Climate change concerns and the performance of green versus brown stocks☆

David Ardia[a], Keven Bluteau[b,*], Kris Boudt[c,d,e], Koen Inghelbrecht[d]

[a] *GERAD & Department of Decision Sciences, HEC Montréal, Canada*

[b] *Department of Finance, Université de Sherbrooke, Canada*

[c] *Solvay Business School, Vrije Universiteit Brussel, Belgium*

[d] *Department of Economics, Ghent University, Belgium*

[e] *School of Business and Economics, Vrije Universiteit Amsterdam, The Netherlands*

National Bank of Belgium, Working paper no. 395
First version: October 2, 2020
This version: May 2, 2022

Abstract

We empirically test the prediction of Pástor, Stambaugh, and Taylor (2021) that green firms outperform brown firms when concerns about climate change increase unexpectedly, using data for S&P 500 companies from January 2010 to June 2018. To capture unexpected increases in climate change concerns, we construct a daily Media Climate Change Concerns index using news about climate change published by major U.S. newspapers and newswires. We find that on days with an unexpected increase in climate change concerns, the green firms' stock prices tend to increase while brown firms' prices decrease. Further, using topic modeling, we conclude that this effect holds for concerns about both transition and physical climate change risk. Finally, we decompose returns into cash flow and discount rate news components and find that an unexpected increase in climate change concerns is associated with an increase (decrease) in the discount rate of brown (green) firms.

Keywords: Asset Pricing, Climate Change, Sustainable Investing, ESG, Greenhouse Gas Emissions, Sentometrics, Textual Analysis
JEL: G11, G18, Q54

☆We thank the participants at the 2020 International National Bank of Belgium conference on "*Climate Change: Economic Impact and Challenges for Central Banks and the Financial System*" and Prof. Dana Kiku for helpful comments. We also thank Elsa Allman, Laurent Barras, Tolga Cenesizoglu, Hannes Mohrschladt, Jérôme Taillard, and participants at the 2021 AFFI, CFE, EC^2, and EFMA conferences, the 2022 Annual QES European Quant and Macro Investment conference, and seminar participants at Quantact, GERAD, Department of Finance at Université de Liège, Department of Finance at Université de Sherbrooke, Department of Economics at Sichuan University, and Department of Economics at Ghent University for their feedback. Finally, we are grateful to the National Bank of Belgium, IVADO, and the Swiss National Science Foundation (grants #179281 and #191730) for their financial support.

*Corresponding author. HEC Montréal, 3000 Chemin de la Côte-Sainte-Catherine, Montreal, QC H3T 2A7. Phone: +1 514 340 6103.

Email addresses: `david.ardia@hec.ca` (David Ardia), `keven.bluteau@usherbrooke.ca` (Keven Bluteau), `kris.boudt@ugent.be` (Kris Boudt), `koen.inghelbrecht@ugent.be` (Koen Inghelbrecht)

1. Introduction

Many consider climate change one of the biggest challenges of our time. However, there is disagreement on the magnitude and the causes of the problem and how to address it. As a result of these differing views, some customers, regulators, and investors have strong preferences for sustainable solutions and investments that tackle the climate change problem, while others do not. Moreover, these preferences can change with new information. These preference shifts can affect the prices of financial assets (Fama and French, 2007). Anecdotal evidence suggests that preference shifts have caused rapid growth in sustainable (green) investing (GSIA, 2018) and a massive fossil fuel (brown) disinvestment campaign (Halcoussis and Lowenberg, 2019). These investment trends can be triggered or accentuated, for instance, by international conferences on climate change (*e.g.*, the 2012 UN Climate Change Conference), international agreements (*e.g.*, the Paris agreements), or new regulatory proposals (*e.g.*, the Climate Action Plan).[1]

Pástor, Stambaugh, and Taylor (2021, hereafter PST) propose a theoretical framework to model the impact of changes in sustainability preferences on asset prices. In the specific case of climate change, their model predicts that green stocks outperform brown stocks *when concerns about climate change strengthen unexpectedly*. The authors posit two mechanisms for this. First, investors can adjust their expectations about future green vs. brown firms' cash flows. This change in expectations results from a change in customers' and regulators' preferences for sustainability solutions. Due to an unexpected increase in climate change concerns, lawmakers are more likely to propose and implement legislation that would harm brown firms' cash flows relative to green firms. Customers are more likely to buy sustainable products. Second, their model assumes that agents care about environmental, social, and governance (ESG) criteria and climate change's social impact. Hence, investors with high sustainability preferences derive utility from owning shares in green firms rather than brown ones. Under their model assumptions, PST then show that the higher (resp. lower) the wealth-weighted mean of ESG tastes of investors, the lower

[1]These events are reflected in large values for the Media Climate Change Concerns index introduced in this paper.

(resp. higher) the expected excess returns of green (resp. brown) firms. Thus, an increase in customer and/or investor preferences for green firms, because of unexpected increasing concerns about climate change, has an immediate effect on stock prices. This can be seen in the discounted cash flow pricing model in PST (Equation 37). When faced with unexpected increases in climate change concerns, the customer channel reduces (boosts) the net cash flows of brown (green) firms, while the investor channel increases (decreases) the discount rate of brown (green) firms. Both channels then contribute to a decrease (increase) in stock prices of brown (green) firms. This paper empirically tests the prediction of PST that green firms outperform brown firms when concerns about climate change increase unexpectedly.

We test this prediction using daily returns of S&P 500 stocks and a novel proxy for unexpected changes in climate change concerns computed from news articles published on the same day. We use news articles for several reasons. First, as mentioned by Nimark and Pitschner (2019), consumers and investors rely on the media as an information intermediary between them and the state of the world. Through the agenda-setting channel, the media influence the concerns about climate change in terms of attention devoted to the issue. In addition, their framing influences people's attitudes and evaluations regarding climate change. A large number of studies have confirmed the underlying hypothesis that media are a powerful tool for increasing public awareness about environmental issues (*e.g.*, Boykoff and Boykoff, 2007; Sampei and Aoyagi-Usui, 2009; Hale, 2010). However, it is crucial to disentangle the expected level of concerns about climate change from the shocks that will drive a change in preferences for products and services provided by green versus brown firms and the investors' tastes holding such stocks. We use regression models to disentangle the effect of shocks in climate change concerns from other factors driving stock market returns.

The transient nature of shocks in climate change concerns requires a *daily* time series of climate change concerns. The choice of a daily horizon strikes a balance between timeliness and feasibility. Short measurement horizons are needed, given the fast reaction of stock markets to news, and diminish practical issues related to potential confounding factors when using lower frequencies such as weekly or monthly horizons.[2] A key contribution of our paper is to propose an algorithm that maps news articles into a daily time series that proxies the latent shocks in climate change concerns. The proposed solution is inspired by the two monthly media-based climate change risk indices introduced in Engle et al. (2020). Their first index captures the *attention* about climate change in the Wall Street Journal (WSJ). Their second index relies on the Crimson Hexagon proprietary sentiment measure to capture the *negative attention* about climate change in a large set of news outlets. None of these indices are available daily.

The first step in our algorithm is to collect news from ten major and highly circulated U.S. newspapers, including the Los Angeles Times, New York Times, Wall Street Journal, and USA Today, and two major newswires: Associated Press Newswires and Reuters News. We only select the media articles for which the subject categorization indicates that the articles discuss climate change. For those articles, we define a novel "concerns score" measuring and combining the levels of negativity and risk discussed in each article. A topic model analysis and an analysis of concern scores per outlet highlight their heterogeneity in terms of coverage, themes, and level of concerns related to climate change, thus confirming the need to consider a broad corpus like ours when proxying climate change concerns captured by U.S. news media. To account for this heterogeneity in our aggregation, we follow Baker, Bloom, and Davis (2016) and normalize each news source separately prior to combining the daily climate change concerns' scores into a daily aggregate Media Climate Change Concerns index (MCCC).[3] Finally, we obtain a proxy of unexpected changes in climate change concerns using the prediction error of an explanatory-variables-augmented autoregressive time series regression model calibrated on the MCCC index, which we refer to as unexpected media climate change concerns (UMC).

In our study, we carefully account for potential endogeneity between the daily UMC and the differential in performance of green and brown stocks. The endogeneity may arise

[2]An intraday analysis as in Boudt and Petitjean (2014) is not feasible since, for newspaper articles, we do not have the relevant intraday time stamp.

[3]The MCCC index is available at `https://sentometrics-research.com/`. Examples of recent research using our index are Alekseev et al. (2021), Ballinari and Mahmoud (2021), Campos-Martins and Hendry (2021), and Pástor, Stambaugh, and Taylor (2022), among others.

due to a potential feedback loop that an exogenous shock leading to increased concern in climate change affects the pricing of green and brown stocks, amplifying the investors' concerns about the impact of climate change and hence strengthening the market impact. We take several actions to mitigate endogeneity. First, we use a conservative approach in selecting the relevant articles. We only retain articles for which the news provider has tagged the topic as "climate change." In addition, we filter this corpus by removing all news articles mentioning keywords related to the stock market and market performance.[4] Second, we define UMC as the shock component in our MCCC index filtered from potential effects of financial-market, energy-related, and macroeconomic variables. Third, we use several sets of contemporaneous controls in our analyses. Throughout the paper, we abstain from any causal interpretation about the relation between the climate change concerns index and the performance of green versus brown stocks.

Our empirical analysis focuses on S&P 500 firms from January 2010 to June 2018. Testing the PST model for this universe requires defining a proxy for the firm's greenness characteristic driving the preferences of consumers and investors for green versus brown stocks. In our study, we quantify a firm's greenness as the ASSET4/Refinitiv carbon-dioxide-equivalent (CO2-equivalent) greenhouse gas (GHG) emissions data scaled by firms' revenue. Thus, the variable measures a firm's emissions intensity: The number of tonnes of CO2-equivalent GHG emissions necessary for a firm to generate $1 million in revenue. We base this choice on recent empirical evidence that the level of greenhouse gas emissions is a driver of returns and can thus be considered as a proxy for the greenness variable relevant to consumers and investors. In particular, Bolton and Kacperczyk (2021) study whether carbon emissions affect the cross-section of the U.S. stock market and find that stocks with higher emissions earn higher returns, consistent with investors demanding compensation for carbon risk.

[4]We acknowledge that while we explicitly exclude news articles discussing the stock market, it may still be that the news published during the calendar day is indirectly influenced by the stock market returns observed during the day. However, given the corpus at our disposal, we believe it is the best we can do.

We first analyze the contemporaneous relation between UMC and the daily return of a green-minus-brown (GMB) portfolio that is long in green firms and short in brown firms. Firms below the 25th percentile of the GHG emissions intensity on a given day are defined as green firms, and firms above the 75th percentile as brown firms. We find a significant positive relation, suggesting that green stocks tend to outperform brown stocks on days for which there is an unexpected increase in climate change concerns. When looking at the green (brown) portfolio returns individually, we find a positive (negative) and significant relation with UMC. This relation is stronger, in absolute terms, for the brown portfolio than for the green portfolio. Hence, when there is an unexpected increase in climate change concerns, investors tend to penalize brown firms more than to reward green firms.

Next, we use firm fixed-effects panel regressions to estimate the exposure of individual firms' stock returns to UMC, conditional on their emissions intensity. Our results align with our previous findings: The lower (higher) the emissions intensity, the more positive (negative) the firm's value changes on days with an unexpected increase in climate change concerns. In related work, Ilhan, Sautner, and Vilkov (2021) show that the industry, to a large extent, explains the variation in GHG emissions intensity. Moreover, Bolton and Kacperczyk (2021) find that institutional investors implement exclusionary screening based on emissions intensity in a few industries. Hence, we test whether the GHG emissions intensity still drives UMC exposure at the industry level. We find that is the case for only a minority of industries. The industry is thus a good predictor for the firm's value change on days with large unexpected increases in climate change concerns. While disclosure of GHG emissions improves over our sample, at least 30% of the S&P 500 firms do not disclose their GHG emissions. We study whether the returns of the non-disclosing firms are also higher for greener than for browner firms when we impute the missing GHG emissions intensity using their industry average. We find that it is the case indicating that the PST prediction also holds for firms that do not disclose their GHG emissions.

We also contribute to understanding the channels through which concerns about climate change news relate to the stock market. In particular, it seems self-evident that not all news articles about climate change lead to an outperformance of green versus brown stocks. Under the PST model, the price reaction is driven by shocks in climate change concerns that influence expectations about firms' cash flows or shocks that change investor preferences. In the final analysis, we investigate the nature of the shocks and study whether the obtained finding with the aggregate UMC variable differs when analyzing the relations for every climate change theme separately. Our analysis identifies four themes (*i.e.*, clusters of topics) related to climate change, namely (in order of prevalence): (i) "Business Impact", (ii) "Environmental Impact", (iii) "Societal Debate", and (iv) "Research". We then use the corresponding topic-probability weights per article to compute an MCCC index per theme, and repeat the above panel regression analysis with each thematic UMC as a covariate. We find that, for each thematic dimension, the sign of the coefficient is consistent with the prediction that, on days with unexpected increase in climate change concerns, green firms tend to have a higher return than brown firms. This relation is highly significant for the transition risk themes "Business Impact" and "Societal Debate." A more granular approach at the level of topics is needed to capture the market-relevant concerns about physical risk. We obtain corroborating findings when doing the analysis using monthly returns. For this frequency, we use the approach by Chen, Da, and Zhao (2013) to decompose the monthly returns into a cash flow and discount rate component. We then find only empirical support for a significant change in the discount rate factor returns component in months with an unexpected increase in climate change concerns.

Our results have practical implications for firm managers. First, we provide them with a tool to monitor the climate change concerns as expressed in media articles. Second, we show that the time variation in the UMC matters as a driver of firm value. In periods of high unexpectedly changes in climate change concerns, we find that green firms outperform brown firms. We mainly find evidence of a discount channel, implying that an increase in UMC manifests itself in an increase (decrease) in the cost of equity of brown (green) firms. Third, as the exposure to the UMC variable is driven by the firm's level of greenhouse gas emissions, the firm can decide to invest in improvements to its climate risk profile to benefit from an increase in investors' tastes for green firms.

By empirically verifying the predictions of PST using our new daily MCCC index, we contribute to a growing body of recent studies that focus on understanding the impact of climate change on financial markets. Giglio, Kelly, and Stroebel (2021) provide an excellent review of the literature. Expanding literature specifically focuses on the relation between climate change shocks and realized stock returns. In particular, Hong, Li, and Xu (2019) find that stock prices of food companies underreact to climate change risks. Choi, Gao, and Jiang (2020) find that in abnormally warm weather, stocks of carbon-intensive firms underperform those of low-emission firms. Bertolotti et al. (2019) analyze the impact of extreme weather events on U.S. electric utilities' stock prices. They find substantial price reactions after a hurricane makes landfall. Ramelli et al. (2021) study firms' stock price reactions and institutional investors' portfolio adjustments following the election of Donald Trump and the nomination of Scott Pruitt as the head of the Environmental Protection Agency, both climate change skeptics. They find that investors rewarded carbon-intensive firms but, surprisingly, also companies demonstrating more responsible climate strategies. A recent comprehensive overview of this strand of the literature is Alekseev et al. (2021).

Our study on the relation between unexpected increases in climate change concerns and realized returns of green versus brown firms also relates to the literature on climate risk premia. The PST model predicts that green stocks have lower expected returns due to investors' preference for green stocks and the ability of green stocks to better hedge climate risk. There is mixed empirical evidence about this prediction. Bolton and Kacperczyk (2021) study whether carbon emissions affect the cross-section of the U.S. stock market and find that stocks with higher emissions earn higher returns, consistent with investors demanding compensation for carbon risk. Görgen et al. (2020) develop and study a carbon risk factor using a long-short portfolio based on a carbon emissions-related measure but do not find evidence of a carbon risk premium. Engle et al. (2020) build a climate change risk proxy using Wall Street Journal news articles to hedge against climate change risks with the mimicking portfolio approach. Although our study is not about climate risk premia, our results do have implications on how to measure it. During periods characterized by increasing concerns about climate change, green stocks can still outperform brown stocks, despite having lower expected returns. This is confirmed by Pástor, Stambaugh, and Taylor (2022), showing that the return spread between environmentally-friendly and unfriendly stock disappears on days without climate change-concerns shocks. Hence, it is essential to account for climate change concerns when measuring climate risk premia or expected returns.

This paper is organized as follows. Section 2 presents our climate change concerns measure. Section 3 describes our data. Section 4 presents the empirical results on the performance of green vs. brown stocks. Section 5 examines which dimensions drive the relation between unexpected increases in climate change concerns and green vs. brown stock returns. Finally, Section 6 concludes.

2. News media and climate change concerns

To empirically study the model of PST, we need to measure unexpected changes in climate change concerns. Formally, given aggregate climate change concerns at time t, CC_t, we aim to capture:

$$\Delta CC_t - \mathbb{E}[\Delta CC_t | I_{t-1}] \,, \tag{1}$$

where ΔCC_t is the change in climate change concerns at time t and I_{t-1} is the information set available at time $t-1$. The challenge is that CC_t is not directly observable.

A potential proxy for CC_t is Gallup's annual Environment poll.[5] One could derive unexpected changes from this survey, in particular unexpected changes in answer to the question about how worried participants are about global warming or climate change. However, this survey (and others) is conducted very infrequently, limiting the measure's usefulness. Instead, we proxy ΔCC_t on a daily basis using news media data.

[5] https://news.gallup.com/poll/1615/environment.aspx

In the remainder of this section, we first present arguments on the validity of using news media information to proxy for (unexpected) changes in climate change concerns. Then, we describe our methodology to compute the daily Unexpected Media Climate Change Concerns (UMC) variable.

2.1. How the media relates to agents' changes in concerns about climate change

Several studies observe that the mass media is a powerful tool for increasing public awareness about environmental issues (*e.g.*, Schoenfeld, Meier, and Griffin, 1979; Slovic, 1986; Boykoff and Boykoff, 2007; Sampei and Aoyagi-Usui, 2009; Hale, 2010). Media can influence a population's perceptions in two ways: (i) via the informational content communicated in news articles and (ii) by the level of news coverage or attention on a particular subject. We hypothesize that this information is sufficient to derive a meaningful proxy of changes in climate change concerns.

Theoretical models of mass media communication support this hypothesis. For example, the dependency model of the media's effects by Ball-Rokeach and DeFleur (1976) implies that information transmitted by the media affects individuals' knowledge and perceptions when they have less information from other sources, such as personal experience. Most people do not directly experience climate change, given that the most severe consequences of climate change are predominantly future outcomes. As such, the media communicate the majority of the informational content about climate change to the public. The framing theory of Chong and Druckman (2007) is an alternative approach that supports the use of informational content communicated by the media. It states that the presentation of information (*i.e.*, how news is framed or presented) influences people's attitudes towards a subject. Based on this theory, the level of concern about climate change portrayed in the media should directly affect a population's concerns about climate change.

The media bias model of Gentzkow and Shapiro (2006) provides theoretical support that the level of media coverage can proxy for the level of attention on climate change. This model implies that in a highly competitive media environment, individual media outlets tend to cater to their readership's prior beliefs to increase their reputation and revenue. Therefore, if the media perceives that its readers are more concerned about a subject (*e.g.*, climate change), the level of coverage will increase.[6] Additionally, the agenda-setting theory of McCombs and Shaw (1972) states that a consumer of news learns how much importance to attach to an issue from the amount of information published about a news event. This theory implies a connection between news coverage about climate change and the level of importance people attach to climate change.

2.2. Method for calculating news article-level concerns

Our goal is to capture unexpected changes in climate change concerns. We define concerns as "the perception of risk and related negative consequences associated with this risk." From this definition, we design a score that measures concerns from the informational content of news articles. We rely on two lexicons: (i) A risk lexicon to determine the level of discussion about (future) risk events and (ii) a sentiment lexicon to assess the increase in (the perception of) risk. These lexicons are retrieved from the LIWC2015 software (Pennebaker et al., 2015).[7] The risk lexicon of this software is also used in Stecula and Merkley (2019) to analyze how the news media shape public opinion about climate change.[8]

With these lexicons, we compute what we refer to as the "concerns score." We assume a media universe of $s = 1, \ldots, S$ news sources. On each day $t = 1, \ldots, T$, source s publishes $n = 1, \ldots, N_{t,s}$ articles discussing climate change. Given the number of risk words $RW_{n,t,s}$, number of positive words $PW_{n,t,s}$, number of negative words $NW_{n,t,s}$, and total number of words $TW_{n,t,s}$ in a news article n published on day t by source s, the article's concerns score is defined as:

$$concerns_{n,t,s} = 100 \times \left(\frac{RW_{n,t,s}}{TW_{n,t,s}}\right) \times \left(\frac{NW_{n,t,s} - PW_{n,t,s}}{NW_{n,t,s} + PW_{n,t,s}} + 1\right) \Big/ 2 \,. \tag{2}$$

[6]See https://www.theguardian.com/environment/2019/apr/22/why-is-the-us-news-media-so-bad-at-covering-climate-change.

[7]The academic version is available at https://liwc.wpengine.com/.

[8]The media sources used in Stecula and Merkley (2019) are the New York Times, Wall Street Journal, Washington Post, and Associated Press. These are also used in our study.

The first ratio of the product, $\left(\frac{RW_{n,t,s}}{TW_{n,t,s}}\right)$, measures the percentage of risk words in the text. Using the percentage rather than the number of risk words accounts for variability in news articles' lengths. The second ratio, $\left(\frac{NW_{n,t,s}-PW_{n,t,s}}{NW_{n,t,s}+PW_{n,t,s}}+1\right)/2$, measures the degree of negativity (with zero being the most positive text and one being the most negative), which allows us to differentiate between negative and positive articles. Thus, our article-level concerns score can be interpreted as a weighted textual risk measure, where a higher (lower) weight is attributed when a text is more negative (positive).

2.3. Aggregation

We construct a daily index that captures changes in climate change concerns by aggregating article-level concerns scores. First, we define the daily concerns score for day t and a given source s as the sum of the article-level concerns scores across $N_{t,s}$ articles related to climate change:

$$concerns_{t,s} = \sum_{n=1}^{N_{t,s}} concerns_{n,t,s} = N_{t,s} \times \overline{concerns}_{t,s}\,. \tag{3}$$

As shown in Equation (3), the sum can be expressed in two parts: (i) $N_{t,s}$ (the number of news articles published about climate change on day t by source s) and (ii) $\overline{concerns}_{t,s}$ (the average concerns score in the news published about climate change on day t by source s). Thus, the index captures both the level of media attention and the (average) level of concerns expressed in news articles on a given day for a given source, two important components as explained in Section 2.1. Note that when no news is published about climate change (*i.e.*, $N_{t,s} = 0$), the concerns score in Equation (3) is zero, which is equivalent to a 100% positive sentiment term in Equation (2). As such, our approach assumes that no news is good news.[9]

[9]In their theoretical analysis of carbon prices over the next hundred years, Gerlagh and Liski (2018) assume that individuals' beliefs that climate change will have a long-term impact decrease over time and increase in the presence of information about the damage of climate change. Thus, they make a similar assumption that no news is good news.

Second, to account for heterogeneity between sources, we follow the source-aggregation methodology of Baker, Bloom, and Davis (2016). For each source s, we compute the standard deviation of the source-specific index over a time range τ_1 to τ_2 ($1 \leq \tau_1 < \tau_2 \leq T$):

$$\sigma_s = \sqrt{\frac{\sum_{\tau=\tau_1}^{\tau_2} (concerns_{\tau,s} - \overline{concerns}_s)^2}{\tau_2 - \tau_1}}\,, \tag{4}$$

where $\overline{concerns}_s$ is the sample mean computed over τ_1 to τ_2. We use the standard deviation to normalize the source-specific index over the $t = 1$ to $t = T$ period:

$$nconcerns_{t,s} = \frac{concerns_{t,s}}{\sigma_s}\,. \tag{5}$$

The normalization is required to aggregate the per-source indices in the next step properly. For instance, consider a source that typically publishes five articles about climate change daily and a competing source that tends to publish one climate change article per day. At some point, however, that second source may publish five articles about climate change. We posit that if the second source suddenly publishes more about climate change than usual, there is a higher probability that a relevant climate-change-related event has occurred. We capture this effect with the by-source normalization. Specifically, we add more weight to the signal available in each source's time-series variation than to differences across sources.

Finally, we compute the Media Climate Change Concerns (MCCC) index at day t by applying an increasing concave function $h(\cdot)$ to the average of the normalized source-specific climate change concerns for that day:

$$MCCC_t = h\left(\frac{1}{S}\sum_{s=1}^{S} nconcerns_{t,s}\right). \tag{6}$$

We use an increasing concave mapping function $h(\cdot)$ to capture the fact that increased media attention always increases climate change concerns, but at a decreasing rate: One concerning article about climate change may increase concerns, but 20 concerning articles are unlikely to increase concerns 20 times more. One reason for this non-linear relationship is the "echo chamber" phenomenon, in which groups tend to read the news that agrees with their views, limiting the reach of alternative information to these groups (*e.g.*, Flaxman, Goel, and Rao, 2016). Another argument comes from the concept of "opinion inertia," which arises, for instance, from the confirmation bias (*e.g.*, Doyle et al., 2016). In this case, individuals have difficulties changing their opinion irrespective of available information. An example of a group with opinion inertia are so-called "global warming skeptics." We set $h(\cdot)$ to the square root function in the rest of the paper.[10]

2.4. Unexpected changes in the Media Climate Change Concerns variable

So far, we have developed a methodology to proxy for changes in climate change concerns, ΔCC_t, using media information. Our aim, however, is to derive *unexpected* changes in climate change concerns. Because the media tends to publish unexpected information, it is reasonable to use $MCCC_t$ as a baseline proxy for unexpected changes in climate change concerns. However, some news might still be expected due to numerous factors, such as pre-announcements (*e.g.*, planned international conferences) or the presence of stale news (*e.g.*, republishing an article with only slight modifications to the text). Additionally, some studies suggest that the current state of the economy may also influence public perception about climate change (Scruggs and Benegal, 2012). Therefore, to capture the shock component in our MCCC index and filter the potential effects of financial-market, energy-related, and macroeconomic variables, we use an explanatory-variables-augmented autoregressive time series model (ARX) to estimate the expected component of $MCCC_t$. We interpret the prediction error as a proxy for the unexpected changes in climate change concerns (*i.e.*, $\Delta CC_t - \mathbb{E}[\Delta CC_t | I_{t-1}]$). We refer to the prediction error as UMC_t in the remainder of the paper. More details are provided in Section 3.2.

2.5. Comparison with existing methodologies

Thanks to the increasing availability of media news, several media-based time series have been proposed over the past years. According to Gentzkow, Kelly, and Taddy (2019),

[10] We also use $h(x) = \log(1 + x)$ as a robustness check, and obtain results and conclusions that are qualitatively similar.

the most influential media-based time series in economics is the EPU index developed by Baker, Bloom, and Davis (2016). This index uses counts of articles containing at least one keyword from the categories economy, policy, and uncertainty. The index is a simple average of the normalized count across various newspapers. Such a count-based approach is a prototypical example of a media attention-based index capturing the intensity of news coverage. An alternative design is to compute an average feature across all articles that satisfy a condition. This leads to an economic sentiment (resp. uncertainty) index in case of averaging the sentiment (resp. uncertainty) in economic news articles. The proposed MCCC index combines the two approaches into a media-based concern index. The more each article about climate change expresses a negative sentiment and high risk, and the more attention the media attaches to it in terms of the number of articles published, the higher the concern.

While several media-based economic time series already exist, the construction of media-based time series for understanding the impact of climate change on the financial market is more recent. The pioneering contribution by Engle et al. (2020) proposes two monthly indices capturing climate change risk using news articles. A first approach relies on WSJ news articles and a lexicon called the "Climate Change Vocabulary" (CCV) derived from authoritative texts about climate change. The method extracts a similarity feature between each news article in the corpus and the CCV. The higher the similarity measure, the more likely an article discusses climate change. This similarity feature is then aggregated monthly to obtain a climate change risk index. Their second approach relies on the natural language proprietary algorithms of Crimson Hexagon to compute news articles' negative sentiments about climate change.

We are the first to propose a daily time series capturing climate change concerns in the media. In Table 1, we summarize our approach and compare it with the indices by Engle et al. (2020) and the more recent proposals by Kapfhammer, Larsen, and Thorsrud (2020), Faccini, Matin, and Skiadopoulos (2021), Bessec and Fouquau (2021), and Bua et al. (2022). Following Ardia, Bluteau, and Boudt (2019), we organize the comparison based on the main steps of constructing a media-based time series: (i) choice of corpus and selection of the relevant news articles, (ii) calculation of the relevant features per article, (iii) cross-sectional aggregation, and (iv) time-series aggregation. For each of the time series, we find that there is always at least one crucial step in which the proposed MCCC index stands out in terms of enabling researchers to test the association between market returns and concerns about climate change expressed by the media on the same day.

[Insert Table 1 about here.]

In the last rows of Table 1, we indicate the correlation of the alternative media-based climate change indices with the MCCC and UMC time series when the alternative is available for download. The correlation is at most 44%, confirming the specificity of each index as described by their unique scope and choices in the index design.

3. Data

Our study relies on climate change news articles published by multiple sources, data on firms' annual greenhouse gas emissions, annual revenue, and daily stock returns.

3.1. Climate change news corpus

We retrieve climate change-related news articles from U.S. newspapers and newswires from January 1, 2003, to June 30, 2018.[11] We select high circulation newspapers so that these sources have a reasonable chance of influencing the population's concerns about climate change. The selection is based on 2007 circulation data from Alliance for Audited Media.[12] We consider newspapers with a daily circulation of more than 500,000: (i) New York Times, (ii) Washington Post, (iii) Los Angeles Times, (iv) Wall Street Journal, (v) Houston Chronicle, (vi) Chicago Tribune, (vii) Arizona Republic, (viii) USA Today, (ix)

[11] We use data from 2003 to 2009 to compute the standard deviation parameter required for the index construction (see Equation (5)) and perform our analyses over the 2010 to 2018 period.

[12] See https://auditedmedia.com/.

New York Daily News, and (x) New York Post. In addition, we consider articles published by major newswires: (i) Associated Press Newswires and (ii) Reuters News.

News articles published by these sources are available in DowJones Factiva, ProQuest, and LexisNexis databases. For DowJones Factiva and ProQuest, we identify climate change-related news articles by picking articles in the "Climate Change" topic category. For LexisNexis, we use the subject "Climate Change" with a relevance score of 85 or more.[13] We filter out short news articles with fewer than 200 words, as lexicon-based methods are typically noisy for short texts. Finally, we exclude news articles discussing the stock market using several keyword-based filters to avoid these articles introducing reverse causality in our analyses. The list of the keywords is presented in the Appendix, Section A.

In Table 2, we report statistics about the number of climate change articles published by the sources in our sample. The source that publishes the most about climate change is the Associated Press Newswires, with 10,061 articles. The Wall Street Journal publishes the most relative to its total number of articles (0.26%). The Chicago Tribune, New York Daily News, and New York Post published the least about climate change relative to their total number of articles. In particular, while the Chicago Tribune has more total articles about climate change than USA Today (482 vs. 234), USA Today publishes more about climate change in relative terms than the Chicago Tribune (0.08% vs. 0.03%). Table 2 also reports information regarding the concern scores extracted from the articles published by the various sources. In particular, we see that the average score ranges from 0.31 for the Arizona Republic to 0.44 for USA Today and the New York Post. The percentage of articles with a zero concerns score is also much larger for the Arizona Republic and the two newswires than the other outlets. This highlights the discrepancies in news reporting, whereby newswire articles are, on average, less opinionated than newspaper articles. This heterogeneity, both in terms of coverage and concern, underlines that standardization

[13]LexisNexis indexes each article with metadata information, such as the topic of the article. These metadata tags are associated with a relevance score, where a score of 60 to 84 indicates a minor reference and a score of 85 and above indicates a major reference.

by sources before aggregation is necessary, as each source covers and treats information related to climate change differently.

[Insert Table 2 about here.]

To get a better overview of climate change topics discussed in our set of articles, we estimate the correlated topic model (CTM) of Lafferty and Blei (2006) on our corpus. The CTM model is an unsupervised generative machine-learning algorithm that infers latent correlated topics among a collection of texts.[14] In particular, each text is a mixture of K topics, and each topic is a mixture of V words. The approach yields: (i) a vector of topic prevalence $\theta_{k,n,t,s}$ for each news article where $\sum_{k=1}^{K} \theta_{k,n,t,s} = 1$ with $\theta_{k,n,t,s} \geq 0$, and (ii) a vector of word probabilities $\omega_{v,k}$ for each topic, where $\sum_{v=1}^{V} \omega_{v,k} = 1$ with $\omega_{v,k} \geq 0$. To calibrate the CTM on our news corpus, we proceed as follows. First, we estimate the CTM for the range of $K \in \{10, 20, \ldots, 100\}$ topics and select the optimal number using semantic coherence and exclusivity metrics. Second, we manually label the topics by: (i) looking at the ten most-probable words for each topic and (ii) looking at the content of the articles with the largest topic prevalence. Third, we organize (group) the topics into clusters that constitute more general themes related to climate change for ease of interpretation. We construct the themes based on clustering and network analysis. We refer to the Appendix, Section B, for details regarding the topics' and clusters' construction.

With our corpus, we find that $K = 30$ is the optimal number of topics and that topics can be grouped into four themes. In Table 3, we report for each theme the labeled topics together with their ten highest-probability keywords. In Table 4, we report the topics' and themes' unconditional prevalence and their average climate change concern score. The unconditional prevalence of a topic is obtained as the average of the topic prevalences across all news articles. For a theme, the unconditional prevalence is the sum of its topics' unconditional prevalences. The average climate change concerns score for a

[14] Hansen, McMahon, and Prat (2018), Larsen and Thorsrud (2017), Larsen (2021), and Faccini, Matin, and Skiadopoulos (2021) estimate latent topics using the popular Latent Dirichlet Allocation (LDA) model of Blei, Ng, and Jordan (2003). However, the LDA model does not account for possible correlations between topics. We find that allowing for non-zero correlation with the CTM model generates more coherent topics.

topic (or a theme) is computed as a weighted sum of the articles' score, where the weights are the topic (or the theme) unconditional prevalences. In addition, we categorize each topic and theme into one of the three types of climate risk put forward in NGFS (2020), namely: (i) physical risk, (ii) transition risk, and (iii) liability risk. Physical risks can be acute if they arise from climate and weather-related events and direct destruction of the environment or chronic if they arise from progressive shifts in climate and weather patterns or gradual loss of ecosystem services. Transition risk results from the process of adjustment towards a lower-carbon economy, arising, for instance, from new regulations, technologies, or social and market sentiment. Finally, liability risk can be considered a subset of either physical or transition risks and results from potential climate change-linked legal liability.

From Table 4, we see that the most prevalent theme is "Business Impact" (prevalence of 51.13%). The topics forming this theme, such as "Renewable Energy" and "Carbon Tax," can be associated with transition risk. The exception is "Legal Actions" which is rather related to liability risk. The second most prevalent theme is "Environmental Impact" (prevalence of 20.17%), the topics of which, such as "Extreme Temperatures" and "Glaciers/Ice Sheets," are related to acute and chronic physical risk, respectively. The third theme is "Societal Debate" (prevalence of 18.14%), constituted of four topics among which "Political Campaign" can be associated with transition risk. The last theme is "Research" (prevalence of 10.56%), formed by topics related to both physical and transition risks. Indeed, while the topic's subject within that theme is often related to (future) physical risk, it is also often accompanied by policy and business recommendations and implications, which enter the realm of transition risk. For instance, the topic "UN/IPCC" captures the content of UN/IPCC reports, for which the IPCC's institution goal is to "provide regular assessments of the scientific basis of climate change, its impacts and future risks, and options for adaptation and mitigation."[15]

[Insert Table 3 and Table 4 about here.]

[15]See https://www.ipcc.ch/about/.

The relation between the topics of discussion in our news corpus is displayed via a correlation network in Figure 1. The plot highlights the correlation of prevalences between the topics (*i.e.*, higher likelihood of topics being discussed together in the same news article). The figure clearly shows two clusters of topics corresponding to "Business Impact" (left part) and "Environmental Impact" (right part). In the middle, linking the two clusters, topics are related to the theme "Societal Debate." Finally, in the bottom right, we see the three topics corresponding to the theme "Research."

[Insert Figure 1 about here.]

In the last four columns of Table 2, we also report the themes' unconditional prevalence for the sources in our sample. The unconditional prevalence of a topic (theme) is obtained as the average of the topic (theme) prevalences across all news articles within a source. While "Research" is the least covered theme among all sources, we see heterogeneity for the three other themes. For instance, "Business Impact" is the major theme in all outlets but the Chicago Tribune, New York Daily News, and New York Post, for which "Societal Debate" is the most prevalent theme of discussion. We also see that "Environmental Impact" is the second most prevalent theme in newswires, while in newspapers, it is "Societal Debate." The news media's discrepancies in the coverage of climate change-related topics emphasize the importance of working with several sources when building a media-based index, the objective of which is to capture climate change concerns.[16]

To better understand how much attention the media devotes to these topics over time, we aggregate the topic weights per article into a monthly time series of "article-equivalents" defined as the total of all article weights per topic. This quantity measures the hypothetical number of news articles uniquely discussing a specific topic for a given period. Formally, the number of article equivalents between dates t_1 and t_2 for topic k is defined as $\sum_{t=t_1}^{t_2} \sum_{s=1}^{S} \sum_{n=1}^{N_{t,s}} \theta_{k,n,t,s}$. We then aggregate the number of article equivalents by theme.

[16]For instance, the attention-based index in Engle et al. (2020) is only based on Wall Street Journal articles. In our corpus, this index would correspond mostly to "Business Impact" as the theme's prevalence is 56.50%.

In Figure 2, we display the monthly number of article-equivalents for each theme from January 2003 to June 2018. We observe significant time variations in the percentage of coverage devoted to each theme. For instance, "Business Impact" tends to have a larger number of article-equivalents during months when there are notable conferences on climate change (*e.g.*, 2009 Copenhagen UN climate change conference), and "Societal Debate" spikes when Trump announced his intention to withdraw from the Paris Agreement.

[Insert Figure 2 about here.]

3.2. Media Climate Change Concerns index

We build the MCCC index following the methodology in Section 2. We compute the source-specific standard deviation σ_s necessary to obtain the standardized source-specific Media Climate Change Concerns with media articles from 2003 to 2009. Then, we aggregate the resulting source-specific indices to obtain the MCCC index for 2010 to 2018. In Figure 3, we display the daily evolution of the index from 2003 to 2018. Note that the 2003-2009 period is forward-looking and is not used in the main analysis but is still of interest for validating the index. We interpret the daily index as a proxy for changes in climate change concerns. We also display a 30-day moving average of the index to help identify trends and events.[17]

[Insert Figure 3 about here.]

First, we see that the index's spikes correspond to climate change events, such as the 2012 Doha United Nations (UN) Climate Change Conference or the Paris Agreement. We also note that climate change concerns, proxied by the moving average, exhibit phases of low and high values. A first period of elevated concerns is observed following the 2007 UN Security Council talks on climate change and lasts until the beginning of 2010, after the

[17]This moving average can be interpreted as a proxy for the level of climate change concerns. This requires an assumption that climate change concerns only decrease because of the passage of time and that news published more than 30 days in the past do not have any effect on current climate change concerns.

Copenhagen UN Climate Change Conference. The second elevated period starts at the end of 2012, near the UN Climate Change conference, and lasts until the Paris Agreement. Later, we note a spike in concerns around the time of U.S. President Donald Trump's announcement that the U.S. will withdraw from the Paris Agreement. These observations suggest that our index captures meaningful events that correlate with increases in climate change concerns.[18]

We extract the unexpected component of the MCCC index as the prediction error of an explanatory-variables-augmented first-order autoregressive model (*i.e.*, ARX) for the MCCC. Specifically, we consider the following model:

$$MCCC_t = \mu + \rho MCCC_{t-1} + \boldsymbol{\gamma}' \mathbf{x}_{t-1} + \epsilon_t \,, \tag{7}$$

where the vector of explanatory variables $\mathbf{x}_t$ is included to mitigate the problem of endogeneity by capturing potential confounders that may affect the MCCC index. The vector includes financial-market, energy-related, and macroeconomic variables, but also variables capturing green/brown performance. We refer to the Appendix, Section C, for the list and description of the variables. We estimate the ARX model (7) on a daily rolling-window basis (of size 1,000) and use the prediction error for UMC_t. Estimation results are available in the Appendix, Section D.

3.3. S&P 500 stock universe and its greenhouse gas emissions intensity

Our analyses require the identification of green and brown firms. We define green (brown) firms as firms that create economic value while minimizing (not minimizing) damages that contribute to climate change. We use the greenhouse gas (GHG) emissions disclosed by firms to quantify these damages. We retrieve these variables from the Asset4/Refinitiv database. Similar to Ilhan, Sautner, and Vilkov (2021), we focus on S&P 500 firms because surveys of greenhouse gas emissions typically target these firms.

[18] As our index is bounded at zero by construction, it is more likely to better capture increases than decreases in climate change concerns.

The GHG emissions variable is separated into three scopes defined by the GHG Protocol Corporate Standard.[19] Scope 1 emissions are direct emissions from owned or controlled sources. Scope 2 emissions are indirect emissions from the generation of purchased energy. Scope 3 emissions are all indirect emissions (not included in Scope 2) that occur in a firm's value chain. These are reported in tonnes of carbon dioxide (CO2) equivalents. We focus on total GHG emissions, defined as the sum of the three emissions scopes.[20] To account for the economic value resulting from a firm's GHG emissions, we scale total GHG emissions by the firm's annual revenue obtained from Compustat. Whether a firm is classified as green or brown depends on its position within the distribution of firms by their total tonnes of CO2-equivalent GHG emissions attributed to $1 million of revenue at a point in time. This scaled-GHG variable is referred to as GHG emissions intensity (see Drempetic, Klein, and Zwergel, 2020; Ilhan, Sautner, and Vilkov, 2021).[21]

In Table 5, Panel A, we report the percentage of firms in the S&P 500 with available GHG emissions. While our GHG emissions source differs from Ilhan, Sautner, and Vilkov (2021), who use the Carbon Disclosure Project database,[22] we see that its coverage of S&P 500 firms is similar, with a yearly average at 63.75%. In Panel B, we report the average and standard deviation of GHG intensities for the industries in our sample, as defined by the 48-industries classification in Fama and French (1997).[23] We can notice the considerable heterogeneity across the industries. While the average emissions intensity is 488.75 tonnes of CO2-equivalent emissions per $1 million in revenue, the 25th and 75th percentiles are 49.92 and 600.28, respectively. The most polluting industry is "Utilities" with an average of 4,072 tonnes, and the least polluting industry is "Construction" with

[19]See https://ghgprotocol.org/standards.

[20]The results from our analysis are similar when excluding Scope 3 emissions.

[21]The environmental dimension of ESG scoring is an alternative variable to classify firms on the green to brown spectrum. However, Drempetic, Klein, and Zwergel (2020) suggest that these scores do not adequately reflect firms' sustainability. Additionally, Berg, Koelbel, and Rigobon (2019) show that the correlations between ESG scores of different data providers are weak, indicating a lack of reliable and consistent scoring methodology across providers.

[22]See https://www.cdp.net/.

[23]We use the 48-industries classification of Fama and French (1997) to strike a balance between a sufficient number of firms and homogeneity in terms of GHG intensities within each industry. Note that we consider 47 instead of 48 industries, as the "Fabricated Products" classification is absent in our sample. Industry classification is retrieved from Kenneth French's website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html.

4.45 tonnes. The table also displays the percentage of GHG emissions' disclosure per industry, defined as the number of years and firms' pairs with GHG emission data divided by the total number of years and firms' pairs in the industry. We see some heterogeneity across industries, but the disclosure rate is relatively high, especially for the most polluting firms. The across-industry average disclosure is about 65%, and the 25th and 75th percentiles are about 50% and 83%, respectively. Finally, we report in Panel B the percentage of institutional ownership retrieved from Thomson-Reuters Institutional Holdings for the firms that disclose or do not disclose their emissions. The average institutional ownership is high, at around 80% for the two groups. Also, the percentages are high and very similar between the two groups for each industry. Given the high institutional ownership, we can say that the typical investor is knowledgeable about the emissions profile of the stocks.

[Insert Table 5 about here.]

We note that GHG emissions are typically reported with a one-year delay. Similar to Ilhan, Sautner, and Vilkov (2021), we account for this by shifting the GHG emissions intensity variable by 12 months in our analyses.

4. Empirical results on the performance of green vs. brown stocks

In this section, we empirically test the prediction of the theoretical model of PST that green firms outperform brown firms when concerns about climate change increase unexpectedly. We first construct portfolios of green and brown stocks and test the prediction both using a conditional mean analysis (Section 4.1) and multivariate factor analysis (Section 4.2). Our main analysis is in Section 4.3. Using a firm fixed-effect panel regression model, we test whether the firm's daily stock returns are explained by an interaction effect between the firm's GHG emissions intensity and the UMC variable. The testable prediction is that firms with a higher GHG emission intensity are more negatively exposed to unexpected changes in climate change concerns (see Section 4.3.1). We further test whether this prediction still holds when considering variations in GHG emissions intensity within industries rather than across industries. Ilhan, Sautner, and Vilkov (2021) show that the variation in GHG emissions intensity is, to a large extent, explained by the industry. Moreover, Bolton and Kacperczyk (2021) find that institutional investors implement exclusionary screening based on direct emissions intensity in a few industries. Hence, the negative relation between the firm's exposure and GHG emission intensity may be driven by industry effects (see Section 4.3.2). Finally, as a non-negligible number of firms do not disclose their GHG emissions (see Table 5), we test whether non-disclosing firms are also affected by climate change concerns based on their industry and if this effect differs from firms that disclose their emissions. Given the result of Ilhan, Sautner, and Vilkov (2021), we do not expect differences in stock return changes between disclosing and non-disclosing firms (see Section 4.3.3).

4.1. Conditional mean analysis

Each day in the sample, we divide stocks into three groups: green, brown, and neutral. Green (brown) stocks are firms with a GHG emissions intensity variable in the lowest (highest) quartile of all firms' values available on that day. Neutral firms are the remainder of firms that disclose GHG emissions data.[24] We then build, for each day, equal-weighted portfolios for these groups. Because of the one-year delay in emissions reporting, we rely on former-year emissions when building the portfolios. As such, our portfolio formation strategy does not suffer from look-ahead bias.[25]

Our first analysis focuses on the average return of the green minus brown (GMB) portfolio conditional on the *UMC* variable. In Figure 4, we display the average performance of the GMB portfolio conditional on threshold values for *UMC*, obtained as the percentiles of *UMC* over the 2010-2018 period. We see a positive relation between the average return

[24]Our definition of neutral firms does not imply that those firms are carbon neutral (*i.e.*, having net-zero GHG emissions), but rather that they are average in terms of GHG emissions intensity across all firms in our dataset.

[25]The GHG emissions are updated yearly but at a different time across firms, and stocks can enter or exit the S&P 500 universe on any day. In the Appendix, Section E, we show that stocks belonging to a given category (green, brown, or neutral) have a high likelihood of remaining in the same category over time. Hence, while the rebalancing is daily, the portfolios' constituents are stable over time.

and *UMC*. In particular, when *UMC* is above its median, we notice strong increases in the GMB portfolio average return as the threshold becomes larger, especially at the extreme. Moreover, the average GMB portfolio return is always higher when the *UMC* is above the threshold than when it is below. These preliminary findings indicate that green firms outperform brown firms when there are unexpected increases in climate change concerns.

[Insert Figure 4 about here.]

4.2. Multivariate factor analysis

We now consider a multivariate linear regression framework to control for other factors that potentially drive stock returns. We regress the green minus brown ($p = GMB$), green ($p = G$), brown ($p = B$), and neutral ($p = N$) portfolios' excess returns, $r_{p,t}$, on UMC_t, and control variables ($\boldsymbol{CTRL}_t$):

$$r_{p,t} = c_p + \beta_p^{UMC} UMC_t + \boldsymbol{\beta}_p' \boldsymbol{CTRL}_t + \varepsilon_{p,t}\,, \tag{8}$$

where c_p is a constant, β_p^{UMC} and $\boldsymbol{\beta}_p$ are regression coefficients, and $\varepsilon_{p,t}$ is an error term. Given the PST model, we expect that $\beta_{GMB}^{UMC} > 0$, $\beta_G^{UMC} > 0$, and $\beta_B^{UMC} < 0$. We consider four sets of controls in specification (8):

- ***CTRL-1***: *MKT*, the excess market return;
- ***CTRL-3***: Set ***CTRL-1*** augmented with *HML*, the high-minus-low factor, and *SMB*, the small-minus-big factor, of Fama and French (1992);
- ***CTRL-6***: Set ***CTRL-3*** augmented with *RMW*, the robust-minus-weak factor, and *CMA*, the conservative-minus-aggressive factor, of Fama and French (2015), and *MOM*, the momentum factor of Carhart (1997);
- ***CTRL-15***: Set ***CTRL-6*** augmented with *WTI*, the crude oil return, *NG*, the natural gas return, *PROP*, the propane return, *EPU*, the economic policy uncertainty index of Baker, Bloom, and Davis (2016), *VIX*, the CBOE volatility index, the *TED*

spread, *TERM*, the term spread factor, and *DFLT*, the default spread factor of Fung and Hsieh (2004), and *FTS*, the flight-to-safety index of Baele et al. (2020).

The variables in ***CTRL-1***, ***CTRL-3***, and ***CTRL-6*** are commonly used in the finance literature. The set of controls in ***CTRL-15*** extends these variables with energy-related and macroeconomic variables to mitigate the potential effect of confounders. We refer to the Appendix, Section C, for more details on the control variables.

Estimation results are reported in Table 6 for the largest set of controls, while results for the alternative specifications are available in Table 7. First, we consider the GMB portfolio. We see that the estimated coefficient for *UMC* aligns with our hypothesis. Specifically, a one-unit increase in *UMC* implies an additional daily positive return of 7.2 basis points. This effect is significant at the 5% level. Looking at the green portfolio, we find a positive and significant exposure to *UMC*. We find a negative coefficient for the brown portfolio, significant at the 10% level.

The estimated coefficients for the control variables indicate that the GMB portfolio is positively related to *MKT*, *HML*, *SMB*, *MOM*, and *TERM*, negatively related to *CMA*, *RMW*, *WTI*, and *PROP*. Thus, the GMB portfolio emphasizes small firms with lower growth, aggressive investment policies, and weak operating profits. The *CMA* coefficient (-0.462) is large compared to the other asset-pricing coefficients. This finding is consistent with green firms investing more and brown firms investing less, which is another implication of the PST model. This prediction arises from the idea that green firms' capital costs are lower than brown firms'. Thus, more investment opportunities for green firms have a positive net present value, resulting in a higher investment level relative to their size than for brown firms. The positive coefficient for *TERM* indicates that green stocks outperform (underperform) brown stocks when there are good (bad) prospects about the economy (as reflected by a positive (negative) term spread). This suggests that investors are more (less) concerned about green investments during good (bad) times when their wealth constraints are less (more) binding, in line with the result of Bansal, Wu, and Yaron (2021). The coefficients for *WTI* and *PROP* show that green (brown) firms are negatively (positively) exposed to oil price and natural gas price shocks. We find that the coefficients for the brown portfolio are larger in absolute value than for the green portfolio. The strong positive effect for the brown portfolio is mainly due to the high presence of energy firms in this portfolio.

[Insert Table 6 about here.]

In Table 7, we report the estimation results when using alternative sets of controls and alternative percentile thresholds (10-90th and 40-60th percentiles of GHG intensities) for the definition of green and brown stocks. In all cases, the coefficients for brown firms are always significant and larger (in absolute terms) than the ones of green firms. Hence, the relation between firms' returns and the unexpected changes in climate change concerns seems stronger for brown firms than green firms. This effect can be explained by the observation of Bolton and Kacperczyk (2021), that institutional investors, whose stock ownership in our sample is high (see Table 5), implement exclusionary screening based on emissions intensity in a few salient industries. In Section 4.3.2, we look more into details about unexpected changes in climate change concerns in relation to industries.

[Insert Table 7 about here.]

4.3. Climate change concerns in the cross-section of stock returns

The previous section showed that the stock returns of a portfolio of firms with low (high) GHG emissions intensity are positively (negatively) associated with unexpected changes in climate change concerns. We now test whether we can recover this relation using stock-level return exposures to *UMC*. Moreover, we test whether the results still hold when considering variations in GHG emissions intensity within industries rather than across industries. Finally, we also analyze whether firms that do not disclose their GHG emissions are affected by unexpected changes in climate change concerns based on their industry and if this effect differs from firms that disclose their emissions.

4.3.1. Baseline model

We first define $lGHG_{i,t}$ as the cross-sectionally standardized logarithm of the GHG emissions intensity of firm i available at time t. The standardization is performed by focusing on the cross-sectional variation across firms. We then estimate the following firm fixed-effect panel regression model:

$$r_{i,t} = c_i + \gamma^{lGHG} lGHG_{i,t} + \left(\gamma^{UMC} + \gamma^{UMC}_{lGHG}\, lGHG_{i,t}\right) UMC_t + \boldsymbol{\beta}_i' \boldsymbol{CTRL}_t + \epsilon_{i,t}\,, \tag{9}$$

where $r_{i,t}$ is the excess stock return of firm i at time t, and $\boldsymbol{CTRL}_t$ are control factors. Again, we consider four different sets of controls in the estimation. Coefficients γ^{lGHG}, γ^{UMC}, and γ^{UMC}_{lGHG} are common to all firms, while c_i and $\boldsymbol{\beta}_i$ are firm-specific coefficients.

In specification (9), the exposure of firms to the unexpected changes in climate change concerns is $\left(\gamma^{UMC} + \gamma^{UMC}_{lGHG}\, lGHG_{i,t}\right)$, including a common component, capturing the exposure of neutral firms (*i.e.*, firms with log-GHG emissions intensity near the cross-sectional average), and one that depends on a firm's level of log-GHG emissions intensity relative to other firms. We expect a negative value for γ^{UMC}_{lGHG}, so that the higher (lower) a firm's level of GHG emissions intensity, the more negative (positive) the firm's exposure is to unexpected increases in climate change concerns. We also include the firm's log-GHG emissions intensity as a covariate to control for agents' willingness to pay more for greener firms and thus accept lower expected returns. Therefore, we expect a positive coefficient for γ^{lGHG}.

Panel regression results are reported in Table 8. For all sets of controls, we find γ^{UMC}_{lGHG} to be negative and highly significant, consistent with the model predictions. The ***CTRL-15*** model's coefficients imply that firms with a one standard deviation log-GHG emissions intensity above the cross-sectional mean have negative exposure to unexpected changes in climate change concerns of about -0.023 (*i.e.*, the sum of the coefficients of UMC and $lGHG \times UMC$). We note that the common factor UMC is not significant across different sets of controls (***CTRL-3*** being the exception). This indicates that the firm with an average level of GHG emissions intensity (*i.e.*, $lGHG = 0$) is not exposed to unexpected changes in climate change concerns. Finally, note that, when we control with the largest set of variables, there is a positive and significant coefficient for *lGHG*, suggesting that brown firms have, on average, higher returns than green firms. This result supports the baseline PST prediction that, in equilibrium, the presence of sustainability preferences leads to a return premium for investments in brown firms.

[Insert Table 8 about here.]

4.3.2. Industry analysis

In recent work, Ilhan, Sautner, and Vilkov (2021) show that most of the variation in GHG emissions intensity across firms can be attributed to their industry. Moreover, Bolton and Kacperczyk (2021) find that institutional investors implement exclusionary screening based on direct emissions intensity in a few industries. Hence, the negative relation between the exposure to unexpected changes in climate change concerns and the GHG emissions intensity may be completely explained by industry effects. To test this, we estimate model (9) for each industry separately.[26] The panel regression estimates are reported in Table 9 for the different sets of controls. We focus the interpretation on the coefficients for *UMC* (measuring the inter-industry effect) and *lGHG*$\times$ *UMC* (capturing the intra-industry effect).[27] For the inter-industry effect, we expect to find similar results as in the portfolio analysis of Section 4.2. The brownest (greenest) industries will be characterized by a negative (positive) coefficient for *UMC*. If there were no intra-industry effect, we would find a mix of positive and negative exposures of the stock returns within the same sector to the *UMC* factor (being positive for the greenest industries and negative for the brownest industries) and no significant effect of the within-industry variation modeled by the interaction effect between *lGHG* and *UMC*.

We find some confirmation of this in Table 9, where we show the results for the industries ranked from brownest to greenest. We find that for four out of the five brownest

[26] We only estimate the model for industries with more than five firms. See Table 5 for the number of companies per industry.

[27] For completeness, we also report the estimated coefficient for *lGHG*. Given the high within-industry similarity of the *lGHG* values, we find that it is insignificant for almost all industries.

industries, the coefficient on *UMC* is negative, of which two are significant. The five greenest industries have a positive coefficient, two of them being significant. For most of the other industries, the coefficient is not significant. A few industries deviate from the expected pattern: The most salient cases are "Transportation" and "Automobiles and Trucks." Despite being among the most polluting industries, they exhibit a positive and significant exposure to *UMC*. This contradicts the prediction of PST that the brownest firms have a negative price reaction to shocks in climate change concerns. We conjecture that, for these sectors, the firm's GHG is not the relevant firm characteristic driving the climate change taste-related decisions of consumers and investors. In fact, some of the firms in these sectors currently have high GHG intensities but are considered important in the transition towards a low-carbon economy and benefit from government support, notably for transport electrification. This can have a positive price impact through both the consumer and investor channel of PST.

Finally, an intra-industry effect (as shown by a significant coefficient for the interaction effect between *lGHG* and *UMC*) is only present for a few industries (*i.e.*, "Machinery," "Business Supplies," "Computers," and "Construction Materials"). For these industries, we find that the browner the firm, the more the firm tends to lose value on days with an unexpected increase in climate change concerns. Overall, we can conclude that the industry is a good predictor of firms' exposure to unexpected changes in climate change concerns.

[Insert Table 9 about here.]

4.3.3. Firms that do not disclose GHG emissions

In our sample, we find that between 28.21% (in 2018) to 44.83% (in 2009) of firms do not disclose their GHG emissions; see Table 5, Panel A. We can expect that, also for non-disclosing firms, greener firms outperform browner firms on days with high shocks in climate change concerns. To test this, we use the industry average GHG emissions intensity as a proxy for the GHG emissions intensity of the non-disclosing firms.[28] For the panel of non-disclosing and disclosing firms, the generalized model becomes:

$$r_{i,t} = c_i + \left(\gamma^{lGHG} + \delta^{lGHG}_{UD} UD_{i,t}\right) lGHG_{i,t} + \Big(\gamma^{UMC} + \left(\gamma^{UMC}_{lGHG} + \delta^{UMC}_{lGHG\text{-}UD} UD_{i,t}\right) lGHG_{i,t}\Big) UMC_t + \boldsymbol{\beta}_i' \boldsymbol{CTRL}_t + \epsilon_{i,t}\,, \quad (10)$$

where $lGHG_{i,t}$ is defined as the industry average of the logarithmic cross-sectionally normalized GHG emissions intensity for firms that do not disclose (and the actual value for firms that do report). The dummy variable $UD_{i,t}$ is equal to one if the GHG emissions intensity of firm i at time t is not disclosed, and zero otherwise. The coefficient of interest is $\delta^{UMC}_{lGHG\text{-}UD}$, which measures the difference in exposure coefficients of the interaction term between $lGHG$ and UMC for the non-disclosing vs. disclosing firms.

Estimation results are reported in Table 10. We find that the difference in exposure coefficient is not significantly different from zero between firms reporting their GHG emissions and the non-disclosing ones for which we use the industry average. This result is not surprising given that the effect of unexpected changes in climate change concern on stock returns is mostly driven by the industry. It confirms that the prediction of PST holds for all firms even if they do not disclose their GHG emissions.

[Insert Table 10 about here.]

5. Dimensions of climate change concerns

So far, we have established a relation between unexpected changes in climate change concerns and returns of greener vs. browner firms. Next, we perform two decompositions to obtain a more fine-grained understanding of the channel through which these concerns are related to the changes in the firm's stock prices.

[28] As shown in Table 5, the stocks in our sample are mainly held by institutional investors. Hence, we can assume that the typical investor is knowledgeable about the emissions profile of the stocks. As we do not find strong evidence of within-industry effects (see Section 4.3.2), it is reasonable to assume that investors impute the average GHG emissions of the industry to firms that do not disclose them.

The first decomposition is at the level of the news content captured by the climate change concerns index. The UMC variable aggregates concerns in all news articles about climate change. However, it seems self-evident that not all climate change topics are equally influential in explaining the difference in the performance of greener and browner firms. In particular, we may expect a difference across the four identified themes ("Business Impact", "Environmental Impact", "Societal Debate", and "Research") and topics that can be associated with either physical or transition climate change risk. To test this, we use in Section 5.1 the estimated topic model on our corpus of climate change news to construct topical and thematic indices of MCCC and UMC. We then use regression analysis to test for which topical and thematic risk dimensions we still find that, on days with an increase in climate change concerns about that topic or theme, there is a significant differential in stock returns explained by the GHG emissions intensity of the firm.

The second decomposition is at the level of monthly stock returns and aims at testing the implication of the PST model that the effect of climate change concerns can arise from two channels: (i) changes in expected cash flows and (ii) changes in the investor sustainability taste leading to a change in the discount factor. The empirical approach in Section 5.2 proceeds in two steps. First, we combine the price and analysts' earnings forecast data to implement the decomposition of Chen, Da, and Zhao (2013) to attribute the panel of monthly returns into their cash flow and discount rate news component. We then study how the shocks in the monthly topical and thematic climate change concerns relate to each return component.

5.1. Topical and thematic MCCC and UMC indices

The topic analysis in Section 3.1 indicates that the corpus of news articles can be summarized using 30 topics split into four themes. These topics differ in terms of prevalence and average level of climate change concerns (see Table 4). To track the heterogeneity in climate change concerns across these topics, we construct daily topical MCCC indices. The building block is the daily topic-attribution weighted concern per source s and topic k:

$$concerns_{k,t,s} = \sum_{n=1}^{N_{t,s}} \theta_{k,n,t,s}\, concerns_{n,t,s}\,, \tag{11}$$

where $\theta_{k,n,t,s}$ is obtained from the estimated CTM (see Section 3.1). We normalize and aggregate the scores for each index, following the steps of Section 2.3. This yields $K = 30$ topical MCCC indices. Moreover, we also compute four MCCC indices for the aggregate themes by weighting the concern of each article in Equation (11) using the sum of the topic-probability weights of all topics belonging to the respective theme. We then estimate the unexpected changes in climate change concerns for each topic and theme, which we denote by $UMC_{k,t}$, using the procedure outlined in Section 2.4 and Section 3.2.

In Table 11, we report the correlations between the aggregate and the four thematic UMC indices. The unconditional correlations in Panel A range from 0.47 to 0.73. The least correlated themes are "Business Impact" and "Environmental Impact," while the most correlated themes are "Environmental Impact" and "Research." Overall conclusions align with the network analysis in Figure 1. In Panel B, we report the correlations in case of large unexpected change in climate change concerns (*i.e.*, when the aggregate UMC is above its 90th percentile). In this case, the indices are much more distinct than in normal times. We can thus expect that different topics in the climate change discourse relate differently to green and brown firms' returns. In particular, some topics might be more relevant than others regarding the climate change concerns of the different economic agents (*e.g.*, customers, regulators, investors). To test this, we repeat our analysis of Section 4.3.1 with the topical and thematic UMC variables instead of the aggregate UMC variable.

[Insert Table 11 about here.]

In the left part of Table 12 (column "Daily"), we report the estimation results for the interaction term in the panel regression (9) using the largest set of controls (*i.e.*,

CTRL-15) and the various topical and thematic *UMC* indices. It is insightful to analyze the results from the dimension of climate change concerns about physical versus transition risks. We find that all themes related to transition risk have a negative and significant coefficient for $lGHG\times UMC$. On days with shocks in average concerns about transitioning to a low-carbon economy, we can thus expect that green firms outperform brown stocks. This interpretation holds for all topics within the themes "Business Impact," "Societal Debate," and "Research," except for the topic "Scientific studies." We only find a similar result for physical risk at the level of specific topics within the "Environmental Impact" theme, namely for "Hurricanes/Floods," "Glaciers/Ice Sheets," and "Tourism." Understanding the market response around concerns about physical risk thus requires a more fine-grained approach disentangling the market-relevant topics from others.

[Insert Table 12 about here.]

5.2. Cash flow and discount rate channels

When concerns about climate change strengthen, the PST model predicts that green firms will gain in popularity among consumers and investors. Through the consumer channel, green firms enjoy an increase in their net cash flows to the detriment of brown firms. As there is also a strengthening of the investor preferences for owning green firms rather than brown firms, the required return for investing in green (brown) firms decreases (increases). This investor channel leads to a reduction in the discount rate of green firms relative to the discount rate of brown firms. An interesting question is how important these two channels are. Additional model assumptions are needed for identification. Chen, Da, and Zhao (2013) propose an approach that requires observing the earnings forecasts revisions, which is not feasible for daily return data. Therefore, we perform the remaining analysis on monthly returns and apply the decomposition of Chen, Da, and Zhao (2013) on firms' monthly capital-gain returns using the implied cost of capital model of Gebhardt, Lee, and Swaminathan (2001).

Formally, denote for firm i at month τ the capital-gain return by $retx_{i,\tau}$. The Chen, Da, and Zhao (2013) decomposition uses analysts' earnings forecasts and firm's accounting data to decompose $retx_{i,\tau}$ into the sum of a cash flow news component, $CF_{i,\tau}$, and discount rate news component, $DR_{i,\tau}$. We refer to the Appendix, Section F, for more details on the decomposition and our implementation. The decomposition requires analysts' earnings forecasts and firm's accounting data that we retrieve from IBES and Compustat, respectively.[29]

To obtain insight into the relative importance of the cash flow and discount channels on the relation between climate change concerns and stock performance, we use a panel regression that models the stock capital-gain return as a function of an interaction effect between the firm's level of GHG emissions intensity and the shock in concern about climate change of that month. The monthly UMC is obtained by first computing monthly MCCC indices following the same methodology of Section 2, but we aggregate at the monthly frequency. We estimate the unexpected monthly changes in climate change concerns using Equation (7) and a rolling estimation window of 60 months. We estimate the panel regression (9) using $retx_{i,\tau}$, $DR_{i,\tau}$, and $CF_{i,\tau}$ as the left-sided variable and the various UMC indices, namely the aggregate, the thematic, and topical indices. As controls, we use ***CTRL-6*** alongside the first three principal components of the remaining variables in ***CTRL-15*** (*i.e.*, excluding the one in ***CTRL-6***).[30] We use the principal components instead of including the full set of control in ***CTRL-15*** due to the limited sample size of this monthly analysis compared to our previous daily analysis.

In the right part of Table 12 (columns "Monthly"), we report the interaction term estimates of panel regression (9) for monthly returns, cash flow news components, and discount rate components when using the aggregate, the thematic, and the topical UMC variables. First, at the aggregate and thematic levels, we see that results for the monthly capital-gain returns are consistent with the results of daily returns in the left part, except

[29] We require that a firm has at least 12 valid monthly observations. An observation is discarded when it is an extreme correlation outlier, implemented as $|CF_{i,\tau}| + |DR_{i,\tau}| > 4 \times |retx_{i,\tau}|$, or when the input accounting or IBES data is missing. A manual check shows that excluding the correlation outliers safeguards our analysis against anomalous earnings forecasts leading to unreliable estimates.

[30] These three principal components explain about 66% of the variation of the remaining variables in ***CTRL-15***.

for the theme "Business Impact."[31] Also, most of the significant terms are negative and related to climate change transition risk at the topics level.

Focusing on the interaction term estimates for the CF and DR news components, we find that the discount rate channel dominates over the cash flow channel in terms of significance. At the aggregate level, the DR news coefficient is negative and significant.[32] At the thematic (topic) level, the DR news coefficient for "Societal Debate" (2 topics out of 4) and "Research" (2 topics out of 3) are negative and significant. For "Business Impact," 2 topics out of 10 are significant and negative for the discount rate channel. Additionally, 1 topic out of 11 is negative and significant for the theme "Environmental Impact." Moreover, even accounting for non-significant results, the cash flow channel only dominates (in absolute value) in 5 out of the 30 topics.

Overall, our results suggest that, for monthly returns, the discount rate channel is the primary channel where the interaction between stock returns, GHG emissions intensity, and unexpected changes in climate change concerns arise. This finding has important consequences for capital budgeting. It implies that green firms enjoy a reduction in the cost of equity in periods of high unexpected climate change concerns, and vice versa for brown firms. A caveat of our analysis on the cash flow channel of PST is that the cash flow effects of news are notoriously difficult to observe using monthly returns. Indeed, from the study of Chen, Da, and Zhao (2013), it can be expected that the cash flow effect manifests itself predominantly on longer horizon returns, such as yearly returns or longer. This could be analyzed using distributed lag models, which is beyond the scope of our paper.

[31] We note that for the daily results, the entire cross-section of the S&P 500 universe is used since the analysis is not limited by the availability of earnings forecast data.

[32] As a robustness exercise, we used the Engle et al. (2020) WSJ and Crimson Hexagon index as a replacement for our aggregate index. While no coefficients are significant for the WSJ index, the Crimson Hexagon index resulted in a negative and significant coefficient for the DR news component, consistent with our result.

6. Conclusion

Our paper empirically verifies the prediction of Pástor, Stambaugh, and Taylor (2021) that green firms outperform brown firms when climate change concerns increase unexpectedly.

Our first contribution is to construct a daily proxy that captures unexpected increases in climate change concerns. We do this by collecting news articles published about climate change from major U.S. newspapers and newswires from 2003 to 2018. We design an article-level concerns score, and aggregate these scores daily across newspapers to obtain our Media Climate Change Concerns (MCCC) index, which proxies for changes in climate change concerns. We show that our index captures several key climate change events that are likely to increase concerns about climate change. Then, we obtain unexpected changes (UMC) as the shock component in our MCCC index filtered from potential effects of financial-market, energy-related, and macroeconomic variables. Combining the index construction framework with a topic model, we obtain topical and thematic UMC variables that we associate with climate change transition and physical risk.

Our second contribution is to show that unexpected changes in climate change concerns help explain differences in the performance of green and brown stocks from 2010 to 2018, where greenness is measured by a firm's greenhouse gas emissions intensity. Multiple analyses lead to the same conclusion: All things being equal, green firms outperform brown firms when there are unexpected increases in climate change concerns.

Our third contribution is to shed light on the channels through which stock returns relate to these shocks in climate change concerns. First, we find that, in the cross-section of firm returns, the conditional exposure to shocks in climate change concerns is, for most industries, the same for firms belonging to the same industry. Second, we investigate whether the exposure to UMC also holds for firms that do not disclose their greenhouse gas emissions. Using the industry average as a proxy for their emissions, we find it is the case. Third, we use a correlated topic model on our news corpus to investigate which types of climate change concerns relate to the performance of green versus brown stocks. We show that there is significant exposure to almost all topics discussing climate change transition risk, but only a subset of physical risk topics explain the performance of green versus brown firms. Finally, we find that high unexpected changes in climate change concerns increase (decrease) the discount factor of brown (green) firms but do not find evidence of a cash flow effect.

A key message for business leaders is that climate change concerns also matter for their firms' equity values and, importantly, that they can manage their exposure by altering their greenhouse gas emissions intensity. As climate change concerns and investor preferences are time-varying, a monitoring system is recommended. The monitoring of thematic news complements the current widespread practice of monitoring reputation in the media (*e.g.*, Fombrun, Ponzi, and Newburry, 2015). In this paper, we propose a first design for such a system using U.S. media news.

References

Alekseev, G., S. Giglio, Q. Maingi, J. Selgrad, and J. Stroebel. 2021. A quantity-based approach to constructing climate risk hedge portfolios. Working Paper.

Ardia, D., K. Bluteau, and K. Boudt. 2019. Questioning the news about economic growth: Sparse forecasting using thousands of news-based sentiment values. *International Journal of Forecasting* 35:1370–86.

Baele, L., G. Bekaert, K. Inghelbrecht, and M. Wei. 2020. Flights to safety. *Review of Financial Studies* 33:689–746.

Baker, S. R., N. Bloom, and S. J. Davis. 2016. Measuring economic policy uncertainty. *Quarterly Journal of Economics* 131:1593–636.

Ball-Rokeach, S. J., and M. L. DeFleur. 1976. A dependency model of mass-media effects. *Communication Research* 3:3–21.

Ballinari, D., and O. Mahmoud. 2021. From chatter to action: An index of sustainability sentiment. Working paper.

Bansal, R., D. A. Wu, and A. Yaron. 2021. Socially responsible investing in good and bad times. *Review of Financial Studies* 35:2067–99.

Berg, F., J. F. Koelbel, and R. Rigobon. 2019. Aggregate confusion: The divergence of ESG ratings. Working Paper.

Bertolotti, A., D. Basu, K. Akallal, and B. Deese. 2019. Climate risk in the US electric utility sector: A case study. Working Paper.

Bessec, M., and J. Fouquau. 2021. A green wave in media, a change of tack in stock markets. Working Paper.

Blei, D. M., A. Y. Ng, and M. I. Jordan. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research* 3:993–1022.

Bolton, P., and M. Kacperczyk. 2021. Do investors care about carbon risk? *Journal of Financial Economics* 142:517–49.

Boudt, K., and M. Petitjean. 2014. Intraday liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. *Journal of Financial Markets* 17:121–49.

Boykoff, M. T., and J. M. Boykoff. 2007. Climate change and journalistic norms: A case-study of US mass-media coverage. *Geoforum* 38:1190–204.

Bua, G., D. Kapp, F. Ramella, and L. Rognone. 2022. Transition versus physical climate risk pricing in European financial markets: A text-based approach. Working Paper.

Campos-Martins, S., and D. Hendry. 2021. Geoclimate, geopolitics, and the geovolatility of carbon-intensive equity returns. Working paper.

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

Chen, L., Z. Da, and X. Zhao. 2013. What drives stock price movements? *Review of Financial Studies* 26:841–76.

Choi, D., Z. Gao, and W. Jiang. 2020. Attention to global warming. *Review of Financial Studies* 33:1112–45.

Chong, D., and J. N. Druckman. 2007. Framing theory. *Annual Review of Political Science* 10:103–26.

Da, Z., Q. Liu, and E. Schaumburg. 2014. A closer look at the short-term return reversal. *Management Science* 60:658–74.

Doyle, C., S. Sreenivasan, B. K. Szymanski, and G. Korniss. 2016. Social consensus and tipping points with opinion inertia. *Physica A* 443:316–23.

Drempetic, S., C. Klein, and B. Zwergel. 2020. The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics* 167:333–60.

Engle, R. F., S. Giglio, B. Kelly, H. Lee, and J. Stroebel. 2020. Hedging climate change news. *Review of Financial Studies* 33:1184–216.

Faccini, R., R. Matin, and G. Skiadopoulos. 2021. Are climate change risks priced in the US stock market? Working Papers.

Fama, E. F., and K. R. French. 1992. The cross-section of expected stock returns. *Journal of Finance* 47:427–65.

———. 1997. Industry costs of equity. *Journal of Financial Economics* 43:153–93.

———. 2007. Disagreement, tastes, and asset prices. *Journal of Financial Economics* 83:667–89.

———. 2015. A five–factor asset pricing model. *Journal of Financial Economics* 116:1–22.

Flaxman, S., S. Goel, and J. M. Rao. 2016. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly* 80:298–320.

Fombrun, C. J., L. J. Ponzi, and W. Newburry. 2015. Stakeholder tracking and analysis: The RepTrak® system for measuring corporate reputation. *Corporate Reputation Review* 18:3–24.

Fung, W., and D. Hsieh. 2004. Hedge funds benchmarks: A risk-based approach. *Financial Analyst Journal* 60:65–80.

Gebhardt, W. R., C. M. Lee, and B. Swaminathan. 2001. Toward an implied cost of capital. *Journal of Accounting Research* 39:135–76.

Gentzkow, M., B. Kelly, and M. Taddy. 2019. Text as data. *Journal of Economic Literature* 57:535–74.

Gentzkow, M., and J. M. Shapiro. 2006. Media bias and reputation. *Journal of Political Economy* 114:280–316.

Gerlagh, R., and M. Liski. 2018. Carbon prices for the next hundred years. *Economic Journal* 128:728–57.

Giglio, S., B. Kelly, and J. Stroebel. 2021. Climate finance. *Annual Review of Financial Economics* 13:15–36.

Görgen, M., A. Jacob, M. Nerlinger, R. Riordan, M. Rohleder, and M. Wilkens. 2020. Carbon risk. Working Paper.

GSIA. 2018. Global Sustainable Investment Review. Technical Report, Global Sustainable Investment Alliance.

Halcoussis, D., and A. D. Lowenberg. 2019. The effects of the fossil fuel divestment campaign on stock returns. *North American Journal of Economics and Finance* 47:669–74.

Hale, B. W. 2010. Using newspaper coverage analysis to evaluate public perception of management in river-floodplain systems. *Environmental Management* 45:1155–63.

Hansen, S., M. McMahon, and A. Prat. 2018. Transparency and deliberation within the FOMC: A computational linguistics approach. *Quarterly Journal of Economics* 133:801–70.

Hong, H., F. W. Li, and J. Xu. 2019. Climate risks and market efficiency. *Journal of Econometrics* 208:265–81.

Ilhan, E., Z. Sautner, and G. Vilkov. 2021. Carbon tail risk. *Review of Financial Studies* 34:1540–71.

Kapfhammer, F., V. Larsen, and L. A. Thorsrud. 2020. Climate risk and commodity currencies. Working Paper.

Lafferty, J. D., and D. M. Blei. 2006. Correlated topic models. In *Advances in Neural Information Processing Systems*, 147–54.

Larsen, V. H. 2021. Components of uncertainty. *International Economic Review* 62:769–88.

Larsen, V. H., and L. A. Thorsrud. 2017. Asset returns, news topics, and media effects. Working Paper, Forthcoming in Scandinavian Journal of Economics.

Martin, F., and M. Johnson. 2015. More efficient topic modelling through a noun only approach. In *Proceedings of the Australasian Language Technology Association Workshop 2015*, 111–5.

McCombs, M. E., and D. L. Shaw. 1972. The agenda–setting function of mass media. *Public Opinion Quarterly* 36:176–87.

Mimno, D., H. Wallach, E. Talley, M. Leenders, and A. McCallum. 2011. Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 262–72.

Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:703–8.

———. 1994. Automatic lag selection in covariance matrix estimation. *Review of Economic Studies* 61:631–53.

NGFS. 2020. Guide for supervisors: Integrating climate-related and environmental risks into prudential supervision. Technical Report, Network of Central Banks and Supervisors for Greening the Financial System.

Nimark, K. P., and S. Pitschner. 2019. News media and delegated information choice. *Journal of Economic Theory* 181:160–96.

Pástor, L., R. F. Stambaugh, and L. A. Taylor. 2021. Sustainable investing in equilibrium. *Journal of Financial Economics* 142:550–71.

———. 2022. Dissecting green returns. Working paper.

Pennebaker, J. W., R. L. Boyd, K. Jordan, and K. Blackburn. 2015. The development and psychometric properties of LIWC2015. Technical Report, University of Texas at Austin.

Ramelli, S., A. F. Wagner, R. J. Zeckhauser, and A. Ziegler. 2021. Investor rewards to climate responsibility: Stock-price responses to the opposite shocks of the 2016 and 2020 U.S. elections. *Review of Corporate Finance Studies* 10:748–87.

Roberts, M. E., B. M. Stewart, and D. Tingley. 2014. **stm**: R package for structural topic models. *Journal of Statistical Software* 10:1–40.

Sampei, Y., and M. Aoyagi-Usui. 2009. Mass-media coverage, its influence on public awareness of climate-change issues, and implications for Japan's national campaign to reduce greenhouse gas emissions. *Global Environmental Change* 19:203–12.

Schoenfeld, A. C., R. F. Meier, and R. J. Griffin. 1979. Constructing a social problem: The press and the environment. *Social Problems* 27:38–61.

Scruggs, L., and S. Benegal. 2012. Declining public concern about climate change: Can we blame the great recession? *Global Environmental Change* 22:505–15.

Slovic, P. 1986. Informing and educating the public about risk. *Risk Analysis* 6:403–15.

Stecula, D. A., and E. Merkley. 2019. Framing climate change: Economics, ideology, and uncertainty in American news media content from 1988 to 2014. *Frontiers in Communication* 4:1–15.

Table 1: Media-based time series for understanding the impact of climate change on the financial markets This table reports the list of media-based indices constructed by scholars to understand the impact of climate change on the financial markets. Column order next to "This paper" is with respect to the first release of the methodology on SSRN.

	This paper	Engle et al.		Kapfhammer et al.	Faccini et al.		Bessec and Fouquau	Bua et al.
Index aims at capturing	Aggregate and thematic media climate change concerns (MCCC)	Intensity of climate change news	Intensity of negative climate change news	Country-specific climate change transition risks	Attention to natural disaster, global warming, international summits and U.S. climate policy	Intensity of transition risks elicited by the U.S. political debate	Media attention, tonality, and uncertainty about environment issues	Intensity of climate change news about physical risk (PR) and transition risk (TR) of climate change
Corpus of news media sources	Ten newspapers: (i) New York Times, (ii) Washington Post, (iii) Los Angeles Times, (iv) Wall Street Journal, (v) Houston Chronicle, (vi) Chicago Tribune, (vii) Arizona Republic, (viii) USA Today, (ix) New York Daily News, (x) New York Post. Two newswires: English articles on (i) Associated Press Newswires and (ii) Reuters News.	Wall Street Journal	Calculation by Crimson Hexagon (CH). Full list not disclosed. The starting corpus has over 1,000 outlets, including New York Times, Washington Post, Wall Street Journal, Reuters, BBC, CNN	Dow Jones corpus (news articles written in English). Includes Wall Street Journal	English articles on Reuters news		Four newspapers: (i) New York Times, (ii) USA Today, (iii) Washington Post, (iv) Wall Street Journal	English articles on Reuters news
Articles' selection	Articles tagged as "climate change" by the publisher are retained. Keyword based filters are used to eliminate articles discussing stock market performance.	None	Articles containing "climate change"	None	Article containing "climate change" or "global warming"	Articles with "domestic policy" prevalence larger than 40%	Articles tagged as "Commodity/financial market news" and "Economic news"	Only news with European regional focus
Relevant feature extraction per article	Measures concerns as a combination of attention (based on number of articles), polarity (using negative words dictionary), and uncertainty (using uncertainty dictionary)	Measures attention as cosine similarity between tf-idf of WSJ editions and a "climate change vocabulary"	Measures the share of all news articles that are both about "climate change" and that have been assigned to a "negative sentiment" score	Measures association between "country" and "climate risk" both expressed in a word2vec 100-dimensional space calibrated to the monthly corpus. Eight countries are considered.	Measures attention using topic shares	Measures transition risk by manual scoring of articles	Measures attention based on dictionary environmental dictionary. Extend with tonality and uncertainty using LM dictionary	Same approach as WSJ index by Engle et al. (2020) but with one physical and one transition risk vocabulary
Sources' aggregation	Equally-weighted of normalized indices per source	N.A.	Pooled	N.A.	N.A.		Pooled	N.A.
Frequency	Daily and monthly	Monthly	Monthly	Monthly	Daily	Monthly	Weekly	Daily
Extraction of shock	Augmented AR model	AR model	AR model	N.A.	N.A.	N.A.	N.A.	AR model
Available for download	Yes	Yes	Yes	No	Yes	Yes	No	No
Correlation with MCCC aggregate index		0.37 for MCCC and 0.44 for UMC	0.23 for MCCC and 0.43 for UMC		Ranges from 0.11 to 0.30 for MCCC and from 0.09 to 0.22 for UMC	-0.40 for MCCC and -0.13 for UMC		

Table 2: Sources of climate change news

This table reports, for each source (Panel A newswires and Panel B newspapers), the number and the percentage of articles discussing climate change from January 2003 to June 2018. The table also reports the average concern score and the percentage of time a concern score is zero. The last four columns report the percentage of themes' prevalence obtained with the correlated topic model (BI: Business Impact, EI: Environmental Impact, SD: Societal Debate, R: Research).

Panel A: Newswires

	Articles		Concern scores		Themes			
Source	*N*	%	Mean	%0	BI	EI	SD	R
Associated Press Newswires	10,061	0.07	0.32	12.97	49.58	22.39	16.97	11.06
Reuters News	9,288	0.08	0.37	11.79	62.43	17.46	10.47	9.64

Panel B: Newspapers

	Articles		Concern scores		Themes			
Source	*N*	%	Mean	%0	BI	EI	SD	R
New York Times	3,472	0.25	0.38	3.77	41.21	22.16	26.14	10.49
Washington Post	2,442	0.25	0.35	5.45	44.30	20.16	23.68	11.86
Los Angeles Times	1,530	0.20	0.39	3.59	36.76	24.06	29.66	9.52
Wall Street Journal	1,412	0.26	0.32	6.23	56.50	11.55	21.09	10.86
Houston Chronicle	1,385	0.16	0.33	9.60	51.17	18.84	18.82	11.17
Chicago Tribune	482	0.03	0.35	6.85	31.04	24.21	34.01	10.74
Arizona Republic	382	0.04	0.31	14.40	36.88	25.18	26.11	11.83
USA Today	234	0.08	0.44	6.41	30.63	27.78	26.98	14.60
New York Daily News	122	0.02	0.43	9.02	32.06	17.82	41.76	8.35
New York Post	111	0.02	0.44	9.91	35.88	11.69	44.24	8.18

Table 3: List of topics together with top ten keywords in terms of probability

This table lists the 30 topics identified in our corpus together with the ten keywords with highest probability for each topic. Topics are regrouped into four themes. For each theme, the topics are sorted by their unconditional prevalence; see Table 4.

Theme 1: Business Impact Topic	Top ten keywords in terms of probability
Climate Summits	agreement, country, climate change, nation, world, talk, deal, meeting, develop country, summit
Agreements/Actions	percent, emission, level, target, greenhouse gas emission, goal, country, government, greenhouse gas, year
Climate Legislation/Regulations	bill, state, cap, legislation, vote, lawmaker, measure, program, global warming, year
Legal Actions	state, administration, rule, regulation, agency, plan, court, decision, law, case
Renewable Energy	oil, energy, natural gas, gas, pipeline, fossil fuel, renewable energy, wind, nuclear power, world
Carbon Reduction Technologies	coal, plant, power plant, electricity, carbon dioxide, technology, power, utility, gas, year
Carbon Credits Market	market, price, scheme, government, credit, euro, tonne, carbon, year, permit
Carbon Tax	cost, tax, carbon, energy, price, policy, fuel, carbon tax, biofuel, economy
Government Programs	project, money, fund, program, year, development, government, budget, funding, plan
Corporations/Investments	company, business, climate change, investor, group, investment, firm, industry, risk, chief executive
Car Industry	car, vehicle, standard, methane, gas, year, fuel, industry, automaker, carbon dioxide
Airline Industry	airline, flight, ship, emission, aviation, plane, air, pollution, shipping, aircraft
Theme 2: Environmental Impact Topic	**Top ten keywords in terms of probability**
Extreme Temperatures	year, record, weather, temperature, winter, day, summer, climate change, heat, global warming
Food Shortage/Poverty	climate change, people, crop, country, farmer, world, food, woman, agriculture, foundation
Hurricanes/Floods	flood, storm, hurricane, climate change, sea level, island, disaster, damage, flooding, risk
Glaciers/Ice Sheets	ice, glacier, year, scientist, foot, ice sheet, mile, melting, sea ice, satellite
Ecosystems	species, animal, plant, bird, disease, climate change, population, year, habitat, extinction
Forests	tree, forests, forest, fire, land, deforestation, carbon, acre, area, soil
Water/Drought	water, state, region, river, rivers, drink, year, lake, area, dam
Tourism	site, town, day, mountain, year, snow, mile, park, foot, people
Arctic Wildlife	polar bear, sea ice, bear, seal, ice, habitat, species, wildlife, year, population
Marine Wildlife	fish, water, sea, oceans, ocean, scientist, coral, alga, year, reef
Agriculture Shifts	food, farm, year, wine, plant, meat, production, farmer, coffee, cow
Theme 3: Societal Debate Topic	**Top ten keywords in terms of probability**
Political Campaign	climate change, issue, leader, president, campaign, election, party, country, speech, policy
Social Events	people, world, time, life, climate change, child, year, student, book, global warming
Controversies	climate change, science, global warming, scientist, climate, issue, question, evidence, research, document
Cities	city, people, building, home, energy, light, resident, community, mayor, group
Theme 4: Research Topic	**Top ten keywords in terms of probability**
Global Warming	degree, global warming, warming, world, scientist, year, carbon dioxide, atmosphere, greenhouse gas, century
UN/IPCC Reports	report, climate change, risk, impact, global warming, panel, effect, government, world, study
Scientific Studies	study, research, scientist, researcher, data, atmosphere, researchers, climate, effect, model

Table 4: Topics' unconditional prevalence, climate change concerns score, and climate risk
This table reports the 30 topics' unconditional prevalence $\bar{\theta}$, average climate change concerns score $\overline{CC}$, and type of climate risk (physical risk, transition risk, or liability risk) following NGFS (2020).

	$\bar{\theta}$	$\overline{CC}$	Climate risk
Theme 1: Business Impact	51.13	0.30	Transition
Climate Summits	11.02	0.32	Transition
Agreements/Actions	6.17	0.28	Transition
Climate Legislation/Regulations	5.28	0.27	Transition
Legal Actions	5.11	0.37	Liability
Renewable Energy	3.81	0.29	Transition
Carbon Reduction Technologies	3.80	0.23	Transition
Carbon Credits Market	3.43	0.23	Transition
Carbon Tax	3.03	0.30	Transition
Government Programs	2.95	0.33	Transition
Corporations/Investments	2.89	0.33	Transition
Car Industry	2.44	0.29	Transition
Airline Industry	1.20	0.32	Transition
Theme 2: Environmental Impact	20.18	0.45	Physical
Extreme Temperatures	3.27	0.34	Physical
Food Shortage/Poverty	2.51	0.62	Physical
Hurricanes/Floods	2.39	0.70	Physical
Glaciers/Ice Sheets	2.35	0.31	Physical
Ecosystems	1.75	0.42	Physical
Forests	1.63	0.39	Physical
Water/Drought	1.55	0.44	Physical
Tourism	1.41	0.36	Physical
Arctic Wildlife	1.18	0.58	Physical
Marine Wildlife	1.15	0.42	Physical
Agriculture Shifts	0.98	0.28	Physical
Theme 3: Societal Debate	18.13	0.35	Transition
Political Campaign	6.01	0.36	Transition
Social Events	4.69	0.35	Transition
Controversies	4.68	0.39	Transition
Cities	2.76	0.28	Transition
Theme 4: Research	10.56	0.40	Physical/Transition
Global Warming	3.80	0.37	Physical/Transition
UN/IPCC Reports	3.50	0.53	Physical/Transition
Scientific Studies	3.26	0.31	Physical/Transition

Table 5: Statistics of GHG intensities and disclosures, and institutional ownership
This table reports statistics of the greenhouse gas intensities used to establish firms' greenness and brownness. Panel A reports the percentage of firms in the S&P 500 universe with available greenhouse gas emissions data for each year (%D). Panel B reports summary statistics for the industries as defined by Fama and French (1997); note that "Fabricated Products" is missing in our sample. For each industry, the table reports the number of companies (N), the average and standard deviation of greenhouse gas intensities, the percentage of firms disclosing their emissions (%D), and the percentage of institutional ownership for firms that disclose (D) or do not disclose (ND) their emissions.

Panel A: Percentage of firms with emissions data

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017	2009-2017 Mean
%D	56.73	62.17	65.24	68.12	60.98	60.79	62.71	66.17	70.83	63.75

Panel B: Summary for the industries

		GHG intensity		GHG disclosure	Inst. ownership	
Industry	N	Mean	Std	%D	D	ND
Utilities	43	4,072.83	2,271.65	80.00	70.54	71.72
Coal	2	3,034.88	3,761.48	83.33	–	–
Other	9	2,142.59	2,347.55	60.00	77.93	85.02
Steel Works Etc	6	1,903.09	1,211.71	24.24	79.41	68.10
Chemicals	14	1,088.50	1,220.14	82.08	81.21	79.74
Petroleum and Natural Gas	34	1,033.31	1,462.40	69.41	79.65	89.54
Precious Metals	1	776.53	306.53	100.00	81.56	–
Non-Metallic and Industrial Metal Mining	4	759.58	712.83	82.14	83.46	96.09
Consumer Goods	14	679.94	1,645.07	86.32	78.66	72.41
Shipping Containers	3	675.60	456.54	83.33	78.41	–
Automobiles and Trucks	6	621.34	1,274.14	86.96	72.34	80.87
Transportation	16	615.70	375.45	77.68	76.61	79.84
Personal Services	5	554.01	671.24	42.86	90.98	89.35
Machinery	15	529.70	2,008.38	63.64	79.87	81.45
Meals Restaurants, Hotels, Motels	8	500.93	753.12	76.92	73.03	74.75
Defense	1	410.60	617.60	100.00	83.98	–
Business Supplies	9	380.62	270.31	91.18	77.77	92.12
Textiles	1	343.13	24.14	60.00	80.54	77.12
Beer and Liquor	6	335.49	413.82	93.18	75.61	85.54
Candy and Soda	5	263.01	244.26	45.16	62.88	56.03
Food Products	19	209.12	283.15	79.86	66.82	70.82
Agriculture	1	207.67	37.09	90.00	82.13	73.24
Rubber and Plastic Products	2	117.29		50.00	87.15	85.60
Computers	16	115.83	176.02	72.64	83.25	85.29
Electronic Equipment	32	109.75	163.45	68.58	80.57	88.47
Pharmaceutical Products	29	99.97	156.82	77.60	75.78	91.71
Tobacco Products	4	88.81	87.21	82.35	62.33	93.15
Wholesale	11	75.62	402.32	44.19	80.65	77.63
Medical Equipment	17	73.44	137.29	53.21	84.81	78.29
Communication	21	69.68	62.67	43.90	63.45	83.53
Apparel	8	64.48	117.47	55.56	84.50	71.67
Entertainment	4	60.54	15.48	12.50	78.98	79.13
Retail	46	57.49	90.43	51.24	75.89	85.46
Construction Materials	6	52.57	34.24	52.78	83.81	87.22
Business Services	55	49.04	135.39	53.55	78.95	86.59
Real Estate	2	42.90	40.28	95.00	–	–
Aircraft	9	41.29	43.11	74.14	76.55	84.63
Recreation	3	39.77	40.36	85.71	88.45	95.42
Measuring and Control Equipment	15	37.38	45.54	63.74	88.28	88.10
Healthcare	8	37.02	5.40	26.00	87.16	93.04
Printing and Publishing	5	36.63	15.02	50.00	100.00	67.19
Electrical Equipment	7	33.18	14.49	83.33	79.21	81.66
Banking	31	21.91	16.24	56.02	78.87	81.08
Trading	19	10.39	8.87	49.64	73.20	77.26
Insurance	28	5.08	3.94	61.93	79.35	75.12
Construction	7	4.45	2.15	24.59	83.48	84.23
Shipbuilding, Railroad Equipment	1	–	–	0.00	–	–
Across-industry mean	12.94	488.75	537.4	64.82	79.27	81.59
Across-industry 25th percentile	4	49.92	40.28	50.62	76.38	77.12
Across-industry 75th percentile	17	600.28	671.24	82.84	83.47	87.22

Table 6: Regression results of daily portfolios' returns

This table reports the results of regressing the daily returns of green-minus-brown (GMB), green, brown, and neutral portfolios on the contemporaneous daily unexpected changes in climate change concerns (UMC) and the daily values of the control variables $\boldsymbol{CTRL\text{-}15}$; see model (8). The composition of the four portfolios is based on greenhouse gas intensities. Newey and West (1987, 1994) standard errors of the estimators are reported in parentheses. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018.

	GMB	Green	Brown	Neutral
Intercept	0.068*	0.019	−0.049*	0.019
	(0.038)	(0.019)	(0.028)	(0.014)
UMC	0.072**	0.029**	−0.042*	0.006
	(0.031)	(0.014)	(0.023)	(0.01)
MKT	0.127***	1.1***	0.973***	1.036***
	(0.016)	(0.008)	(0.013)	(0.005)
HML	0.112***	0.178***	0.066***	−0.095***
	(0.036)	(0.018)	(0.024)	(0.011)
SMB	0.071***	0.016	−0.055***	0.017**
	(0.026)	(0.011)	(0.02)	(0.008)
CMA	−0.462***	−0.08***	0.382***	0.231***
	(0.048)	(0.028)	(0.035)	(0.016)
RMW	−0.296***	−0.122***	0.174***	0.137***
	(0.04)	(0.019)	(0.031)	(0.013)
MOM	0.078***	−0.088***	−0.165***	−0.076***
	(0.023)	(0.011)	(0.017)	(0.007)
WTI	−8.458***	−2.956***	5.502***	0.124
	(0.695)	(0.336)	(0.51)	(0.221)
NG	−0.361	−0.065	0.296	−0.099
	(0.27)	(0.113)	(0.208)	(0.074)
$PROP$	−1.252***	−0.404	0.848**	−0.148
	(0.482)	(0.251)	(0.351)	(0.135)
EPU	0.013	−0.005	−0.018	−0.016***
	(0.019)	(0.009)	(0.013)	(0.006)
VIX	−0.003	0.001	0.004***	0.001
	(0.002)	(0.001)	(0.001)	(0.001)
TED	−0.063	−0.069*	−0.006	−0.064**
	(0.081)	(0.041)	(0.062)	(0.027)
$TERM$	3.324***	1.122***	−2.202***	0.217*
	(0.366)	(0.164)	(0.273)	(0.111)
$DFLT$	0.291	0.146	−0.145	0.145*
	(0.211)	(0.103)	(0.159)	(0.082)
FTS	0.078	−0.044	−0.122	−0.002
	(0.125)	(0.057)	(0.097)	(0.053)

Table 7: Regression results of daily portfolios' returns – Alternative controls and setups
This table reports the results of regressing the daily returns of green-minus-brown (GMB), green, brown, and neutral portfolios on the contemporaneous daily unexpected changes in climate change concerns (*UMC*) for different sets of controls (***CTRL-1***, ***CTRL-3***, ***CTRL-6***, ***CTRL-15***) and green/brown stock classifications (Panel A: 25-75th percentiles, Panel B: 10-90th percentiles, Panel C: 40-60th percentiles of the greenhouse gas emissions intensity). Newey and West (1987, 1994) standard errors of the estimators are reported in parentheses. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018.

Panel A: 25-75th percentiles				
	UMC exposure			
	GMB	Green	Brown	Neutral
CTRL-1	0.085^{**}	0.03^{*}	-0.056^{*}	0.004
	(0.038)	(0.015)	(0.029)	(0.012)
CTRL-3	0.086^{**}	0.03^{**}	-0.056^{**}	0.004
	(0.036)	(0.015)	(0.026)	(0.011)
CTRL-6	0.081^{**}	0.031^{**}	-0.05^{**}	0.006
	(0.034)	(0.015)	(0.024)	(0.01)
CTRL-15	0.072^{**}	0.029^{**}	-0.042^{*}	0.006
	(0.031)	(0.014)	(0.023)	(0.01)
Panel B: 10-90th percentiles				
	UMC exposure			
	GMB	Green	Brown	Neutral
CTRL-1	0.113^{**}	0.039	-0.074^{*}	−0.001
	(0.056)	(0.026)	(0.039)	(0.011)
CTRL-3	0.113^{**}	0.038	-0.075^{**}	−0.001
	(0.051)	(0.025)	(0.037)	(0.011)
CTRL-6	0.111^{**}	0.04	-0.071^{*}	0.002
	(0.052)	(0.025)	(0.037)	(0.009)
CTRL-15	0.116^{**}	0.041^{*}	-0.075^{**}	0.004
	(0.05)	(0.024)	(0.037)	(0.009)
Panel C: 40-60th percentiles				
	UMC exposure			
	GMB	Green	Brown	Neutral
CTRL-1	0.061^{**}	0.023^{*}	-0.038^{*}	0.006
	(0.029)	(0.013)	(0.023)	(0.017)
CTRL-3	0.061^{**}	0.023^{*}	-0.038^{*}	0.006
	(0.028)	(0.013)	(0.021)	(0.015)
CTRL-6	0.057^{**}	0.025^{**}	-0.032^{*}	0.007
	(0.028)	(0.012)	(0.019)	(0.015)
CTRL-15	0.05^{**}	0.023^{*}	−0.027	0.006
	(0.025)	(0.012)	(0.018)	(0.015)

Table 8: Panel regression results of daily firms' returns

This table reports the estimation results for the firm fixed-effect panel regression of daily stock returns on daily standardized logarithmic GHG emissions intensity ($lGHG$), daily unexpected changes in climate change concerns (UMC), and their interaction ($lGHG \times UMC$); see model (9). We use four different sets of controls (***CTRL-1***, ***CTRL-3***, ***CTRL-6***, ***CTRL-15***). Standard errors of the estimators are reported in parentheses. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018.

	$lGHG$	UMC	$lGHG \times UMC$
CTRL-1	0.005^{*}	-0.006	-0.027^{***}
	(0.003)	(0.004)	(0.005)
CTRL-3	0.004	-0.008^{**}	-0.024^{***}
	(0.003)	(0.004)	(0.005)
CTRL-6	0.005	-0.005	-0.023^{***}
	(0.003)	(0.004)	(0.005)
CTRL-15	0.007^{**}	-0.002	-0.021^{***}
	(0.003)	(0.004)	(0.005)

Table 9: Panel regression results of daily firms' returns – Industries

This table reports the estimation results for the fixed-effect panel regression of daily stock returns on daily standardized logarithmic GHG emissions intensity ($lGHG$), daily unexpected changes in climate change concerns (UMC), and their interaction ($lGHG \times UMC$); see model (9). One panel regression is estimated per industry, and the standardization of $lGHG$ is also done per industry. We rely on Fama and French (1997) for the industry classification and constraint the estimation to industries with more than five firms. The regressions use controls ***CTRL-15***. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018. Results are sorted by decreasing industries' GHG emissions intensity.

Industry	$lGHG$	UMC	$lGHG \times UMC$
Utilities	0.00	-0.075^{***}	0.022
Steel Works Etc	0.093	-0.072	0.185
Chemicals	0.011	-0.035	0.028
Petroleum and Natural Gas	0.011	-0.045^{**}	0.033
Consumer Goods	0.008	0.024	0.013
Automobiles and Trucks	0.016	0.114^{***}	0.063
Transportation	-0.06	0.06^{**}	0.044
Machinery	-0.038^{**}	-0.015	-0.066^{*}
Meals Restaurants, Hotels, Motels	0.001	-0.005	-0.019
Business Supplies	-0.005	0.028	-0.05^{*}
Beer & Liquor	0.028	-0.036	0.017
Food Products	-0.007	0.026	-0.002
Electronic Equipment	0.023	-0.009	0.011
Computers	0.004	-0.016	-0.073^{**}
Pharmaceutical Products	-0.015	0.018	0.005
Wholesale	0.037	0.039^{*}	-0.04
Medical Equipment	0.024	-0.001	-0.012
Communication	-0.038	0.025	-0.007
Construction Materials	0.064^{**}	0.046	-0.258^{***}
Apparel	-0.011	0.025	0.022
Retail	0.02	0.004	-0.003
Business Services	-0.01	-0.004	0.005
Aircraft	-0.004	0.012	-0.025
Measuring and Control Equipment	-0.001	-0.045^{**}	0.028
Healthcare	0.033	-0.103^{**}	-0.225
Electrical Equipment	0.005	0.024	0.004
Banking	-0.002	0.04^{***}	-0.002
Trading	0.032^{*}	0.034^{*}	0.014
Insurance	0.014	0.02	-0.019
Construction	-0.013	0.053	0.003

Table 10: Panel regression results of daily firms' returns – Non-disclosure

This table reports the estimation results for the firm fixed-effect panel regression of daily stock returns on daily standardized logarithmic GHG emissions intensity ($lGHG$), daily unexpected changes in climate change concerns (UMC), their interaction ($lGHG{\times}UMC$), and two interactions terms with the undisclosure dummy variable UD ($lGHG{\times}UD$ and $lGHG{\times}UD{\times}UMC$); see model (9). UD takes a value of one when emissions data is not disclosed. The standardized GHG emissions intensity of firms that do not disclose their greenhouse gas emissions level is set to the average of the firm's industry. Standard errors of the estimators are reported in parentheses. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018.

	$lGHG$	$lGHG{\times}UD$	UMC	$lGHG{\times}UMC$	$lGHG{\times}UD{\times}UMC$
CTRL-1	0.006^{*}	0.026^{*}	-0.006	-0.027^{***}	-0.011
	(0.003)	(0.014)	(0.004)	(0.005)	(0.028)
CTRL-3	0.005	0.024^{*}	-0.009^{**}	-0.024^{***}	-0.015
	(0.003)	(0.014)	(0.004)	(0.005)	(0.027)
CTRL-6	0.005^{*}	0.022	-0.006	-0.023^{***}	-0.019
	(0.003)	(0.014)	(0.004)	(0.005)	(0.027)
CTRL-15	0.005^{*}	0.022	-0.006	-0.023^{***}	-0.019
	(0.003)	(0.014)	(0.004)	(0.005)	(0.027)

Table 11: Correlation matrix between daily aggregate and thematic UMC indices
This table reports the pairwise correlations between the daily aggregate and thematic UMC indices. Panel A reports the unconditional correlations while Panel B reports the correlations when daily aggregate UMC is highed than its 90th percentile (*i.e.*, high unexpected changes in climate change concerns). Themes are BI: Business Impact, EI: Environmental Impact, SD: Societal Debate, and R: Research.

Panel A: Unconditional correlations				
	BI	EI	SD	R
Aggregate	0.85	0.79	0.82	0.81
BI		0.47	0.66	0.57
EI			0.51	0.73
SD				0.58
Panel B: Correlations when aggregate UMC is high				
	BI	EI	SD	R
Aggregate	0.58	0.41	0.53	0.54
BI		−0.25	0.34	0.04
EI			−0.22	0.37
SD				0.03

Table 12: Interaction term estimates – Thematic and topical UMC indices
This table reports the interaction term ($lGHG \times UMC_k$) regression coefficient estimates of firm fixed-effect panel regression (9) on daily stock returns (left column "Daily"), and monthly capital-gain returns, cash flow news components, and discount rate components (right columns "Monthly") when using the aggregate, the thematic, and the topical UMC variables (reported in rows). The daily firm fixed-effect panel regressions use controls ***CTRL-15***. To deal with the low number of observations at the monthly frequency, the regressions use controls ***CTRL-6*** and the first three principal components (explaining 66% of the total variance) of the additional variables in ***CTRL-15***. The signs *, **, and *** indicate significant coefficients at the 10%, 5%, and 1% levels, respectively. The model is estimated with data from January 2010 to June 2018.

	$lGHG \times UMC_k$			
	Daily	Monthly		
	Return	Return	CF news	DR news
Aggregate *UMC*	-0.021^{***}	-0.376^{*}	0.23	-0.606^{**}
Theme 1: Business Impact	-0.028^{***}	-0.165	0.141	-0.306
Climate Summits	-0.028^{***}	-0.166	0.099	-0.265
Agreements/Actions	-0.021^{***}	-0.41^{*}	0.147	-0.557^{**}
Climate Legislation/Regulations	-0.026^{***}	-0.06	0.324	-0.385
Legal Actions	-0.013^{***}	0.143	-0.088	0.231
Renewable Energy	-0.02^{***}	0.171	0.123	0.048
Carbon Reduction Technologies	-0.015^{***}	0.368^{*}	0.203	0.165
Carbon Credits Market	-0.018^{***}	-0.51^{**}	-0.16	-0.35
Carbon Tax	-0.014^{***}	-0.553^{**}	-0.102	-0.45
Government Programs	-0.019^{***}	-0.193	0.059	-0.253
Corporations/Investments	-0.015^{***}	-0.048	-0.048	0.000
Car Industry	-0.001	-0.141	0.021	-0.162
Airline Industry	-0.012^{***}	-0.39^{**}	-0.059	-0.33^{*}
Theme 2: Environmental Impact	-0.005	-0.318	0.164	-0.481
Extreme Temperatures	0.003	0.059	0.113	-0.055
Food Shortage/Poverty	0.001	-0.247	0.091	-0.339^{*}
Hurricanes/Floods	-0.01^{***}	-0.065	0.097	-0.162
Glaciers/Ice Sheets	-0.016^{***}	-0.287	-0.064	-0.223
Ecosystems	0.004	-0.363^{*}	-0.149	-0.214
Forests	-0.006	0.135	-0.078	0.213
Water/Drought	-0.01^{**}	-0.376^{*}	-0.161	-0.215
Tourism	-0.025^{***}	0.106	0.109	-0.003
Arctic Wildlife	-0.002	-0.163	0.017	-0.18
Marine Wildlife	0.001	-0.191	0.047	-0.238
Agriculture Shifts	0.012^{**}	0.116	-0.041	0.157
Theme 3: Societal Debate	-0.023^{***}	-0.309^{*}	0.103	-0.412^{*}
Political Campaign	-0.024^{***}	-0.238^{*}	0.111	-0.35^{**}
Social Events	-0.01^{**}	-0.379^{*}	0.096	-0.475^{*}
Controversies	-0.022^{***}	-0.22	0.043	-0.263
Cities	-0.014^{***}	0.047	0.085	-0.038
Theme 4: Research	-0.01^{*}	-0.578^{**}	0.113	-0.692^{**}
Global Warming	-0.014^{**}	-0.586^{***}	0.182	-0.768^{***}
UN/IPCC Reports	-0.009^{*}	-0.382^{**}	0.087	-0.468^{*}
Scientific Studies	-0.002	-0.343	-0.024	-0.319

Figure 1: Correlation network of climate change topics
This figure displays the Spearman correlation network for the 30 climate change topics obtained with the correlated topic model. To keep the network readable, we display only correlations above 0.35. Each topic is assigned to a thematic cluster (Theme 1: Business Impact, Theme 2: Environmental Impact, Theme 3: Societal Debate, and Theme 4: Research).

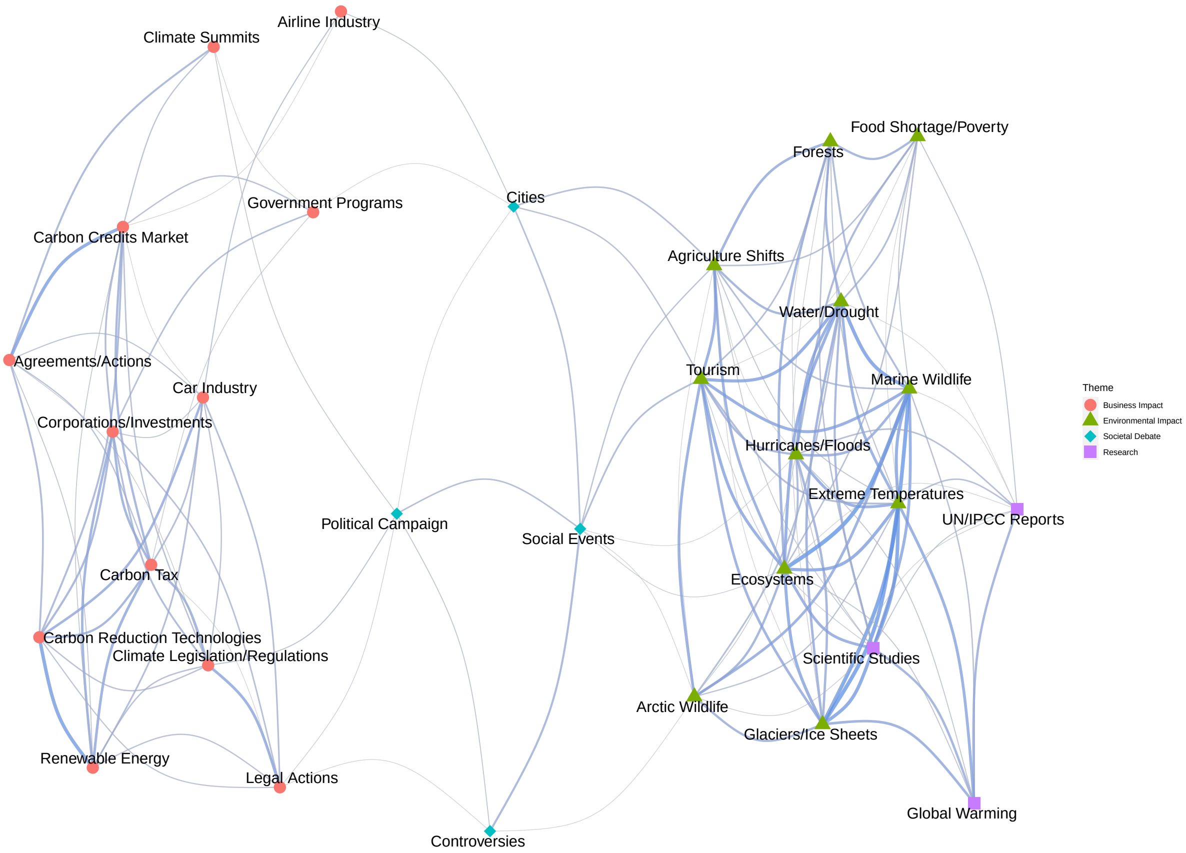


Figure 2: Number of article-equivalents by climate change theme
This figure displays the monthly number of article-equivalent publications for each theme from January 2010 to June 2018.

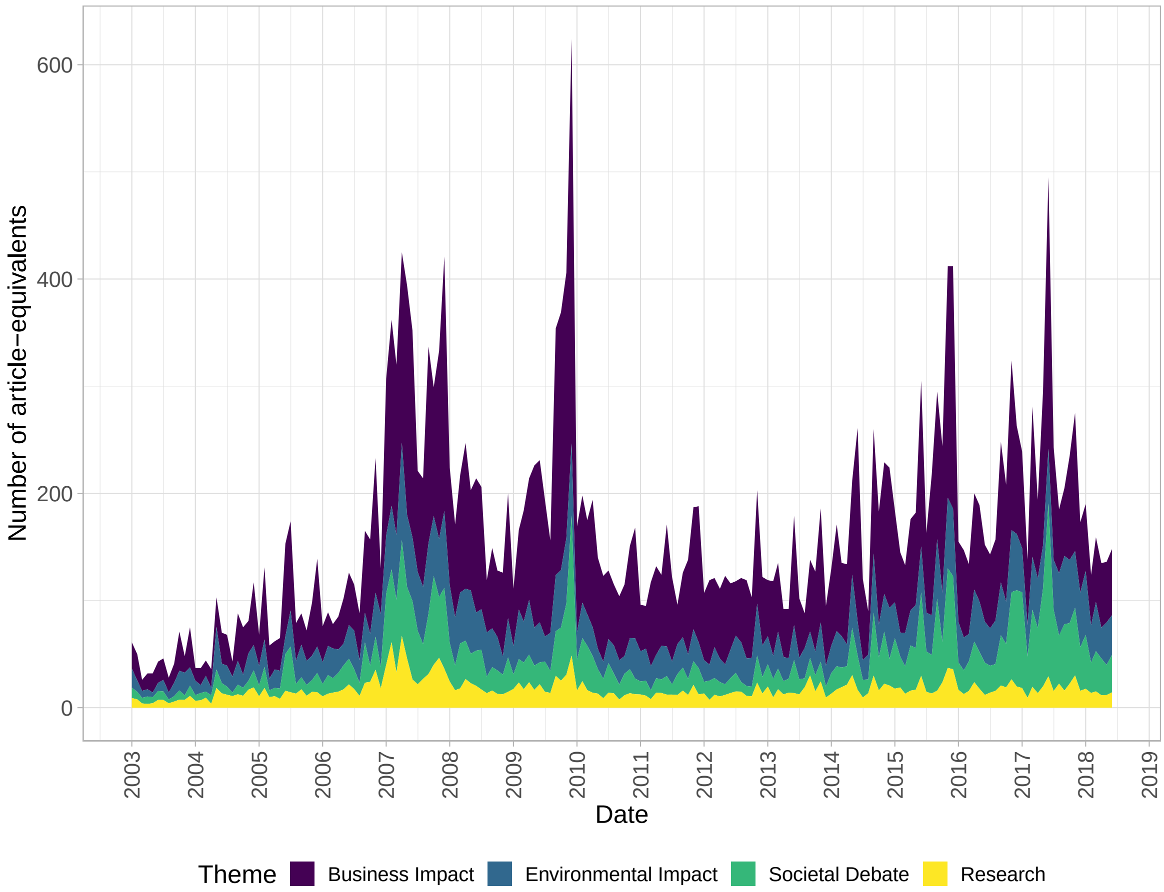


Figure 3: Media Climate Change Concerns index
This figure displays the daily MCCC index (gray points) together with its 30-day moving average (bold line) from January 2003 to June 2018. We also report several major events related to climate change (in boxes). The observations before January 1, 2010 (*i.e.*, at the left of the black dotted line) are considered forward-looking, since the data from that period is used to compute the source-specific standard deviation estimate required to normalize the source-specific indices before aggregation into the MCCC index. The observations from January 1, 2010, to the end of the time series (*i.e.*, at the right of the black dotted line) are not forward-looking and correspond to the period for our main analysis.

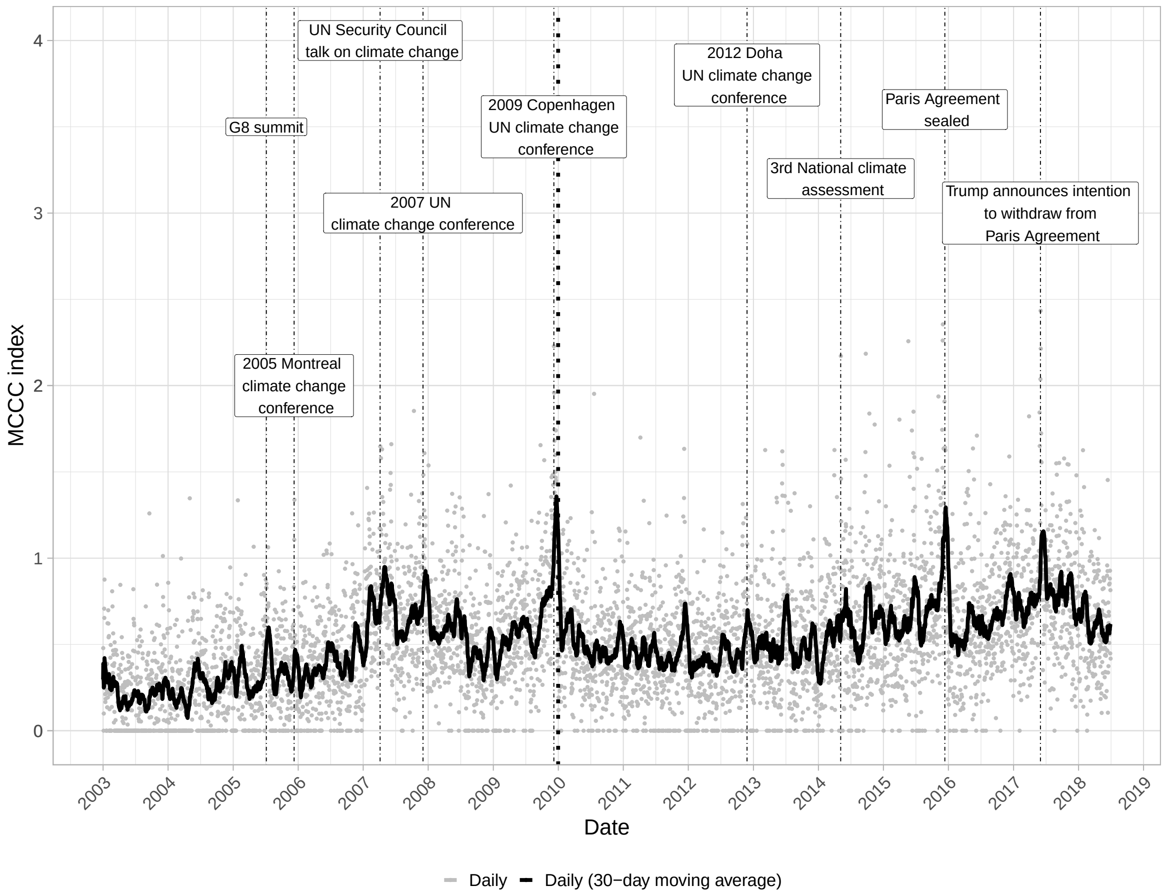
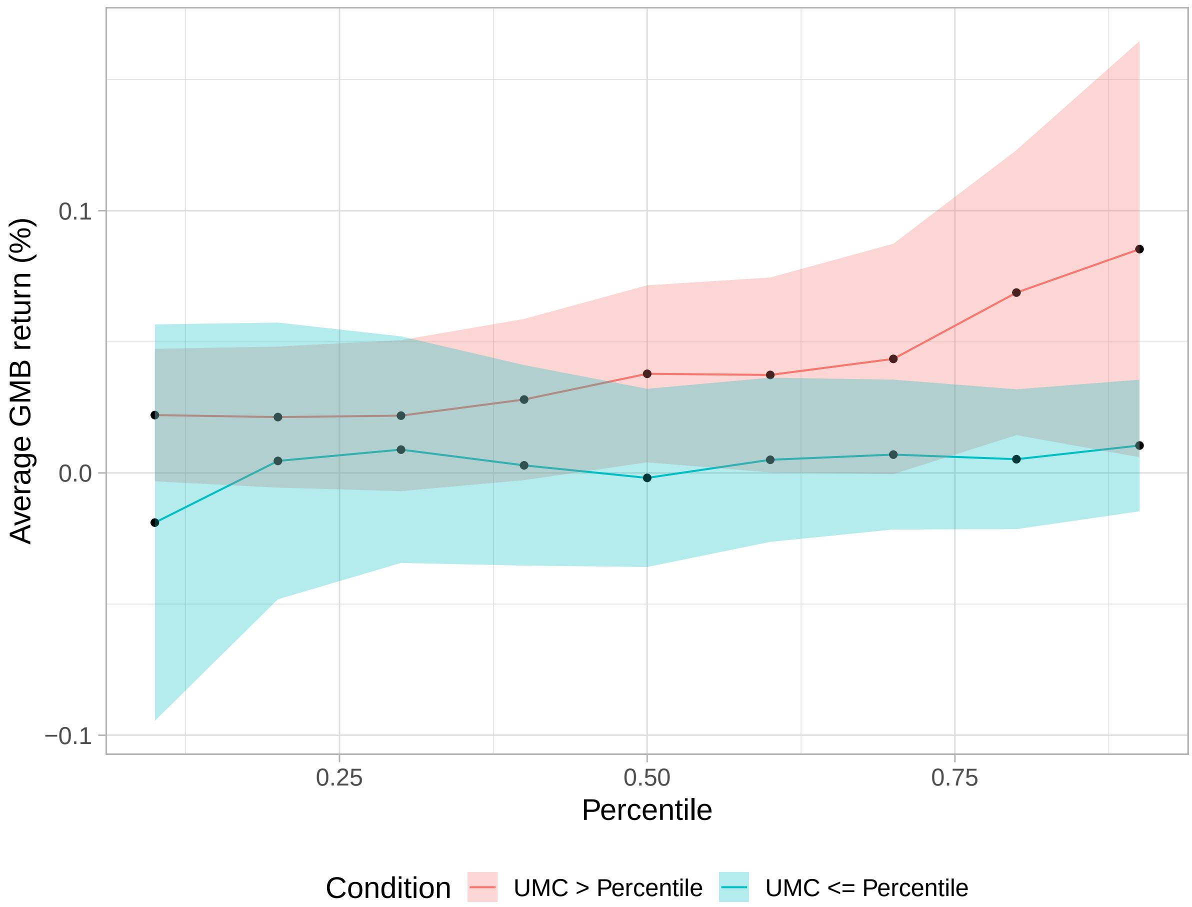


Figure 4: Green minus brown portfolio average return
This figure displays the average daily return of the GMB portfolio (vertical axis) conditional on *UMC* being above or below a specific threshold (horizontal axis). Thresholds are set as percentiles of *UMC*. The colored bands report the 95% confidence interval.



Appendix

A. Keywords used as negative filters to eliminate stock market news articles

In Table A.1, we report the keywords used to remove the news articles that can be related to the stock market (Panel A) and the number of news articles removed by the filters (Panel B for newswires and Panel C for newspapers). The list was compiled manually from a more extensive list of trigrams, bigrams, and unigrams extracted automatically from the news corpus.

Table A.1: Keywords used for the filters

This table reports the keywords used to remove the news articles that can be related to the stock market (Panel A) and the number of news articles removed by the filters (Panel B in newswires and Panel C in newspapers).

Panel A: Number of detected keywords used in the filters

Keyword	#	Keyword	#	Keyword	#
financial crisis	440	investment portfolio	28	share falling	2
market price	162	commodity markets	26	share fell	2
financial market	149	share prices	24	stock closed	2
green investment	144	shares fell	22	stock fall	2
green fund	141	boost investment	19	stock fell	2
capital market	127	the returns	17	stock were down	2
financial markets	118	stock index	16	market returns	1
market share	117	shares rose	15	markets closed	1
the return	113	green funds	14	share rises	1
capital markets	111	stocks fall	13	share was down	1
investment fund	108	shares fall	9	shares rises	1
stock exchange	103	drive investments	9	stock jumped	1
stock market	93	shares were down	8	stock return	1
market value	92	the performances	8	stock returns	1
market prices	84	boost investments	6	stocks indices	1
bullish	80	stocks fell	6	stocks moved	1
the performance	74	green stock	5	stocks price	1
no return	71	shares were up	5	stocks prices	1
investment funds	64	driving investments	4	stocks rally	1
stock price	60	green stocks	4	stocks rebound	1
share price	47	market shares	4		
nasdaq	46	share fall	4		
green investments	45	shares jumped	4		
drive investment	42	share rose	3		
new york stock exchange	37	shares closed	3		
stock prices	37	stock indices	3		
commodity market	37	stock rally	3		
the crash	32	stock rose	3		
bearish	31	growth stock	2		

Panel B: Number of articles removed by the filters in the newswires

Source	#
Reuters News	1,087
Associated Press Newswires	480

Panel C: Number of articles removed by the filters in the newspapers

Source	#
New York Times	224
Wall Street Journal	118
Washington Post	92
Houston Chronicle	61
Los Angeles Times	61
Chicago Tribune	26
USA Today	10
Arizona Republic	8
New York Post	3
New York Daily News	2

B. Topic modeling and climate change themes construction

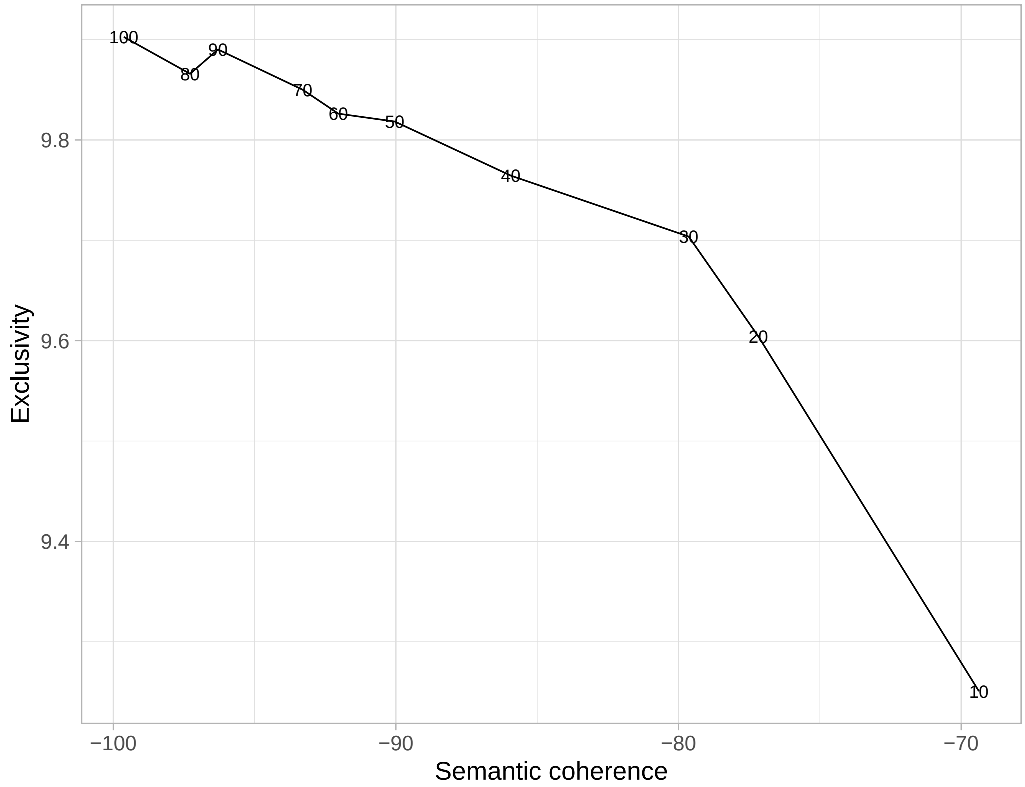
We follow Martin and Johnson (2015) and only use nouns (including proper nouns) in our vocabulary. Moreover, following Hansen, McMahon, and Prat (2018), we also identify collocation, which is a sequence of words (in our case, a sequence of nouns) that have a specific meaning. We only identify two-word collocations. We then calculate the number of times these collocations appear and create a single term for those that appear more than 100 times in the climate change corpus. An example of such collocation is "climate change."

Next, we lemmatize every standalone word (*i.e.*, excluding collocations). We use vocabulary and morphological analysis of words to remove inflectional endings and transform them into their base or dictionary form. This step helps delete non-informative variations of words. We then remove rare words (*i.e.*, words that appear in less than 0.05% of the texts in the corpus) and common words (*i.e.*, words that appear in more than 50% of the texts in the corpus).

We use two common metrics to determine the optimal number K of topics in our corpus: (i) the semantic coherence and (ii) the exclusivity. Semantic coherence is maximized when the most probable words in a given topic frequently co-occur together, and it is a metric that correlates well with the human judgment of topic quality (Mimno et al., 2011). However, having high semantic coherence is relatively easy when very common words dominate only a few topics. Roberts, Stewart, and Tingley (2014) propose a complementary metric called exclusivity. It measures the exclusiveness of the words that make up a topic. The objective is to find the number of topics that balances semantic coherence and exclusivity. To do so, we calibrate the CTM with $K \in \{10, 20, \ldots, 100\}$.

In Figure B.1, we display the semantic coherence and exclusivity for each of those models. We can observe that a very high (low) number of topics leads to a high exclusivity but low (high) semantic coherence. It seems that 30 topics appears to be around the point where lowering (increasing) the number of topics would significantly reduce exclusivity (semantic coherence). Thus, we choose $K = 30$ as the optimal number of topics for the CTM estimation.

Figure B.1: Semantic coherence vs. exclusivity
This figure displays the values of semantic coherence (horizontal axis) and exclusivity (vertical axis) for various numbers of topics ($K \in \{10, 20, \ldots, 100\}$) is the CTM.



Next, we manually label the 30 topics in our corpus. We proceed by: (i) looking at the ten most probable words for each topic and (ii) looking at the content of the articles with the largest topic prevalence. In Table 3, we report the ten highest-probability words for each topic in our list.

Finally, identification of the themes is based on the correlation of the topics' prevalences. Manual inspection of the topics via hierarchical clustering (average distance) leads to four themes that have an intuitive meaning, as also shown in the correlation network in Figure 1.

C. Sets of variables

In Table C.1, we list the variables used in the ARX model, the portfolio regressions, and the firm fixed-effect panel regressions.

Table C.1: Sets of variables

	ARX	***CTRL-1***	***CTRL-3***	***CTRL-6***	***CTRL-15***
MKT	X	X	X	X	X
HML	X		X	X	X
SMB	X		X	X	X
CMA	X			X	X
RMW	X			X	X
MOM	X			X	X
WTI	X				X
NG	X				X
PROP	X				X
EPU	X				X
VIX	X				X
TED	X				X
TERM	X				X
DFLT	X				X
FTS	X				X
GB	X				
GMB	X				

The variables are:

1. *MKT*: the daily excess market return;
2. *HML*: the daily high-minus-low factor of Fama and French (1992);
3. *SMB*: the daily small-minus-big factor of Fama and French (1992);
4. *CMA*: the daily conservative-minus-aggressive factor of Fama and French (2015);
5. *RMW*: the daily robust-minus-weak factor of Fama and French (2015);
6. *MOM*: the daily momentum factor of Carhart (1997);
7. *WTI*: the daily crude oil return (West Texas Intermediate crude oil price, DCOILWTICO);
8. *NG*: the daily natural gas return (Henry hub natural gas spot price, DHHNGSP);
9. *PROP*: the daily propane return (Mont Belvieu Texas price, DPROPANEMBTX);
10. *EPU*: the daily U.S. economic policy uncertainty index of Baker, Bloom, and Davis (2016);
11. *VIX*: the daily CBEO volatility index;
12. *TED*: the daily TED spread;
13. *TERM*: the daily term factor of Fung and Hsieh (2004);
14. *DFLT*: the daily default factors of Fung and Hsieh (2004);
15. *FTS*: the daily flight-to-safety index of Baele et al. (2020);
16. *GB*: the daily returns of a green bonds' portfolio;
17. *GMB*: the daily return of the green-minus-brown stocks' portfolio (25-75th percentile).

The risk-free rate and asset-pricing variables are retrieved from Kenneth French's website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html. The rates (used to construct the term and default factors) as well as the TED, EPU, VIX, and energy-related variables are retrieved from the FRED library at https://fred.stlouisfed.org. The flight-to-safety index is kindly provided by the authors. For the green bond, we use the S&P green bond U.S. Dollar Select Index retrieved from Datastream. Finally, returns of the 25-75th percentile green-minus-brown stocks' portfolio are obtained following the steps in Section 4.2.

D. ARX model estimation

In Table D.1, we report statistics about the ARX coefficients estimated over the 3,103 rolling windows (each rolling window is of size 1,000). Specifically, we report the mean and the 25-75th percentiles of the estimates and the percentage of time an estimate is significantly positive or negative, at the 5% significance level.

We note the autocorrelation in the *MCCC* series. Lagged variables that are significant more than 25% of the time are *GB*, *WTI*, and *NG*. We also note significant coefficients for *TERM*, *CMA*, and *RMW*.

Table D.1: ARX coefficients

This table reports statistics about the ARX coefficients estimated over the 3,103 rolling windows (each rolling window is of size 1,000). We report the mean and the 25-75th percentiles of the estimates and the percentage of time an estimate is significantly positive or negative, at the 5% significance level. *MCCC* is the autoregressive coefficient; see the Appendix, Section C, for the definition of the other variables. Note that for *GMB* and *GB* we have 2,103 and 2,498 estimations, respectively.

				5% level	
	Mean	25th	75th	%+	%−
Constant	0.417	0.360	0.458	100.00	0.00
MCCC	0.312	0.276	0.359	100.00	0.00
MKT	-0.003	-0.013	0.008	0.00	0.00
HML	0.010	-0.003	0.023	0.00	0.00
SMB	0.007	-0.001	0.013	0.00	0.00
CMA	0.024	-0.026	0.067	19.30	0.06
RMW	0.053	0.041	0.069	15.63	0.00
MOM	0.011	0.004	0.017	0.00	0.00
WTI	0.462	-0.226	0.889	40.83	10.12
NG	0.265	0.201	0.328	30.84	0.00
PROP	0.029	-0.172	0.255	0.26	1.55
EPU	-0.010	-0.032	-0.002	5.06	16.37
VIX	0.002	-0.001	0.004	0.00	1.48
TED	-0.137	-0.251	-0.027	3.25	50.37
TERM	0.226	0.034	0.430	24.04	0.00
DFLT	0.121	0.039	0.192	7.03	0.00
FTS	0.001	-0.063	0.034	3.67	3.80
GB	0.966	0.030	2.023	49.80	15.53
GMB	0.032	0.027	0.045	7.56	0.00

E. Green/brown classification and GHG emissions intensity distribution over time

In Table E.1, we report the yearly transition matrices of moving from one category (green, neutral, or brown) to another category (green, neutral, or brown). Panel A displays results for the baseline 25-75th percentiles GHG emissions intensity split. Panel B (Panel C) reports results for the 10-90th (40-60th) percentiles GHG emissions intensity split.

Table E.1: Yeary transition matrices of moving from one category to another one

	Green	Neutral	Brown
Panel A: 25-75th percentiles			
Green	88.98	3.62	0.00
Neutral	10.86	91.34	9.15
Brown	0.16	5.04	90.85
Panel B: 10-90th percentiles			
	Green	Neutral	Brown
Green	83.19	1.23	0.00
Neutral	16.81	98.03	6.79
Brown	0.00	0.74	93.21
Panel C: 40-60th percentiles			
	Green	Neutral	Brown
Green	91.54	10.7	0.98
Neutral	6.14	78.4	5.56
Brown	2.32	10.89	93.46

In Figure E.1, we display the evolution of the logarithmic GHG emissions intensity percentiles.

Figure E.1: Evolution of the logarithmic GHG emissions intensity percentiles

F. Implied cost of capital, cash flow news, and discount rate news estimation

We use the return decomposition framework proposed by Chen, Da, and Zhao (2013) to decompose the monthly capital gain return into their cash flow and discount rate news component. The methodology requires an estimate of the implied cost of capital. We estimate it with the model by Gebhardt, Lee, and Swaminathan (2001). Below, we describe both approches. For simplicity, we omit the firm subscript in our exposition.

F.1. Implied cost of capital

Under the valuation model by Gebhardt, Lee, and Swaminathan (2001), the stock price P_τ of a firm in month τ given cash flow and book value forecasts embedded in vector $\mathbf{c}_\tau$ (available in month τ), and the cost of capital, q_τ (expressed annually), is defined as:

$$P_\tau = f(\mathbf{c}_\tau, q_\tau) = B_\tau + \sum_{y=1}^{11} \frac{FROE_{\tau,y} - q_\tau}{(1+q_\tau)^y} B_{\tau,y-1} + \frac{FROE_{\tau,12} - q_\tau}{q_\tau (1+q_\tau)^{11}} B_{\tau,11}\,, \tag{F.1}$$

where:

- B_τ is the book value per share in the latest financial statement available in month τ. The data is retrieved from Compustat.
- $FROE_{\tau,y}$ is the return-on-equity forecast for yearly period y at time τ. For the first two yearly periods ($y = 1, 2$), $FROE_{\tau,y} = FEPS_{\tau,y}/B_{\tau,y-1}$, where $FEPS_{\tau,y}$ is the IBES average analysts' earnings-per-share forecast for yearly period y and $B_{\tau,y-1}$ is the book value per share for yearly period $y-1$, both in month τ. For the third yearly period ($y = 3$), we set $FEPS_{\tau,3}$ to $FEPS_{\tau,2}(1 + LTG_\tau)$, where LTG_τ is the average analysts' estimate of the long-term earnings growth rate of the firm in month τ. For the fourth yearly period and beyond, $FEPS_{\tau,y} = FEPS_{\tau,y-1}[1 + LTG_\tau + \frac{(y-3)}{9}(g_\tau - LTG_\tau)]$, where g_τ is the cross-sectional average of LTG_τ across all firms in the sample. Thus, for yearly periods $y = 4, \ldots, 12$, the long-term earnings growth linearly converges to the cross-sectional long-term earnings growth average in month τ.[33]
- $B_{\tau,y} = B_{\tau,y-1} + FEPS_{\tau,y} - FDPS_{\tau,y}$, where $FDPS_{\tau,y} = kFEPS_{\tau,y}$. The factor k is defined as the latest dividend payout ratio (*i.e.*, common dividend paid divided by net income). For firms experiencing negative earnings, we divide the dividends paid by (0.06 $\times$ total assets) to derive the payout ratio. Payout lesser (greater) than zero (one) are assigned a value of zero (one).

For each month and firm in our sample, we use Equation (F.1) to estimate q_τ. This implied cost of capital is an input in the decomposition approach, as shown below.

[33] In the spirit of Chen, Da, and Zhao (2013), we cross-sectionally winsorize $FEPS_{\tau,1}$, $FEPS_{\tau,2}$, and LTG_τ at the 1% and 99% level each month. In addition, we process earnings forecasts following Da, Liu, and Schaumburg (2014, Footnote 7).

F.2. Cash flow and discount rate news

We follow Chen, Da, and Zhao (2013) to decompose capital gain returns into cash flow (CF) and discount rate (DR) news components. The capital gain return at time τ is defined as:

$$retx_\tau = \frac{P_\tau - P_{\tau-1}}{P_{\tau-1}} = \frac{f(\mathbf{c}_\tau, q_\tau) - f(\mathbf{c}_{\tau-1}, q_{\tau-1})}{P_{\tau-1}} = CF_\tau + DR_\tau \,, \tag{F.2}$$

where P_τ is the observed stock price (adjusted for stock splits) in month τ, and:

$$CF_\tau = \left(\frac{f(\mathbf{c}_\tau, q_\tau) - f(\mathbf{c}_{\tau-1}, q_\tau)}{P_{\tau-1}} + \frac{f(\mathbf{c}_\tau, q_{\tau-1}) - f(\mathbf{c}_{\tau-1}, q_{\tau-1})}{P_{\tau-1}} \right) /2 \,, \tag{F.3}$$

$$DR_\tau = \left(\frac{f(\mathbf{c}_{\tau-1}, q_\tau) - f(\mathbf{c}_{\tau-1}, q_{\tau-1})}{P_{\tau-1}} + \frac{f(\mathbf{c}_\tau, q_\tau) - f(\mathbf{c}_\tau, q_{\tau-1})}{P_{\tau-1}} \right) /2 \,. \tag{F.4}$$

The function $f(\cdot, \cdot)$ is defined in Equation (F.1).