN}$	$\Delta MCCC_{FO}$	$\Delta MCCC_{MSN}$	$\Delta MCCC_{NB}$	$\Delta MCCC_{PB}$	$\Delta MCCC_{rad}$	$\Delta MCCC_{A}$
$\Delta MCCC_{WP}$	1.00													
$\Delta MCCC_{WSJ}$	0.55	1.00												
$\Delta MCCC_{NYT}$	0.60	0.25	1.00											
$\Delta MCCC_{USAT}$	0.71	0.49	0.36	1.00										
$\Delta MCCC_{LAT}$	0.75	0.53	0.57	0.64	1.00									
$\Delta MCCC_{ABC}$	0.69	0.43	0.46	0.61	0.64	1.00								
$\Delta MCCC_{CBS}$	0.64	0.44	0.50	0.59	0.60	0.63	1.00							
$\Delta MCCC_{CNN}$	0.71	0.49	0.58	0.58	0.67	0.63	0.56	1.00						
$\Delta MCCC_{FOX}$	0.65	0.57	0.53	0.57	0.58	0.57	0.46	0.61	1.00					
$\Delta MCCC_{MSN}$	0.46	0.33	0.43	0.43	0.39	0.33	0.39	0.43	0.62	1.00				
$\Delta MCCC_{NBC}$	0.61	0.44	0.50	0.54	0.58	0.58	0.54	0.70	0.56	0.45	1.00			
$\Delta MCCC_{PBS}$	0.55	0.42	0.41	0.40	0.37	0.37	0.39	0.51	0.49	0.32	0.46	1.00		
$\Delta MCCC_{radi}$	0.75	0.51	0.59	0.64	0.73	0.59	0.57	0.70	0.60	0.46	0.65	0.50	1.00	
$\Delta MCCC_{AP}$	0.85	0.58	0.61	0.66	0.79	0.63	0.59	0.74	0.64	0.47	0.66	0.51	0.78	1.00
$\Delta MCCC_{UPI}$	0.61	0.40	0.52	0.45	0.48	0.55	0.39	0.57	0.54	0.32	0.46	0.42	0.52	0.58

Table 3: Forecasting market return with individual media climate change concern measures

This table presents the regression slope, Newey–West t-value with 12 lags, in-sample R^2, and out-of-sample R_{os}^2 of predicting market returns with individual media climate change concern measures.

$$R_{t,t+h} = \alpha + \beta X_t + \varepsilon_{t,t+h}$$

where $R_{t,t+h}$ is the cumulative market return (%) between months t and $t + h$ (h = 1, 3, 6 or 12), and X_t is one of the unexpected changes in media climate change concern from 15 sources. The in-sample period for first 13 individual measures is 2002:01–2021:09 and for last two variables is 2006:01-2021:09. The out-of-sample period is 2009:01–2021:09. Statistical significance for R_{os}^2 is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing H_0: $R_{os}^2 \leq 0$ against H_A: $R_{os}^2 > 0$. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

	h=1				h=3				h=12			
	β(%)	t-stat.	R^2(%)	R_{os}^2(%)	β(%)	t-stat.	R^2(%)	R_{os}^2(%)	β(%)	t-stat.	R^2(%)	R_{os}^2(%)
$\Delta MCCC_{WP}$	-0.44 ***	-2.48	1.09	0.70	-1.51 ***	-2.67	3.86	-0.92	-5.51 **	-1.94	12.08	2.57 ***
$\Delta MCCC_{WSJ}$	-0.41	-1.54	0.94	0.46	-1.15	-1.39	2.24	-1.27	-5.87 **	-1.95	13.58	-12.17 ***
$\Delta MCCC_{NYT}$	-0.31	-1.57	0.52	-0.30	-0.98 **	-1.99	1.64	-3.44	-3.15	-1.61	3.96	-21.40 *
$\Delta MCCC_{USAT}$	-0.47 **	-2.04	1.21	1.27 *	-1.24 *	-1.86	2.61	-2.07	-5.10 **	-1.99	10.41	14.94 ***
$\Delta MCCC_{LAT}$	-0.34	-1.55	0.64	0.15	-1.56 ***	-3.15	4.09	4.45 ***	-6.23 ***	-2.46	15.37	5.52 ***
$\Delta MCCC_{ABC}$	-0.35 *	-1.87	0.67	0.00	-0.97	-1.48	1.58	1.67 **	-3.28 **	-2.06	4.21	10.59 ***
$\Delta MCCC_{CBS}$	0.02	0.06	0.00	-1.93	0.01	0.02	0.00	-5.36	-2.27	-1.26	2.05	8.93 ***
$\Delta MCCC_{CNN}$	-0.49 *	-1.72	1.32	0.49	-1.38	-1.50	3.18	-3.58	-3.88	-1.27	5.86	-25.25 *
$\Delta MCCC_{FOX}$	-0.52 **	-2.13	1.48	1.25 *	-1.28 *	-1.72	2.79	0.65 **	-5.65 **	-2.33	12.69	-16.84 **
$\Delta MCCC_{MSNBC}$	-0.61 **	-2.34	2.03	2.22 **	-0.54	-1.12	0.49	-1.30	-2.99	-1.24	3.56	-6.06 *
$\Delta MCCC_{NBC}$	-0.75 ***	-3.05	3.14	4.04 ***	-1.88 ***	-2.80	5.91	0.11 *	-4.37	-1.49	7.33	-3.92 ***
$\Delta MCCC_{PBS}$	-0.34	-1.43	0.63	0.39	-1.79 **	-2.44	5.40	-4.67	-3.03	-1.35	3.66	-9.87
$\Delta MCCC_{radio}$	-0.71 ***	-3.67	2.74	2.31 **	-1.67 ***	-2.81	4.73	-1.02	-6.60 **	-1.98	17.54	-13.99 **
$\Delta MCCC_{AP}$	-0.67 ***	-2.88	2.37	-1.28	-2.30***	-2.83	8.71	-6.63	-8.87 **	-2.23	28.85	-5.67 ***
$\Delta MCCC_{UPI}$	-0.07	-0.28	0.03	-4.90	-1.11	-1.46	2.01	-4.24	-3.75	-1.56	5.07	4.44 ***

Table 4: Forecasting market return with comprehensive media climate change concern

This table presents the regression slope, Newey–West t-value with 12 lags, in-sample R^2, and out-of-sample R^2_{os} of predicting market returns with comprehensive media climate change concern.

$$R_{t,t+h} = \alpha + \beta \Delta CMCCC_t + \epsilon_{t,t+h}$$

where $R_{t,t+h}$ is the cumulative market return (%) between months t and $t + h$ (h = 1, 3, 6, or 12), and $\Delta CMCCC_t$ is the PLS, equal-weight and volatility weight aggregation index. The in-sample period is 2002:01–2021:09 and the out-of-sample period is 2009:01–2021:09. Statistical significance for R^2_{os} is based on the p-value of the Clark and West (2007) MSFE-adjusted statistic for testing $H_0: R^2_{os} \leq 0$ against $H_A: R^2_{os} > 0$. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

	h=1	h=3	h=6	h=12
Panel A: Results for $\Delta CMCCC^{PLS}$				
β(%)	-0.68 ***	-1.82 ***	-3.18 **	-6.36 **
t-stat.	-2.99	-2.53	-2.23	-2.07
R^2(%)	2.53	5.60	7.71	16.26
R^2_{os}(%)	2.86 **	3.19 **	6.29 ***	10.07 ***
Panel B: Results for $\Delta CMCCC^{Equ}$				
β(%)	-0.57 ***	-1.66 **	-3.05 **	-5.97 **
t-stat.	-2.58	-2.44	-2.22	-2.06
R^2(%)	1.77	4.65	7.08	14.36
R^2_{os}(%)	2.19 **	3.58 **	6.41 ***	12.02 ***
Panel C: Results for $\Delta CMCCC^{Vol}$				
β(%)	-0.56 ***	-1.68 **	-3.02 **	-5.86 **
t-stat.	-2.56	-2.44	-2.21	-2.06
R^2(%)	1.74	4.76	6.92	13.81
R^2_{os}(%)	2.13 **	3.58 **	5.84 ***	12.98 ***

Table 5: Comparison with economic predictors

Panel A reports the in-sample and out-of-sample estimation results for the predictive regression of the monthly excess market return on one of the 14 economic predictors in Welch and Goyal (2008) $Z_t^k$and their first principal component $ECON^{PC}$asfollowing:

$$R_{t+1} = \alpha + \psi Z_t^k + \varepsilon_{t+1}$$

where R_{t+1} denotes the monthly excess market return (%). Panel B reports the results of forecasting excess market returns with the comprehensive media climate change concern constructed using PLS, and one of the economic predictors as following:

$$R_{t+1} = \alpha + \psi Z_t^k + \beta \Delta CMCCC_t^{PLS} + \varepsilon_{t+1}$$

We report the regression coefficients and R^2s. The significance of the estimates is based on Newey-West t-statistics with 12 lags. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is 2002:01–2020:12. (because the economic variables on Amit Goyal website are only available until the end of 2020).

	Panel A: Univariate			Panel B: Bivariate		
Economic Predictor	ψ(%)	R^2(%)	R^2_{os}(%)	β(%)	ψ(%)	R^2(%)
Dividend-price ratio (DP)	0.43	1.01	-3.09	-0.64 ***	0.37	3.23
Dividend yield (DY)	0.56	1.72	-1.14	-0.61 ***	0.48	3.70
Earning-price ratio (EP)	0.06	0.02	-7.72	-0.70 ***	0.15	2.61
Dividend-payout ratio (DE)	0.08	0.04	-11.23	-0.68 ***	-0.01	2.49
Sample variance (SVAR)	-0.07	0.03	-11.97	-0.68 ***	-0.08	2.52
Book-to-market ratio (BM)	0.34	0.62	-1.45	-0.63 **	0.17	2.64
Net equity expansion (NTIS)	-0.31	0.51	-3.23	-0.65 **	-0.11	2.55
Treasury bill rate (TBL)	0.44 *	1.06	0.32	-0.63 ***	0.10	2.53
Long-term bond yield (LTY)	0.77 ***	3.18	1.06 **	-0.62 **	0.72 ***	5.27
Long-term bond return (LTR)	0.29	0.44	-1.2	-0.69 ***	0.31	3.00
Term spread (TMS)	-0.23	0.29	-0.86	-1.08 ***	-0.78 **	4.94
Default yield spread (DFY)	-0.22	0.26	-5.9	-0.73 ***	-0.33	3.07
Default return spread (DFR)	0.37	0.73	-9.2	-0.64 ***	0.27	2.87
Inflation rate (INFL)	-0.3	0.48	-0.26	-0.71 ***	-0.36	3.19
$ECON^{PC}$	-0.21	0.23	-7.17	-0.67 ***	-0.07	2.52

Table 6: Comparison with Sentiment and Climate-related Attention variables

Panel A reports the in-sample estimation results for the predictive regression of the monthly excess market return on one of four sentiment measures, including the Baker and Wurgler (2006) investor sentiment index ($SENT^{BW}$) from 2002:01-2018:12, the Huang et al. (2015) aligned investor sentiment index ($SENT^{PLS}$) from 2002:01-2020:12,the Calomiris and Mamaysky (2019) average article sentiment ($SENT^{news}$) and sentiment on market topic ($SENT^{mkt\_topic}$) both from 2002:01-2019:05. Panel B reports the in-sample estimation results for the predictive regression of the monthly excess market return on one of six climate-related concern measures: *Natural Disaster Concern* and *Environmental Concern* which are constructed by Bybee et al., (2020) based on a topic modelling of Wall Street Journal articles. The AR (1) innovations in climate change news index(EGKLS^{wsj}) and the AR (1) innovations in the CH Negative Climate Change News Index($\mathrm{EGKLS}^{\mathrm{chneg}}$),both of which are constructed by Engle, Giglio, Kelly, Lee and Stroebel (2020). AR (1) innovations in Media Climate Change Concern ($\mathrm{ABBI}^{overall}$) from Ardia, Bluteau, Boudt and Inghelbrecht (2021) which is constructed based on news about climate change on eight major U.S. newspapers. The last one measure (PST^{decay_change}) is the change in climate concern measure from Pastor, Stambaugh and Taylor (2021) which is constructed based on Ardia, Bluteau, Boudt and Inghelbrecht (2021) measures. The sample period is from 2002:01-2017:06. We run the univariate regression as following:

$$R_{t+1} = \alpha + \psi D_t^k + \varepsilon_{t+1}$$

where R_{t+1} denotes the monthly excess market return (%), D_t^k is one of the sentiment measures on panel A or climate-related attention measures on panel B. We also run the bivariate regressionswith the comprehensive media climate change concern constructed using PLS, and one of sentiment or climate-related attention predictors as following:

$$R_{t+1} = \alpha + \psi D_t^k + \beta \Delta CMCCC_t^{PLS} + \varepsilon_{t+1}$$

We report the regression coefficients and R^2s. The significance of the estimates is based on Newey-West t-statistics with 12 lags. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

	Univariate		Bivariate		
	ψ(%)	R^2(%)	β(%)	ψ(%)	R^2(%)
Panel A: Investor sentiment					
$SENT^{BW}$	-0.50 **	1.48	-0.45 *	-0.30	2.42
$SENT^{PLS}$	-0.51	1.42	-0.83 ***	-0.70 *	4.98
$SENT^{news}$	0.35	0.72	-0.76 ***	0.60	3.77
$SENT^{mkt\_topic}$	0.53	1.63	-0.67 ***	0.63	4.17
Panel B: Climate-related concern					
Natural Disaster Concern	-0.08	0.04	-0.59 ***	-0.15	2.06
Environmental Concern	-0.28	0.45	-0.55 **	-0.05	1.96
EGKLS^{wsj}	0.07	0.03	-0.62 **	0.21	2.19
EGKLS$^{\text{chneg}}$	-0.16	0.14	-0.55 *	-0.03	1.65
ABBI^{overall}	0.19	0.25	-0.73 ***	0.37	3.64
PST^{decay_change}	0.12	0.08	-1.03 ***	0.69	4.44

Table 7: Asset allocation results

This table reportsannualized CER gains (in percentage) and annualized Sharpe ratiosfor a mean-variance investor with a risk-aversion coefficient γ= 1, 3 or 5, respectively. The investor allocates assets monthly between the stock market and the risk-free asset by applying the out-of-sample forecasts based on $\Delta CMCCC^{PLS}$,$\Delta CMCCC^{Equ}$ and $\Delta CMCCC^{Vol}$. We consider two scenarios: zero transaction cost and a proportional transaction cost of 50 basis points per transaction. The investment period is from January 2009 through September 2021.

	no transaction cost		50 bps transaction cost	
	CER gain (%)	Sharp ratio	CER gain (%)	Sharp ratio
Panel A: Risk aversion γ= 1				
$\Delta CMCCC^{PLS}$	4.39	1.20	2.86	1.09
$\Delta CMCCC^{Equ}$	5.10	1.20	3.71	1.10
$\Delta CMCCC^{Vol}$	5.29	1.22	4.03	1.12
Panel B: Risk aversion γ= 3				
$\Delta CMCCC^{PLS}$	5.40	1.09	3.44	0.95
$\Delta CMCCC^{Equ}$	4.02	1.03	2.12	0.88
$\Delta CMCCC^{Vol}$	4.00	1.03	2.07	0.88
Panel C: Risk aversion γ= 5				
$\Delta CMCCC^{PLS}$	3.55	1.01	1.78	0.84
$\Delta CMCCC^{Equ}$	2.84	1.00	1.27	0.83
$\Delta CMCCC^{Vol}$	2.85	1.01	1.28	0.84

Table 8: Out-of-sample R_{os}^2 with alternative methods

This table presents the out-of-sample R_{os}^2(%) of predicting h-month ahead market returns (h = 1, 3, 6, or 12) based on all available 15 individual media climate change concern using alternative methods: equal-weight Combine, elastic net (Enet) and Lasso, respectively. The out-of-sample evaluation period is 2009:01–2021:09. Statistical significance for R_{os}^2 is based on the *p*-value of the Clark and West (2007) MSFE-adjusted statistic for testing H_0: $R_{os}^2 \leq 0$ against H_A: $R_{os}^2 > 0$. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

Method	h=1	h=3	h=6	h=12
$Combine$	1.83**	6.00 **	9.58 ***	15.11 ***
$Elastic\ net$	1.36 *	8.20 ***	11.41 ***	19.42 ***
$Lasso$	1.35 **	9.51 **	13.01 ***	18.54 ***

Table 9: Forecasting channel of Comprehensive Media Climate Change Concern

This table reports the estimation results for the predictive regression as following:

$$y_{t+1} = \alpha + \beta_y \Delta CMCCC_t^{PLS} + \epsilon_{t+1}$$

where y_t is one of three estimated components of the S&P 500 log return for month t and $\Delta CMCCC_t^{PLS}$ is the PLS comprehensive media climate change concern index. The two estimated components of the S&P 500 log return are the expected return, cash flow news and discount rate news, corresponding to $\hat{\beta}_{\hat{E}}$, $\hat{\beta}_{CF}$, and $\hat{\beta}_{DR}$, respectively. We estimate these components using Campbell (1991) and Campbell and Ammer (1993) vector autoregression (VAR) approach based on the variable in the first column: the S&P 500 log return (r), 14 representative macroeconomic variables, and the first principal components extracted from the14 macroeconomic variables (PC). We report the regression slopes and Newey-West t-statistics with 12 lags. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively. The sample period is from 2002:01-2021:09

VAR variables	$\hat{\beta}_{CF}$	t-stat.	$\hat{\beta}_{DR}$	t-stat.
r, DP	-0.28 **	-2.12	0.21 **	2.07
r, DP, DY	-0.28 **	-2.10	0.21 **	2.02
r, DP, EP	-0.57 ***	-2.76	-0.05	-0.59
r, DP, DE	-0.57 ***	-2.76	-0.05	-0.59
r, DP, RVOL	-0.33 **	-2.36	0.10	0.99
r, DP, BM	-0.34 **	-2.44	0.10	1.00
r, DP, NTIS	-0.10	-0.66	0.14	0.67
r, DP, TBL	-0.18	-1.41	0.25 *	1.88
r, DP, LTY	-0.35 **	-1.97	0.13	0.62
r, DP, LTR	-0.29 **	-2.13	0.20 *	1.83
r, DP, TMS	-0.6 ***	-3.07	0.07	0.52
r, DP, DFY	-0.38 **	-2.08	0.20	1.17
r, DP, DFR	-0.28 **	-2.1	0.2 **	2.00
r, DP, INFL	-0.29 **	-2.15	0.23 **	2.26
r, DP, PC	-0.61 ***	-2.86	0.02	0.21

Table 10: Forecasting GMB returns with comprehensive media climate change concern

This table presents the regression slope, Newey–West *t*-value with 12 lags (in parentheses), and $R^2(\%)$ of predicting portfolio returns with both contemporaneous and lagged comprehensive media climate change concern.

$$R_{t+1}^{j} = \alpha + \beta X_{t+1}^{k} + \psi X_{t}^{k} + \epsilon_{t+1}$$

where R_{t+1}^{j} denotes monthly excess return (%) for the green-minus-brown portfolio GMB, the green leg G, and the brown leg B, respectively, X_t^k represents $\Delta CMCCC_t^{PLS}$ and $\Delta CMCCC_t^{Equ}$ in Panel A and Panel B, respectively. The sample period is 2009:08–2018:10. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

	Panel A: Results for $\Delta CMCCC^{PLS}$			Panel B: Results for $\Delta CMCCC^{Equ}$		
	GMB	G	B	GMB	G	B
$\beta(\%)$	-0.23	-0.01	0.23	-0.25	-0.06	0.19
	(-1.26)	(-0.02)	(0.53)	(-1.40)	(-0.21)	(0.46)
$\psi(\%)$	0.47**	-0.15	-0.61	0.48**	-0.06	-0.54
	(2.09)	(-0.49)	(-1.54)	(2.29)	(-0.22)	(-1.42)
$R^2(\%)$	3.53	0.20	1.39	3.76	0.11	1.12

Appendix

Table A1: **AR (1) residual for comprehensive media climate change concern measure**

This table presents the regression slope, Newey–West *t*-value with 12 lags, in-sample R^2, and out-of-sample R_{os}^2 of predicting market returns with comprehensive media climate change concern based on AR (1) model.

$$R_{t,t+h} = \alpha + \beta \Delta CMCCC_t + \epsilon_{t,t+h}$$

where $R_{t,t+h}$ is the cumulative market return (%) between months *t* and *t* + *h* (*h* = 1, 3, 6, or 12), and $\Delta CMCCC_t$ is the PLS, equal-weight and volatility weight aggregation index of 15 individual measures. For each individual measure, we first take square root of the number of news coverages and then take 24 months moving average. Next, we use the residual from AR (1) model to construct the individual media climate change concern. The in-sample period is 2002:01–2021:09 and the out-of-sample period is 2009:01–2021:09. Statistical significance for R_{os}^2 is based on the *p*-value of the Clark and West (2007) MSFE-adjusted statistic for testing H_0: $R_{os}^2 \leq 0$ against H_A: $R_{os}^2 > 0$. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

	h=1	h=3	h=6	h=12
Panel A: Results for $\Delta CMCCC^{AR(1)\_PLS}$				
β(%)	-0.69 ***	-1.86 ***	-3.26 **	-6.45 **
t-stat.	-3.00	-2.50	-2.22	-2.07
R^2(%)	2.62	5.85	8.08	16.73
R_{os}^2(%)	2.96 ***	5.32 ***	7.33 ***	12.53 ***
Panel B: Results for $\Delta CMCCC^{AR(1)\_Equ}$				
β(%)	-0.58 ***	-1.7 **	-3.13 **	-6.09 **
t-stat.	-2.61	-2.41	-2.21	-2.07
R^2(%)	1.86	4.90	7.45	14.91
R_{os}^2(%)	2.27 **	3.53 **	6.78 ***	14.95 ***
Panel C: Results for $\Delta CMCCC^{AR(1)\_Vol}$				
β(%)	-0.57 ***	-1.71 **	-3.07 **	-5.94 **
t-stat.	-2.58	-2.39	-2.18	-2.05
R^2(%)	1.81	4.93	7.18	14.21
R_{os}^2(%)	2.15 **	3.49 **	6.10 ***	15.96 ***

Table A2: Detailed Description of 14 Economic Variables

In the robustness check, we control for the following 14 economic variables of Goyal and Welch (2008).

Variable	Description
Log dividend-price ratio (DP)	log of a 12-month moving sum of dividends paid on the S&P 500 index minus the log of stock prices (S&P 500 index).
Log dividend yield (DY)	log of a 12-month moving sum of dividends minus the log of lagged stock prices.
Log earnings-price ratio (EP)	log of a 12-month moving sum of earnings on the S&P 500 index minus the log of stock prices.
Log dividend-payout ratio (DE)	log of a 12-month moving sum of dividends minus the log of a 12-month moving sum of earnings.
Excess stock return volatility (RVOL)	computed using a 12-month moving standard deviation estimator, as in Mele (2007).
Book-to-market ratio (BM)	book-to-market value ratio for the Dow Jones Industrial Average.
Net equity expansion (NTIS)	ratio of a 12-month moving sum of net equity issues by NYSE-listed stocks to the total end-of-year market capitalization of NYSE stocks.
Treasury bill rate (TBL)	interest rate on a three-month Treasury bill (secondary market).
Long-term yield (LTY)	long-term government bond yield.
Long-term return (LTR)	return on long-term government bonds.
Term spread (TMS)	long-term yield minus the Treasury bill rate.
Default yield spread (DFY)	difference between Moody's BAA- and AAA-rated corporate bond yields.
Default return spread (DFR)	long-term corporate bond return minus the long-term
Inflation (INFL)	calculated from the Consumer Price Index (CPI) for all urban consumers